Advances in Intelligent Systems and Computing 597

Neville A. Stanton *Editor*

Advances in Human Aspects of Transportation

Proceedings of the AHFE 2017 International Conference on Human Factors in Transportation, July 17–21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA

Springer

Advances in Intelligent Systems and Computing

Volume 597

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Chairman

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
e-mail: nikhil@isical.ac.in

Members

Rafael Bello Perez, Universidad Central "Marta Abreu" de Las Villas, Santa Clara, Cuba
e-mail: rbellop@uclv.edu.cu

Emilio S. Corchado, University of Salamanca, Salamanca, Spain
e-mail: escorchado@usal.es

Hani Hagras, University of Essex, Colchester, UK
e-mail: hani@essex.ac.uk

László T. Kóczy, Széchenyi István University, Győr, Hungary
e-mail: koczy@sze.hu

Vladik Kreinovich, University of Texas at El Paso, El Paso, USA
e-mail: vladik@utep.edu

Chin-Teng Lin, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ctlin@mail.nctu.edu.tw

Jie Lu, University of Technology, Sydney, Australia
e-mail: Jie.Lu@uts.edu.au

Patricia Melin, Tijuana Institute of Technology, Tijuana, Mexico
e-mail: epmelin@hafsamx.org

Nadia Nedjah, State University of Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: nadia@eng.uerj.br

Ngoc Thanh Nguyen, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Ngoc-Thanh.Nguyen@pwr.edu.pl

Jun Wang, The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: jwang@mae.cuhk.edu.hk

More information about this series at http://www.springer.com/series/11156

Neville A. Stanton
Editor

Advances in Human Aspects of Transportation

Proceedings of the AHFE 2017 International Conference on Human Factors in Transportation, July 17–21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA

Springer

Editor
Neville A. Stanton
Human Factors Engineering, Transportation Research Group
Boldrewood Innovation Campus, University of Southampton
Southampton
UK

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-60440-4 ISBN 978-3-319-60441-1 (eBook)
DOI 10.1007/978-3-319-60441-1

Library of Congress Control Number: 2017943063

© Springer International Publishing AG 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Advances in Human Factors and Ergonomics 2017

AHFE 2017 Series Editors

Tareq Z. Ahram, Florida, USA
Waldemar Karwowski, Florida, USA

8th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences

Proceedings of the AHFE 2017 International Conference on Human Factors in Transportation, July 17–21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA

Advances in Affective and Pleasurable Design	*WonJoon Chung and Cliff (Sungsoo) Shin*
Advances in Neuroergonomics and Cognitive Engineering	*Carryl Baldwin*
Advances in Design for Inclusion	*Giuseppe Di Bucchianico and Pete Kercher*
Advances in Ergonomics in Design	*Francisco Rebelo and Marcelo Soares*
Advances in Human Error, Reliability, Resilience, and Performance	*Ronald L. Boring*
Advances in Human Factors and Ergonomics in Healthcare and Medical Devices	*Vincent G. Duffy and Nancy Lightner*
Advances in Human Factors in Simulation and Modeling	*Daniel N. Cassenti*
Advances in Human Factors and System Interactions	*Isabel L. Nunes*
Advances in Human Factors in Cybersecurity	*Denise Nicholson*
Advances in Human Factors, Business Management and Leadership	*Jussi Kantola, Tibor Barath and Salman Nazir*
Advances in Human Factors in Robots and Unmanned Systems	*Jessie Chen*
Advances in Human Factors in Training, Education, and Learning Sciences	*Terence Andre*
Advances in Human Aspects of Transportation	*Neville A. Stanton*

(continued)

(continued)

Advances in Human Factors, Software, and Systems Engineering	*Tareq Z. Ahram and Waldemar Karwowski*
Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries	*Paul Fechtelkotter and Michael Legatt*
Advances in Human Factors, Sustainable Urban Planning and Infrastructure	*Jerzy Charytonowicz*
Advances in the Human Side of Service Engineering	*Louis E. Freund and Wojciech Cellary*
Advances in Physical Ergonomics and Human Factors	*Ravindra Goonetilleke and Waldemar Karwowski*
Advances in Human Factors in Sports, Injury Prevention and Outdoor Recreation	*Tareq Z. Ahram*
Advances in Safety Management and Human Factors	*Pedro Arezes*
Advances in Social & Occupational Ergonomics	*Richard Goossens*
Advances in Ergonomics of Manufacturing: Managing the Enterprise of the Future	*Stefan Trzcielinski*
Advances in Usability and User Experience	*Tareq Ahram and Christianne Falcão*
Advances in Human Factors in Wearable Technologies and Game Design	*Tareq Ahram and Christianne Falcão*
Advances in Communication of Design	*Amic G. Ho*
Advances in Cross-Cultural Decision Making	*Mark Hoffman*

Preface

Human Factors and Ergonomics have made a considerable contribution to the research, design, development, operation, and analysis of transportation systems which includes road and rail vehicles and their complementary infrastructure, aviation, and maritime transportation. This book presents recent advances in the Human Factors Aspects of Transportation. These advances include accident analysis, automation of vehicles, comfort, distraction of drivers (understanding of distraction and how to avoid it), environmental concerns, in-vehicle systems design, intelligent transport systems, methodological developments, new systems and technology, observational and case studies, safety, situation awareness, skill development and training, warnings, and workload.

This book brings together the most recent human factors work in the transportation domain, including empirical research, human performance, and other types of modeling, analysis, and development. The issues faced by engineers, scientists, and other practitioners of human factors in transportation research are becoming more challenging and more critical.

The common theme across these sections is that they deal with the intersection of the human and the system. Moreover, many of the chapter topics cross section boundaries, for instance by focusing on function allocation in NextGen or on the safety benefits of a tower controller tool. This is in keeping with the systemic nature of the problems faced by human factors experts in rail and road, aviation and aerospace, and maritime research—it is becoming increasingly important to view problems not as isolated issues that can be extracted from the system environment, but as embedded issues that can only be understood as a part of an overall system.

In keeping with a system that is vast in its scope and reach, the chapters in this book cover a wide range of topics. The chapters are organized into 14 sections:

Section 1: Aviation: Human Factors Safety in Aviation and Aerospace
Section 2: : Maritime- Human Factors in Maritime
Section 3: Road and Rail: Driver Safety
Section 4: Road and Rail: Human-Machine Interaction
Section 5: Road and Rail: Automation

Section 6: Road and Rail: Infrastructure and Road Safety
Section 7: Road and Rail: Vulnerable Road Users
Section 8: Road and Rail: Driving Simulation and Test Track
Section 9: Road and Rail: Eco-Driving
Section 10: Road and Rail: Rail Passengers
Section 11: Road and Rail: Public Transport
Section 12: Road and Rail: Driving Automation and Human Factors Issues
Section 13: Road and Rail: Rail Safety and Design
Section 14: Road and Rail: Transportation Logistics

This book will be of interest and use to transportation professionals who work in the road and rail, aviation and aerospace, and maritime domains as it reflects some of the latest Human Factors and Ergonomics thinking and practice. It should also be of interest to students and researchers in these fields, to help stimulate research questions and ideas. It is my hope that the ideas and studies reported within this book will help to produce safer, more efficient and effective transportation systems in the future.

We are grateful to the Scientific Advisory Board which has helped elicit the contributions and develop the themes in this book. These people are experts and academic leaders in their respective fields, and their help is very much appreciated, especially as they gave their time to the project.

Road and Rail

Giles Balbinotti, Brazil
Klaus Bengler, Germany
Gary Burnett, UK
Peter Chapman, UK
Fang Chen, Sweden
Denis Coelho, Portugal
Benjamin Colucci Rios, Puerto Rico
Guillaume Craveur, France
Laurel Dickson-Bull, USA
L. Dorn, UK
Ian Glendon, Australia
Iwona Grabarek, Poland
R. Happee, Netherlands
S. Jamson, UK
Dave Kaber, USA
Josef Krems, Germany
Mike Lenné, Australia
Franck Mars, France
D. McAvoy, USA
Ann Mills, UK

Ralf Risser, Austria
Paul Salmon, Australia
Shafiq ur Rehman, Sweden
Didier Valdes Diaz, Puerto Rico
Guy Walker, Scotland
Kristie Young, Australia

Aviation and Aerospace

Marcus Biella, Germany
Clark Borst, The Netherlands
Tamsyn Edwards, USA
Michael Feary, USA
Andreas Haslbeck, Germany
Becky Hooey, USA
David Kaber, USA
Kara Latorella, USA
Arnab Majumdar, UK
Lynne Martin, USA
Joey Mercer, USA
Max Mulder, The Netherlands
Jon Lars Syversen, Norway
Savvy Verma, USA
Kim Vu, USA

Maritime

Ahmet Dursun Alkan, Turkey
David Andrews, UK
Giuseppe Di Bucchianico, Italy
Dawn Gray, USA
Marc Grootjen, The Netherlands
Thomas Koester, Denmark
Scott Netson MacKinnon, Canada
Massimo Musio Sale, Italy
Stella (Styliani) Parisi, Germany
Gesa Praetorius, Sweden
Andrea Ratti, Italy
Andrea Vallicelli, Italy

July 2017 Neville A. Stanton

Contents

Aviation: Human Factors Safety in Aviation and Aerospace

Influencing Factors on Error Reporting in Aviation - A Scenario-Based Approach 3
Sebastian Sieberichs and Annette Kluge

Quantifying Pilot Contribution to Flight Safety During Hydraulic Systems Failure 15
Lynda J. Kramer, Timothy J. Etherington, Randall E. Bailey, and Kellie D. Kennedy

Perceptions and Affective Responses to Alternative Risk-Based Airport Security 27
Kenneth D. Nguyen and Richard S. John

Autonomous Stall Recovery Dynamics as a Prevention Tool for General Aviation Loss of Control 38
Tiziano Bernard, Lucas Stephane, and Guy A. Boy

Planning of General Aviation Pilots Using Interviews 48
Elena Psyllou, Arnab Majumdar, and Washington Ochieng

HMIM: A Method to Study the Information and Control Flow Exchange in the Flight Deck 58
Yiyuan Zheng, Yanyu Lu, Yuwen Jie, and Shan Fu

Analyzing Positive and Negative Effects of Salience in Air Traffic Control Tasks 70
Haruka Yoshida, Hisae Aoyama, Satoru Inoue, Taro Kanno, and Kazuo Furuta

Engine Failure Induced Task Load Transient for Simulation Based Certification Aiding for Aircraft 79
Jens Alfredson, Luís G. Trabasso, Niklas Blomstrand, Maria Eckerberg, Linda Klamer, Johanna Ledin, Jasmine Tarander, and Magnus Bang

An Analysis of Human Factor Aspects in Operational Fuel Saving 87
Daniel Vogel, Ivan Sikora, and Hans-Joachim Ruff-Stahl

Transition from Conventionally to Remotely Piloted Aircraft – Investigation of Possible Impacts on Function Allocation and Information Accessibility Using Cognitive Work Analysis Methods 96
Max Friedrich, Anne Papenfuß, and Andreas Hasselberg

Workload and the En Route Controller – An Overblown Issue? 108
Richard W. Rohde

Data-Driven Pilot Behavior Modeling Applied to an Aircraft Offset Landing Task 117
Felipe M.S. Turetta, Helon Vicente Hultmann Ayala, Luís G. Trabasso, Leandro S. Coelho, and Jens Alfredson

Motion and Time Study on Space Maintenance Mission 128
Hanzhao Qiu, Weining Fang, Li Wang, Bo Wang, and Beiyuan Guo

Air Traffic Controller Additional Work Load as a Result of Aircraft Dynamic Separations 139
Julio Roa

Habitability Issues in Long Duration Space Missions Far from Earth 145
Giorgio Musso, Simona Ferraris, Franco Fenoglio, Antonio Zafarana, Adriana Salatino, and Raffaella Ricci

Maritime

Identifying Gaps, Opportunities and User Needs for Future E-navigation Technology and Information Exchange 157
Nicole A. Costa, Monica Lundh, and Scott N. MacKinnon

Inshore Navigation's Yacht: Study Cases of Sustainable and Ergonomic Yacht Design 170
Massimo Di Nicolantonio

"A-Shaped" Mast Sailing Yacht: Technological Solution for Easy Navigation 182
Jessica Lagatta and Andrea Vallicelli

Assessing Sonar and Target Motion Analysis Stations in a Submarine Control Room Using Cognitive Work Analysis 191
Daniel Fay, Neville A. Stanton, and Aaron P.J. Roberts

Visualization in Maritime Navigation: A Critical Review 199
Victor Conceição, Maria Beatriz Carmo, Joakim Dahlman, and Ana Ferreira

Gaps Between Users and Designers: A Usability Study About a Tablet-Based Application Used on Ship Bridges 213
Yemao Man, Margareta Lützhöft, Nicole A. Costa, Monica Lundh, and Scott N. MacKinnon

Road and Rail: Driver Safety

ATV Safety in Agriculture: Injury, Illness, Analysis and Interventions 227
David Gilkey, John Rosecrance, Hali Neves, and Elise Lagerstrom

Study of Spain's Vehicle Fleet Safety Level Based on the Results of Euro NCAP Tests 234
Oriol Vaquer, Núria Parera, and Alba Fornells

Effects of Actigraphically Acquired Sleep Quality on Driving Outcomes in Obstructive Sleep Apnea Patients and Control Drivers: A Naturalistic Study 242
Nazan Aksan, Robert Marini, Jon Tippin, Jeffrey Dawson, and Matthew Rizzo

Integrating Traffic Safety in Vehicle Routing Solution 251
Iyad Sahnoon, Mohamed Shawky, and Abdulla Al-Ghafli

Structural Study of the Side Impact Test with a Fully-Instrumented Pole 264
Jacob Sanchez, Melcior Malivern, Núria Parera, and Fernando Bernal

Neurobehavioural Evaluation of Rehabilitation Programs for Dangerous Drivers 275
Pavel Řezáč, Veronika Kurečková, Petr Zámečník, Daniel Shaw, Milan Brázdil, Kristína Czekóová, Beáta Špiláková, and Miguel Salazar

Different Learning Conditions and Their Impact on Hazard Perception Training 282
Jana Hilz, Sarah Malone, and Roland Brünken

Effects of Using a Cell Phone on Gaze Movements During Simulated Car Driving: Hand-Held and Hands-Free Conditions 289
Bethânya Graick Carizio, Gisele Chiozi Gotardi, Juliana Cristina de Ângelo, Paula Fávaro Polastri, Fábio Augusto Barbieri, Luis Carlos Paschoarelli, and Sérgio Tosi Rodrigues

Risky Behavior and the Attitudes Towards the Police Safety Operations in the Czech Republic 300
Petr Zámečník, Veronika Kurečková, Pavel Řezáč, and Vít Gabrhel

Does the Familiarity of Road Regulation Contribute to Driving Violation? A Simulated Study on Familiar and Unfamiliar Road Intersections among Young Chinese Drivers 307
Siyuan Huang, Daniele Ruscio, Dedy Ariansyah, Jun Yi, and Monica Bordegoni

Research on Blinking-Luminescence Travel Support for Visually Impaired Persons 319
Tomohiro Okada, Takashi Miyamoto, Shin-ichi Ito, Katsuya Sato, Norihiro Ikeda, Osamu Sueda, and Shoichiro Fujisawa

An Emerging Framework to Inform Effective Design of Human-Machine Interfaces for Older Adults Using Connected Autonomous Vehicles 325
Phillip L. Morgan, Alexandra Voinescu, Craig Williams, Praminda Caleb-Solly, Chris Alford, Ian Shergold, Graham Parkhurst, and Anthony Pipe

Road and Rail: Human-Machine Interaction

Identifying Difference Thresholds of Haptics and Acoustics of Control Devices 337
Tobias Glohr, Andreas Zimmermann, and Thomas Maier

Digital Touchpoints in Campus Slow Traffic Service System 349
Jintian Shi and Ke Ma

Cognitive Vehicle Design Guided by Human Factors and Supported by Bayesian Artificial Intelligence 362
Ata Khan

Assessing In-Vehicle Secondary Tasks with the NHTSA Visual-Manual Guidelines Occlusion Method 373
Mikael Ljung Aust, Robert Broström, and Sofia Lindvall

The Benefit of Touchless Gesture Control: An Empirical Evaluation of Commercial Vehicle-Related Use Cases 383
Michael Stecher, Britta Michel, and Andreas Zimmermann

How Does Awareness Affect Performance in an Automotive Dual Task Condition? 395
Lee Skrypchuk, Patrick Langdon, Alex Mouzakitis, and P. John Clarkson

Intelligent Driving Interface Layout and Design Research 407
Linghua Ran, Hong Luo, Yunhao Yan, Mengli Yu, and Xin Zhang

Human-Machine Interface Design Development for Connected and Cooperative Vehicle Features 415
William Payre and Cyriel Diels

How Does Eye-Gaze Relate to Gesture Movement in an Automotive Pointing Task? 423
Bashar I. Ahmad, Patrick M. Langdon, Lee Skrypchuk, and Simon J. Godsill

Determination of Auditory Stimuli for the Auditory Variant of the Detection Response Task Method 435
Kristina Stojmenova and Jaka Sodnik

Perceptions Towards the Use of Intelligent Transport System Technologies by Earthquake Victims 444
Izza Anwer and Susan Grant-Muller

Blind Driving by Means of a Steering-Based Predictor Algorithm 457
Pavlo Bazilinskyy, Charles Beaumont, Xander van der Geest, Reinier de Jonge, Koen van der Kroft, and Joost de Winter

Using Interior Design Principles to Improve the User's Perception of Vehicle Interiors: A Study on Visual Parameters 467
Susan Shaw, Claudia Escobar, and Shawn Lee

Possibility of Automobile Seat Evaluation with Seat Fidgets and Movements 474
Junya Tatsuno, Setsuo Maeda, George Sammonds, and Neil Mansfield

A Study on the Relationship Between Passengers' Visual Searching Efficiency in Bus-Waiting Process and the Elements of Bus Front Face Design 484
Tao Xiong, Jianxin Cheng, and Tengye Li

Vision and Driving Support for Shielded Vehicles – Implementation and Test of an Electronic Vision Replacement System with Augmented Reality 490
Daniel López, Marcel C.A. Baltzer, Christian Lassen, and Frank Flemisch

Measuring the Cognitive Demands of In-vehicle Dashboard and Centre Console Tasks 503
Madeleine Conaghan, Ian Colwill, Edward Elton, Periklis Charchalakis, and Elias Stipidis

Cognitive Load Assessment of Tractor Driver Based on Cognitive Task Workload Model 511
Yeqing Pei, Xiaoping Jin, Zhenghe Song, Haoyang Li, Ling Luo, and Bowen Zheng

Road and Rail: Automation

A Review of Non-driving-related Tasks Used in Studies on Automated Driving 525
Frederik Naujoks, Dennis Befelein, Katharina Wiedemann, and Alexandra Neukum

Driverless Pods: From Technology Demonstrators to Desirable Mobility Solutions 538
Joscha Wasser, Cyriel Diels, Anthony Baxendale, and Michael Tovey

Testing Scenarios for Human Factors Research in Level 3 Automated Vehicles 551
Christian Gold, Frederik Naujoks, Jonas Radlmayr, Hanna Bellem, and Oliver Jarosch

Autonomous Vehicles: Reliability of Their Perception of the World Around Them and the Role of Human Driver 560
Ata Khan

Connecting Rural Road Design to Automated Vehicles: The Concept of Safe Speed to Overcome Human Errors 571
Pasquale Colonna, Paolo Intini, Nicola Berloco, and Vittorio Ranieri

A Longitudinal Simulator Study to Explore Drivers' Behaviour During Highly-Automated Driving 583
David R. Large, Gary Burnett, Andrew Morris, Arun Muthumani, and Rebecca Matthias

Understanding and Applying the Concept of "Driver Availability" in Automated Driving 595
Claus Marberger, Holger Mielenz, Frederik Naujoks, Jonas Radlmayr, Klaus Bengler, and Bernhard Wandtner

What to Expect of Automated Driving: Expectations and Anticipation of System Behavior 606
Johanna Josten, Teresa Schmidt, Ralf Philipsen, Lutz Eckstein, and Martina Ziefle

The Design of a Vibrotactile Seat for Conveying Take-Over Requests in Automated Driving 618
Sebastiaan M. Petermeijer, Paul Hornberger, Ioannis Ganotis, Joost C.F. de Winter, and Klaus. J. Bengler

Learning from the Best – Naturalistic Arbitration for Cooperative Driving 631
Gina Weßel, Constanze Schreck, Eugen Altendorf, Yigiterkut Canpolat, and Frank Flemisch

Modelling the Dynamics of Driver Situation Awareness in Automated Driving 643
Zhenji Lu, Riender Happee, and Joost de Winter

Evaluation of an Autonomous Vehicle External Communication System Concept: A Survey Study 650
Jingyi Zhang, Erik Vinkhuyzen, and Melissa Cefkin

Machine Learning and Big Data Analytics in Support of Fleet Safety During Severe Weather 662
Zachary Spielman, David I. Gertman, Haoran Liu, Ira Pray, Justin Traiteur, Scott Wold, and Steven Wysmuller

A Study on the Human and the Automation in Automated Driving: Getting to Know Each Other 672
Eugen Altendorf, Raphael Schütz, Yigiterkut Canpolat, Gina Weßel, and Frank Flemisch

How Automation Level and System Reliability Influence Driver Performance in a Cut-In Situation 684
Jonas Radlmayr, Veronika Weinbeer, Cathrine Löber, Mehdi Farid, and Klaus Bengler

Looking at Drivers and Passengers to Inform Automated Driver State Monitoring of In and Out of the Loop 695
Christopher D.D. Cabrall, Veronika Petrovych, and Riender Happee

Taxonomy of Traffic Situations for the Interaction between Automated Vehicles and Human Road Users 708
Tanja Fuest, Lenja Sorokin, Hanna Bellem, and Klaus Bengler

Effect of Warning Levels on Drivers' Decision-Making with the Self-driving Vehicle System 720
Wenjie Cui, Ronggang Zhou, Yunhao Yan, Linghua Ran, and Xin Zhang

Does Shifting Between Conditionally and Partially Automated Driving Lead to a Loss of Mode Awareness? 730
Anna Feldhütter, Christoph Segler, and Klaus Bengler

Information Expectations in Highly and Fully Automated Vehicles 742
Cyriel Diels and Simon Thompson

Obtaining Design Requirements from the Public Understanding of Driverless Technology 749
Patrick Langdon, Ioannis Politis, Mike Bradley, Lee Skrypchuk, Alex Mouzakitis, and John Clarkson

Manual Takeover and Handover of a Simulated Fully Autonomous Vehicle Within Urban and Extra-Urban Settings 760
Phillip L. Morgan, Chris Alford, Craig Williams, Graham Parkhurst, and Tony Pipe

Road and Rail: Infrastructure and Road Safety

Driving Performance, Adaptation, and Cognitive Workload Costs of Logo Panel Detection as Mediated by Driver Age 775
Mei Ying Lau and David Kaber

Comparison Between Young Male Drivers' Self-assessed and Objectively Measured Driving Skills 787
Laila M. Martinussen, Mette Møller, and Carlo G. Prato

Identification of Criteria for Drivers' State Detection 798
Nicolas D. Herzberger, Gudrun M.I. Voß, and Maximilian Schwalm

A Real Case-Based Study Exploring Influence of Human Age and Gender on Drivers' Behavior and Traffic Safety 807
Nazha R. Ghadban, Galal M. Abdella, Khalifa N. Al-Khalifa, Abdel Magid Hamouda, and Khadija B. Abdur-Rouf

Verification of Installed Position of LED Block Equipped with Projections to Indicate Travel Direction 817
Tomoyuki Inagaki, Hideaki Nagahama, Norihiro Ikeda, Kazuya Takahashi, Kiyohito Takeuchi, Hiroshi Ogino, Katsuya Sato, Sin-Ichi Ito, and Shoichiro Fujisawa

Do End Users Really Have a Place in the Design Arena When Safe Design Is Critical? 825
Elise Crawford, Yvonne Toft, Ryan L. Kift, and Geoff Dell

Seat Belts Unfastened: Non-seat Belt Use in the Czech Republic 834
Vit Gabrhel, Petr Zamecnik, Veronika Kureckova, and Pavel Rezac

The Effect of See-Through Truck on Driver Monitoring Patterns and Responses to Critical Events in Truck Platooning 842
Bo Zhang, Ellen S. Wilschut, Dehlia M.C. Willemsen, Tom Alkim, and Marieke H. Martens

Road and Rail: Vulnerable Road Users

Driver–Cyclist Interaction Under Different Bicycle Crossroad Configurations 855
Francesco Bella and Manuel Silvestri

Operational and Safety Analysis of Signage and Pavement Marking Treatments in Puerto Rico Dynamic Toll Lane Using a Driving Simulator 867
Bryan Ruiz, Didier Valdés, Benjamín Colucci, Johnathan Ruiz, Ricardo García, and Enid Colón

Is the Driving Behaviour of Young Novices and Young Experienced Drivers Under Alcohol Linked to Their Perceived Effort and Alertness? 878
Catherine Berthelon and Edith Galy

Injury Severity Analysis in Vehicle-Pedestrian Crashes 884
Younshik Chung, Tai-Jin Song, and Juyoung Kim

Road and Rail: Driving Simulation and Test Track

Motion Sickness Measurements for Young Male Adults in Vitality, Endurance, Profiles and Sensitivity 895
Aie Liu, Li Ding, and Yan Li

Modeling the Real World Using STISIM Drive® Simulation Software: A Study Contrasting High and Low Locality Simulations 906
Craig K. Allison, Katie J. Parnell, James W.H. Brown, and Neville A. Stanton

What Driving Abilities Do Racing Video Games Stimulate? Rating the Levels of Realism Experienced in Commercial Racing Video Games 916
Daniele Ruscio

Multivariate Differences in Driver Workload: Test Track Versus On-Road Driving 927
Jack L. Auflick

Road and Rail: Eco-Driving

Range Makes All the Difference? Weighing up Range, Charging Time and Fast-Charging Network Density as Key Drivers for the Acceptance of Battery Electric Vehicles 939
Julian Halbey, Ralf Philipsen, Teresa Schmidt, and Martina Ziefle

Fast-Charging Stations or Conventional Gas Stations: Same Difference? - Variations of Preferences and Requirements 951
Ralf Philipsen, Teresa Schmidt, and Martina Ziefle

Fuzzy Logic Based Merging Gap Acceptance Model Incorporating Driving Styles and Drivers' Personalities 963
Chen Chai, Xuesong Wang, Yiik Diew Wong, and Yidan Gao

Road and Rail: Rail Passengers

Providing Improved Crowding Information to Provide Benefits for Rail Passengers and Operators 973
James Pritchard

User Experience Evaluation on Ticket Gage of Subway Station: A Repertory Grid Approach 985
Linna Hu, Yimu Yang, Long Liu, and Hua Dong

A Study on Quantitative Evaluation of Relationship Between Longitudinal Shock and Ride Comfort by Train Braking 998
Toshiro Miura and Toshiyuki Izumi

Guidelines for Electronic Systems Designed for Aiding the Visually Impaired People in Metro Networks 1010
Eliete Mariani and Marcelo Eduardo Giacaglia

Road and Rail: Public Transport

A Comparison of Auditory and Visual in-Vehicle Stop Signals in Philippine Public Road-Based Transportation 1025
Josiel Catherine V. Almeda, Majella Cayley J. Domingo, and Alyannah Bianca L. Saringan

Measuring Mobility and Transport Services: The METPEX Project 1036
Andree Woodcock, Yusak Susilo, Marco Diana, Roberto Abenoza, Miriam Pirra, and Michael Tovey

From Task Analysis to Innovation 1046
Denis Miglianico, Luc Moyart, and Mathieu Mouchel

Road and Rail: Driving Automation and Human Factors Issues

Three Driver and Operator Behaviour Models in the Context of Automated Driving – Identification of Issues from Human Actor Perspective 1059
Pirkko Rämä and Hanna Koskinen

The Right Moment for Braking as Informal Communication Signal Between Automated Vehicles and Pedestrians in Crossing Situations 1072
Matthias Beggiato, Claudia Witzlack, Sabine Springer, and Josef Krems

Older Drivers and Driving Automation 1082
Anabela Simoes

Road and Rail: Rail Safety and Design

A Comparison of Three Systemic Accident Analysis Methods Using 46 SPAD (Signals Passed at Danger) Incidents 1097
Kate Dixon, Patrick Waterson, and Jo Barnes

Understanding Railway Employees' Perceptions of Senior Managers' Safety Commitment 1109
Kate C. Bowers and Mark Fleming

Proposal of Ergonomic Intervention in Horizontal Traffic Signaling 1121
Claudemilson dos Santos and Galdenoro Botura Jr.

Road and Rail: Transportation Logistics

Collective Cars in Cities: Optimal Management in Real Time Through DES Approach. 1133
Jennie Lioris

Human Factors in the Design of Automated Transport Logistics 1145
Matthias Neubauer and Oliver Schauer

Erratum to: Advances in Human Aspects of Transportation E1
Neville A. Stanton

Author Index. 1157

Aviation: Human Factors Safety in Aviation and Aerospace

Influencing Factors on Error Reporting in Aviation - A Scenario-Based Approach

Sebastian Sieberichs(✉) and Annette Kluge

Department of Work, Organizational and Business Psychology,
Ruhr-University Bochum, Universitätsstraße 150, 44801 Bochum, Germany
sebastian.sieberichs@rub.de

Abstract. As crew reports represent an important data source for safety management in aviation, the present study investigated influencing factors on error reporting using a scenario-based approach with 33 prototypical aviation scenarios including eleven threat categories and three risk levels (low, medium, high). Ninety-three pilots from eight countries participated in an online survey to assess their attitude towards reporting. The results show that errors attributed to environmental factors are said to be reported more often than errors related to the individual pilot, e.g. proficiency (d = 2.29). High-risk events are more often reported than low-risk events (d = 3.54) and medium-risk events (d = 1.86). Airlines should promote the appreciation of reports about individual related errors and events with low risks.

Keywords: Error reporting · Aviation · Safety management · Type of error · Risk level

1 Introduction

High Reliability Organizations (HRO) like airlines strive for extremely low failure rates to avoid high costs and to improve the overall safety level [1, 2]. To achieve highest possible safety standards, proactive Safety Management Systems (SMS) emphasize a distinct identification of potential hazards by using multiple data sources [3].

As human error contributes to over 70% of aviation accidents, crew reports represent an important source of data, since they facilitate understanding the reasons why an event has happened [4]. Moreover, reports can help to picture a risk distribution specific to the respective organization [5]. In addition to the fact that most error-reporting systems are voluntary, recent research underlines that underreporting rates of adverse events is very high (ranging from 50% to 90%) leaving a high level of uncertainty in airline's safety databases [6, 7]. In contrast, today's airlines are also facing over-reporting issues in certain areas [4].

Therefore, crew reports cannot provide exact data that reflect amount and rates of threats and errors, although reporting rates themselves are commonly used as a safety

The original version of the book was revised: Missed out corrections have been incorporated. The erratum to the book is available at 10.1007/978-3-319-60441-1_108

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_1

performance indicator to measure the safety level [4, 8]. Due to the fact that reporting systems were never originated being a safety-measuring tool and changing reporting rates cannot be interpreted as an indicator for an improving safety level or a concerning organizational "blind spot", more exact data about reporting rates is needed to implement effective barriers and defenses and to eliminate unmanageable, undiscovered risk areas [9, 10].

The purpose of the study was to get more detailed information about reporting rates applying scenario-based items for the investigation of influencing factors on error reporting, focusing on different threat categories in combination with different risk levels.

2 Error Reporting and Its Importance for Safety Management in Aviation

2.1 An Integrated Threat Model

Human error has been involved in most civil and military aviation accidents [5]. First attempts to build hazard-based databases out of crew reports were already accomplished in 1975 [11]. In the following decades, several aviation-specific reporting systems were implemented, but most systems were not designed using any theoretical concept of human error [5, 12]. Since then, we learned that the type of error matters, as it determines whether the error can be detected by the organization [23].

Literature shows a variety of human error definitions: Reason [14] suggests a three-fold classification of types of errors, distinguishing between skill-based slips, rule-based mistakes and knowledge-based mistakes. Extending the scope to organizational error in aviation context, Wiegmann and Shappell [5] define five different perspectives on the nature and causes of human error (Cognitive, Ergonomics and systems design, Aeromedical, Psychosocial and Organizational). The "Model of Flight Crew Error Management" by Helmreich [8] defines flight crew error as an action or inaction that leads to deviation from organizational or crew-intentions or expectations [15]: The model differentiates between

- intentional non-compliance errors or conscious violations of SOPs or regulations,
- procedural errors in which the intention is correct but the execution flawed,
- communication errors,
- proficiency errors that indicate a lack of knowledge or stick and rudder skill and
- operational decision errors.

Combining these different human error definitions, an Integrated Threat Model was developed for this study to structure possible errors during pilot's work on flight decks:

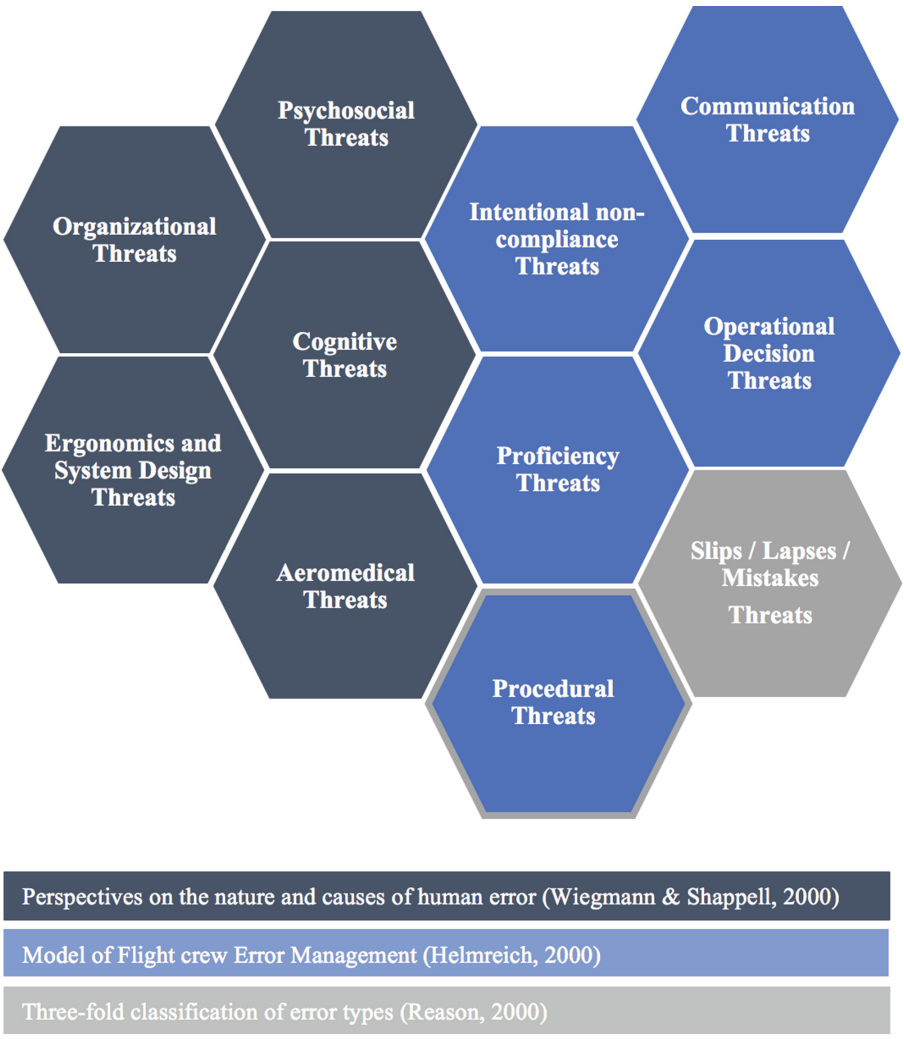


Fig. 1. Integrated Threat Model

2.2 Error Reporting

The application of error reporting systems is essential for the safety level of many HROs like nuclear power, petrochemical industry or aviation [16, 17]. Since the best method to prevent future errors is to learn from previous errors, reports reveal a unique opportunity to answer the question why something has happened [9, 18]. Following Reason (2016), the detailed analysis of mishaps, incidents, near misses, but also "free lessons" is needed to discover error traps and to find the edge that one might fall over [17].

The motivation of individuals to file a report is influenced by the seriousness of errors, a direct or indirect involvement, the working environment, the legal protection, the motive of the wrongdoer and the relationship to the wrongdoer [19]. The estimated

dark figures concerning underreporting differ among safety-related events; events with bad outcomes are more likely to be reported than events with innocuous outcomes [7, 9]. Also, the study of Siebericchs and Kluge [20] emphasizes different reporting rates depending on the type of error. Although scientific results about individual influencing factors on underreporting are available, there is still a confusion for pilots regarding what matters should be reported [19].

2.3 Error Reporting as Part of Safety Management

Safety within the context of aviation can be described as "the state in which the possibility of harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and safety risk management" [21]. To allow a structured approach towards managing flight-related risks, the collection of data using various data sources is one key element of the work within safety departments and a fundamental prerequisite of Safety Management Systems (SMS) [10, 21].

Reports can facilitate a mutual and exact understanding of errors by receiving accurate and timely reports on incidents or accidents [21, 22]. In order to increase the amount of data collected and facilitate their quantitative analysis, reporting systems shall be designed voluntary, non-punitive and confidential [6]. Thus, the organization will not only receive reports on obvious incidents, but also on hidden events [6].

A classification of the risks reflected by these reports is usually performed by experts evaluating probability and the severity of the event, and results in a so-called risk level, usually ranging from "low" to "severe" [22]. Since events with a low risk level may still contain relevant data as they could mark the possibility of a more severe outcome, it is important to reveal information about both, high- and low-risk events to sustain a database pointing out risk distributions, although reporting systems were originally not designed for safety monitoring [4, 21, 23].

2.4 Hypotheses

The present study investigates how the probability that an error is reported by the crew depends on the underlying threat category and risk level.

As reporting rates differ between safety related events, the results of Haslbeck, Schmidt-Moll and Schubert [7] and Tani [19] suggest that events attributed to environmental factors will more often be reported than individual-related errors and events with a higher risk level will more likely be reported than events with a low risk level.

The following hypotheses are investigated:

H1: Errors being attributed to environmental factors will be said to be reported more often than personal-related errors.

H2: The higher the risk level within an event, the more positive the attitude towards reporting.

3 Materials and Methods

3.1 Sample

One hundred four pilots participated in the present survey. Eleven incomplete datasets were excluded. The sample of N = 93 (44 captains and 48 officers) consisted of seven females and 86 males serving on short-range (53) and long-range flights (39).

Figure 2 shows the country of the participants:

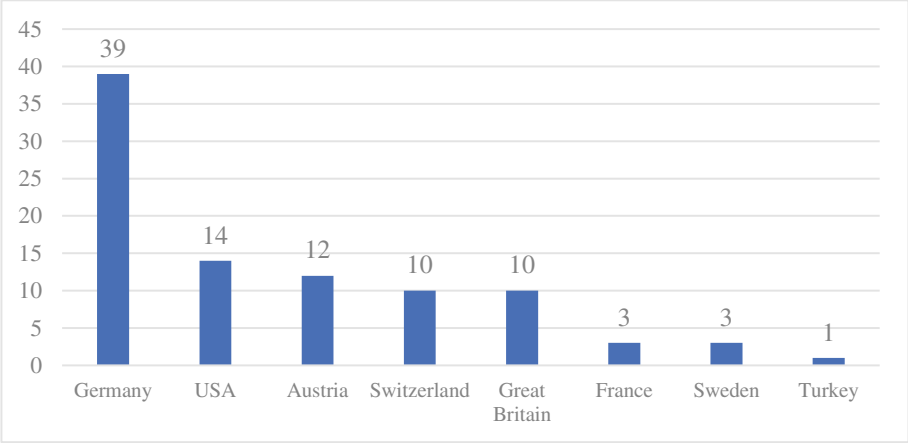


Fig. 2. Country of the participants

Figure 3 shows the age of the participants:

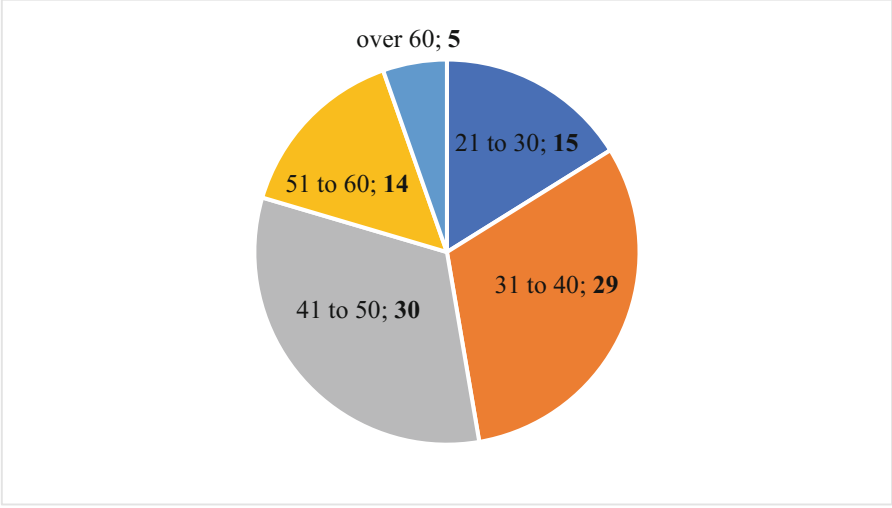


Fig. 3. Age of the participants

3.2 Instrument Development and Pretest

To investigate the attitude towards reporting, a group of five safety-experts developed 33 scenario-based items, based on the Integrated Threat Model (see Fig. 1). These experts were active commercial pilots with over ten years of experience and safety service officers with a major expertise in safety management, safety reporting and safety report evaluation. The group's task was to identify typical errors appertaining to the different threat categories of the Integrated Threat Model and to translate them into scenarios. The scenarios were supposed to be prototypical for the corresponding threat and needed to mirror different events corresponding to three underlying risk categories "low", "medium" and "high". Besides using an evidence-based source (Operations-Manual-B, Airbus A320 Fam., FCOM-5-010 "Preventing identified risks"), the group was asked to take into consideration past events that have actually been reported in previous safety reports.

The following set of rules was given to the group: The scenarios are supposed to lead to an increased risk and not to a real damage. They shall not provide concrete details about crew position. The scenarios need to include details, such as whether the event happened on a short- or long-range flight, whereas short-range scenarios shall be used preferably, as pilots normally start their career on short-range aircraft. The scenarios shall be, as far as possible, equal in length, and they shall be ranked assuming that reporting is non-confidential and not mandatory.

After development of a set of 33 scenarios, a pretest was conducted to assess comprehensibility, plausibility and the time needed for the assessment with a group of seven randomly picked active commercial pilots without experience in safety management. Additionally, a qualitative item asking for individual factors influencing reporting behavior was added to the pretest.

As a result, the scenarios of the four threat categories "Intentional Non-Compliance", "Lapses", "Mistakes" and "Communication" were changed entirely, since they were not clearly understandable or were rated not to be the best solution for a prototypical threat. In addition, the order of items was changed. Five additional items were generated, based on the stated factors having an influence on reporting behavior (Example item: "*My attitude towards reporting is influenced by political issues"*).

The following list summarizes the content of the scenarios after the pretest (Table 1):

Table 1. List of scenarios sorted by threat category

Threat category	Scenario
Organizational	Cost cutting by the company leads to incorrect flight documents and wrong fuel calculation
Ergonomics/system design	Error-prone cockpit layout leads to inadvertent setting of parking brake
Aeromedical	Night flight leads to fatigue and to a temporarily uncontrolled aircraft during cruise flight
Psychosocial	Brake-down of communication on flight deck leads to impaired CRM

(*continued*)

Table 1. (*continued*)

Threat category	Scenario
Intentional non-compliance	Omitted or late briefings lead to flying a wrong departure route
Slips	Wrong flight level selections lead to altitude excursion
Lapses	Very low weather radar brightness leads to a thunderstorm incursion
Mistakes	Mistakes during go-around lead to increased workload and violations
Communication	Communication ambiguities lead to incorrect flap settings
Proficiency	Lack of manual flying-skills leads to unsafe aircraft state
Operational decision	Unclear weather indication during take-off leads to an unsafe aircraft state

Example item for an Organizational threat with a high risk level:

"*To reduce the overall costs, your company has decided to calculate and deliver the Operational Flight Plans (OFP) by a new contractor. You are on a medium range flight. After 30 min flight time, you recognize that the OFP has been calculated assuming an Airbus A319 instead of an Airbus A321. Due to the difference in fuel flow, you recalculate your fuel demand and decide to continue to your destination. After parking, the remaining fuel allows an endurance of 5 min above final reserve fuel".*

Example item for an Intentional Non-compliance threat with a medium risk level:

"You are on a short-range rotation together with a colleague you know from past flights. Due to the tight schedule, you decide to deviate from your SOPs and perform your briefings (departure, engine-failure and rejected take-off) during taxi-out and not at the gate. Shortly after lift-off, ATC instructs you to pick up a heading, as you are flying the wrong Standard Instrument Departure (SID). A comparison of the SID on your departure chart with your FMS/FMC reveals that the briefed departure is not the departure in your FMS/FMC. The flight is continued without further incident".

3.3 Variables

The scenarios were grouped in two scales, named "Environmental Threats" and "Personal Threats":

The scale "Environmental Threats" consists of twelve items investigating organizational, ergonomics and system design, communication and aeromedical threats, since these items represent human errors that can be attributed to environmental factors. The internal consistency of the scale lies in an acceptable range (Cronbach's $\alpha = .78$).

The scale "Personal Threats" consists of 21 items for intentional non-compliance, slips, lapses, mistakes, proficiency, operational decision and psychosocial threats, as these items represent human errors that can be related to the individual pilot. The internal consistency of the scale lies in a good range (Cronbach's $\alpha = .85$).

The participants were asked to state their probability of reporting the 33 scenarios on a Five-Point Likert scale with the endpoints 1 = "Extremely unlikely" and 5 = "Extremely likely".

3.4 Procedure

Participants of the study were persons from the professional and personal social environment of the author who have been invited by e-mail to answer an online survey in English language taking 30 min. Additionally, they were asked to forward the e-mail and to share the link of the survey with further persons. No remuneration was offered for participation. The survey was available online for eight weeks between January 28th and March 25th, 2017.

4 Results

4.1 Descriptive Statistics

Table 2 sums up the descriptive statistics of the reporting rate with respect to the various threat categories:

Table 2. Reporting rate with respect to threats, listed in descending order (Scale E = Environmental Threats, P = Personal Threats; 1 = "Extremely unlikely" and 5 = "Extremely likely")

Threat categories	Scale	N	Min	Max	*M*	*SD*
Organizational	E	93	1.33	5	4.11	.70
Communication	E	93	1	5	4.03	.88
Ergonomics/system design	E	93	1.67	4.67	3.27	.61
Aeromedical	E	93	1	5	3.13	.79
Psychosocial	P	93	1	5	2.82	.85
Lapses	P	93	1	4.33	2.68	.55
Intentional non-compliance	P	93	1	4	2.51	.56
Slips	P	93	1	5	2.47	.68
Mistakes	P	93	1	4	2.20	.53
Operational decision	P	93	1	4	2.14	.64
Proficiency	P	93	1	5	1.90	.88

The likelihood of Organizational Threats and Communication Threats proved to be highest, Proficiency Threats and Operational Decision Threats were reported least. The full range of the scale was used for most of the threat categories.

The scale Environmental Threats shows a mean of $M = 3.63$ ($SD = .58$), the scale Personal Threats shows a mean of $M = 2.39$ ($SD = .50$).

Table 3 sums up the descriptive statistics about the reporting rate with respect to the three risk levels low, medium and high:

Table 3. Reporting rate with respect to the risk level listed in descending order (1 = "Extremely unlikely" and 5 = "Extremely likely")

Risk level	N	Min	Max	*M*	*SD*
High	93	1.73	5	3.81	.58
Medium	93	1.27	4.73	2.77	.54
Low	93	1	3.73	1.94	.47

4.2 Hypothesis Testing

Hypothesis 1 assumed errors being attributed to environmental factors will be said to be reported more often than personal-related errors. Since the t-test showed significantly higher values for Environmental Threats (M = 3.63, SD = .58) than for Personal Threats (M = 2.39, SD = .50), $t(92)$ = 24.32, p < .001, d = 2.29, 99% CI [1.11, 1.38], the hypothesis can be kept.

Hypothesis 2 assumed that a higher risk level leads to a more positive attitude towards reporting. Figure 4 sums up the results of the t-tests:

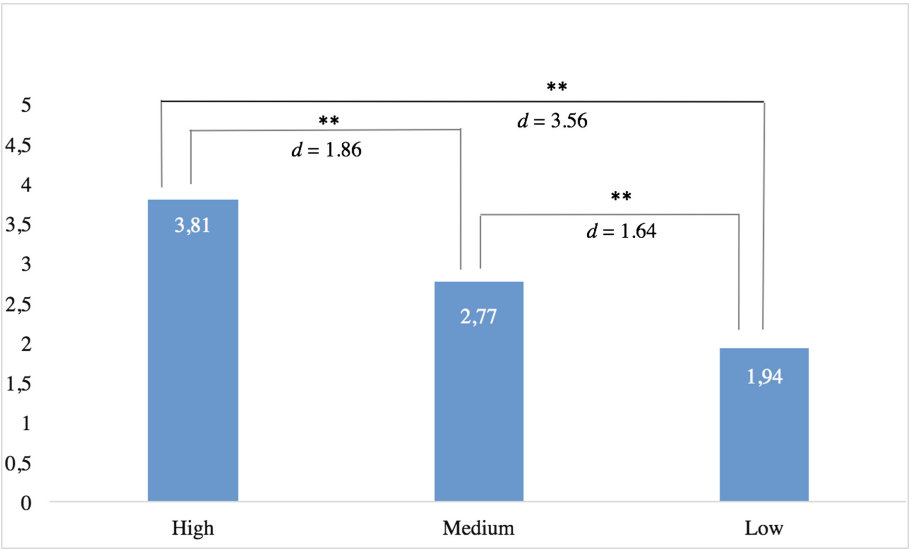


Fig. 4. Likelihood of reporting and risk level, results of t-test, Annotation: ** p < .001

The t-test showed

- significantly higher values for a High Risk Level (M = 3.81, SD = .58) than for a Low Risk Level (M = 1.94, SD = .47), $t(92)$ = 32.21, p < .001, d = 3.54, 99% CI [1.72, 2.03],
- significantly higher values for a High Risk Level (M = 3.81, SD = .58) than for a Medium Risk Level (M = 2.77, SD = .54), $t(92)$ = -27.48, p < .001, d = 1.86, 99% CI [.94, 1.14] and
- significantly higher values for a Medium Risk Level (M = 2.77, SD = .54) than for a Low Risk Level (M = 1.94, SD = .47), $t(92)$ = 23.61, p < .001, d = 1.64, 99% CI [.74, .93].

5 Discussion

The purpose of the study was to investigate influencing factors on error reporting in aviation using a scenario-based approach.

The results point out that errors which are attributed to environmental factors are more likely to be reported than individual-related errors: Safety-infringing events caused by incorrect flight documentation, ambiguous procedural communication requirements, error-prone cockpit layouts or night-time working environments are more likely to be reported than individual errors like operational decision errors or a lack of manual flying skills.

The very high effect sizes of $d > 2$ underline an unbalanced reporting situation that is tantamount to possibly undiscovered risk areas. Additionally, a lack of background information can be experienced by flight safety service offices if events are derived from other sources as flight data analysis, and no corresponding report has been filed adding background information. From a safety management point of view, the results underline that reporting systems were originally not designed for safety monitoring [4] as the risk distribution is incomplete and details about individual-related errors are underrepresented in comparison to events attributed to environmental factors.

High- and medium-risk events are more often reported than events with low risk levels. This is valid for high-risk events compared to medium- and low-risk events, but also for medium-risk events compared to low-risk events.

This clear underreporting situation coincide with the results of Tani (2010) [19], who states that the more an individual perceives an event being a problem for flight safety, the more likely she/he will report it. Although low-risk events do not infringe flight safety directly, they can be precursors of worse situations. Furthermore, the knowledge of low-risk events can discover unidentified risk areas and thus support preventing future errors by redesigning cockpit layouts, procedures or flight-time-schemes (Example: A reported low-risk error due to an error-prone cockpit design might prevent a high-risk error of someone else). Again, this result underlines the difficulty to use crew reporting as the only means of building safety databases, since the picture of the risk distribution is incomplete, as details about events with a low risk factor are not available in the database.

From a methodological point of view, the number of 33 scenarios with approximately 100 words each turned out to be demanding for some participants. Written comments suggest that a uniform assessment of all scenarios cannot be assured since it is hard to step into 33 situations within 30 min. This goes along with the fact that more than 10% of the participants did not complete the survey entirely.

Common risk assessment in aviation is usually performed by (subjective) expert judgment. In terms of the used scenarios, both the scenarios themselves and the associated risk levels are a subjective evaluation of a prototypical situation concerning these two criteria. Consequently, there is a good chance of a different group of safety experts developing an entirely different set of prototypical items, and rating the scenarios differently regarding risk. Additionally, a direct comparison of two items of different threat categories it is unfavorable, as the potential damage differs between the threat categories even if they share the same risk level (Example: A psychosocial threat like a communication breakdown vs. a near mid-air collision of two aircraft).

The external validity of the study might be impaired since its participants come from eight different countries, and the groups are not equal in size.

5.1 Implications for Safety Management and Further Investigations

The study demonstrated that threat categories and risk levels influence the attitude towards reporting. The results are important for the identification of organizational risk blind-spots and risk areas where over-reporting may distort the risk landscape. This chimes with the aim of current regulatory projects like the European PROSPERO project that aims for an enhanced perspective on human error, providing a proactive anti-cipation of risks [24].

To support safety management, pilots should be encouraged to file reports not only about events that are favored by environmental factors but also about individual-related events. Additionally, the appreciation of reports on low-risk events also needs to be promoted, as these reports may contain important information as well.

This study proved differences in reporting; therefore, further research should concentrate on a more differentiated analysis between the different threat categories to get a better picture of the risk distribution within the organization's structures. Furthermore, the influence of the risk level needs to be investigated in higher detail, since this study assumed the influence to be equally strong for all eleven threat categories.

Further research should also investigate the impact of personality traits and safety climate on error-reporting, following a risk-based approach in combination with these scenarios.

Day after day pilots make errors and need to decide whether they file a report or not. If organizations enhance appreciation for reports on individual-related errors and events with low risks, and provide more exact details on what reports are helpful and necessary, a balanced ratio between over- and under-reporting can be achieved, and flight safety can be further improved.

References

1. Hales, D.N., Chakravorty, S.S.: Creating high reliability organizations using mindfulness. J. Bus. Res. **69**(8), 2873–2881 (2016)
2. La Porte, T.R., Consolini, P.M.: Working in practice but not in theory: theoretical challenges of "high-reliability organizations". Institute of Governmental Studies, University of California, Berkeley (1991)
3. Liou, J.J., Yen, L., Tzeng, G.H.: Building an effective safety management system for airlines. J. Air Transp. Manag. **14**(1), 20–26 (2008)
4. Macrae, C.: The problem with incident reporting. BMJ Qual. Saf. **25**(2), 71–75 (2016)
5. Shappell, S.A., Wiegmann, D.A.: A Human Error Approach to Aviation Accident Analysis: The Human Factors Analysis and Classification System. Ashgate Publishing, Ltd., Farnham (2012)
6. Barach, P., Small, S.D.: Reporting and preventing medical mishaps: lessons from non-medical near miss reporting systems. BMJ **320**(7237), 759 (2000)
7. Haslbeck, A., Schmidt-Moll, C., Schubert, E.: Pilots' Willingness to report aviation incidents. In: International Symposium on Aviation Psychology (2015)
8. Helmreich, R.L.: On error management: lessons from aviation. BMJ **320**(7237), 781 (2000)
9. Gilbey, A., Tani, K., Tsui, W.H.K.: Outcome knowledge and under-reporting of safety concerns in aviation. Appl. Cogn. Psychol. **30**(2), 141–151 (2016)

10. Hollnagel, E., Nemeth, C. P., Dekker, S. (Eds.).: Resilience engineering perspectives: remaining sensitive to the possibility of failure, vol. 1. Ashgate Publishing, Ltd., Farnham (2008)
11. Zhao, B., Olivera, F.: Error reporting in organizations. Acad. Manag. Rev. **31**(4), 1012–1030 (2006)
12. Billings, C.E., Lauber, J.K., Funkhouser, H., Lyman, E.G., Huff, E.M.: NASA aviation safety reporting system (1976)
13. Wears, R.L.: Learning from near misses in aviation: so much more to it than you thought (2016)
14. Reason, J.: Human error: models and management. BMJ **320**(7237), 768–770 (2000)
15. Helmreich, R.L.: Culture and error in space: implications from analog environments. Aviat. Sp. Environ. Med. **71**(9; PART 2), A133–A139 (2000)
16. Furukawa, H., Bunko, H., Tsuchiya, F., Miyamoto, K.I.: Voluntary medication error reporting program in a Japanese national university hospital. Ann. Pharmacother. **37**(11), 1716–1722 (2003)
17. Reason, J.: Managing the Risks of Organizational Accidents. Routledge, Abingdon (2016)
18. Beaubien, J.M., Baker, D.P.: A review of selected aviation human factors taxonomies, accident/incident reporting systems and data collection tools. Int. J. Appl. Aviat. Stud. **2**(2), 11–36 (2002)
19. Tani, K.: Under-reporting in aviation: an investigation of factors that affect reporting of safety concerns: a thesis presented in partial fulfilment of the requirements for the degree of Doctorate of Philosophy in Aviation at Massey University, Manawatu, New Zealand (2010)
20. Sieberichs, S., Kluge, A.: Good sleep quality and ways to control fatigue risks in aviation - an empirical study with commercial airline pilots. In: Advances in Physical Ergonomics and Human Factors, pp. 191–201. Springer, Berlin (2016)
21. International Civil Aviation Organization: Safety Management Manual (SMM) (2013)
22. Hollnagel, E., Woods, D.D., Leveson, N.: Resilience engineering: concepts and precepts. Ashgate Publishing, Ltd., Farnham (2007)
23. Roelen, A.L.C., Klompstra, M.B.: The challenges in defining aviation safety performance indicators. In: Presentado en la 11a Conferencia de International Probabilistic Safety Assessment and Management y la Conferencia annual de European Safety and Reliability, Helsinki, Finlandia (2012)
24. McDonald, N., Corrigan, S., Ulfvengren, P., Baranzini, D.: Proactive safety performance for aviation operations. In: International Conference on Engineering Psychology and Cognitive Ergonomics, pp. 351–362. Springer, Berlin (2014)

Quantifying Pilot Contribution to Flight Safety During Hydraulic Systems Failure

Lynda J. Kramer[1(✉)], Timothy J. Etherington[2], Randall E. Bailey[1], and Kellie D. Kennedy[1]

[1] NASA Langley Research Center, Hampton, VA, USA
lynda.j.kramer@nasa.gov
[2] Rockwell Collins, Cedar Rapids, IA, USA

Abstract. Accident statistics cite the flight crew as a causal factor in over 60% of large transport aircraft fatal accidents. Yet, a well-trained and well-qualified pilot is acknowledged as the critical center point of aircraft systems safety and an integral safety component of the entire commercial aviation system. The latter statement, while generally accepted, cannot be verified because little or no quantitative data exists on how and how many accidents/incidents are averted by crew actions. A joint NASA/FAA high-fidelity motion-base human-in-the-loop test was conducted using a Level D certified Boeing 737-800 simulator to evaluate the pilot's contribution to safety-of-flight during routine air carrier flight operations and in response to aircraft system failures. To quantify the human's contribution, crew complement (two-crew, reduced crew, single pilot) was used as the independent variable in a between-subjects design. This paper details the crew's actions, including decision-making, and responses while dealing with a hydraulic systems leak.

Keywords: Flight crew error · Aviation Safety · Increasingly autonomous systems · Automation · Crew complement · Reduced crew operations · Hydraulic system

1 Introduction

Accident statistics cite the flight crew as a causal factor in over 60% of large transport aircraft fatal accidents [1]. Yet, the Air Line Pilots Association says that "a well-trained and well-qualified pilot is acknowledged as the critical center point of aircraft systems safety and an integral safety component of the entire commercial aviation system" [2]. The latter statement, while generally accepted, cannot be verified because little or no quantitative data exists on how and how many accidents/incidents are averted by crew actions. Anecdotal evidence suggests crews handle routine failures on a daily basis and Aviation Safety Action Program (ASAP) data [3, 4] supports this assertion but its data is not publicly releasable. Without hard data, the contribution and methods employed by pilots to improve the safety of flight is difficult to define. Developing ways to augment and/or improve a pilot's ability to contribute to flight safety is similarly ill-defined and is hard to characterize in the absence of quantifiable data.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_2

A joint NASA/FAA high-fidelity motion-base simulation experiment specifically addressed this void by collecting data to quantify the human (pilot) contribution to safety-of-flight and the methods used by pilots in today's National Airspace System as they handled normal and non-normal conditions during typical revenue-like flight operations. These data are fundamental to and critical for the design and development of future increasingly autonomous systems that can better support the human in the cockpit. Different crew complement configurations were tested to gain understanding of the safety afforded by having two crewmembers on the flight deck. Normal two-crew operations were contrasted and compared to conditions where the second crew member was unavailable when the non-normal condition occurred but became re-engaged after returning to the flight deck and another case where only a single pilot was on the flight deck (i.e., simulating an incapacitated pilot). This paper details preliminary results and analysis of one of six non-normal events tested – a hydraulic leak in the System A reservoir.

2 Methodology

Crew complement (single pilot and crewed configurations) was experimentally manipulated during normal and increasingly challenging non-normal airline operations to quantify the pilot contribution to flight safety.

2.1 Experiment Design

The test objectives of the experiment were as follows:

- Establish "baseline" levels of performance and safety with nominal two-crew configuration as well as collect data to assess the performance and safety decrements in reduced crew and single pilot crew complements for present-day flight deck design and certification; and,
- Identify technology requirements from these data for increasingly autonomous systems that might assist future two-crew operations and eventually, enable reduced crew or ultimately, single pilot operations.

To assess human performance and safety, the experiment contrasted two-crew operations to conditions when one of the pilots was absent from the flight deck. If the condition included a temporary absence, it was designated as reduced crew operations (RCO). If the condition included a permanent absence, it was designated as single pilot operations (SPO).

The independent variables were crew complement and scenario. The three crew complement configurations were: Two-Crew, RCO, and SPO. Two normal scenarios and six non-normal scenarios were flown over the two days of data collection. The non-normal scenarios were grouped into three categories (A, B, and C), with two non-normal runs in each category. Category A featured failures, initially unannunciated, with autopilot available; Category B featured annunciated failures with autopilot available; and, Category C featured annunciated failures with autopilot not available. Alert type and autopilot state were used to identify workload and automation issues

(i.e., by availability of autopilot) and flight crew awareness and monitoring for normal / non-normal operations (i.e., alerting). All flights were flown to landing.

Failures were triggered near top of climb or top of descent. This paper details one Category B failure - a loss of System A hydraulic system. Reference 5 provides a detailed description of the experiment design (factors, metrics, and run matrix) and details one Category C failure.

The data shown here is taken from 18 nominal Two-Crew runs, 18 nominal SPO runs, 6 nominal RCO runs (with the Captain resting), and 18 hydraulic leak non-normal runs (6 SPO, 6 RCO, 6 Two-Crew). For the RCO configuration, the non-normal started out with First Officer flying from the right seat and the Captain resting in the left seat, isolated in sight and sound from the cockpit. Two minutes after the flying pilot was alerted to the hydraulic failure, the resting pilot returned to flying duties in the cockpit. For the SPO configuration, each pilot flew from the left seat.

2.2 Participants

Thirty-six pilots (18 crews total), representing 5 airlines, participated in this experiment. Each pilot held an Airline Transport Pilot rating and was current in the 737-800 aircraft as either Captain or First Officer. All participants were male. Crews were paired by function (Captain or First Officer) and employer to minimize conflicts in training, standard operating procedures, and crew resource management techniques. Crews were instructed to bring their company's paper and/or electronic charts and 737-800 checklists with them to further reduce conflicts in training and standard operating procedures.

2.3 Simulator

The research was conducted using the B-737-800 simulator operated by the FAA AFS-440 at Oklahoma City, OK. The simulator is Level D-certified and can be used for both initial and recurrent training. The simulator, although a Level D training device, is also fitted with experimental controls, modifications, and recording capability to support AFS-440's research mission. The fidelity of the simulator and the recording capability were both critical to this research effort.

The test was set-up to replicate a normal airline operation in today's National Airspace System. An air carrier flight from Denver (KDEN) to Albuquerque (KABQ) was used. Dispatch paperwork for the flight was provided to the crews and constituted the flight release.

The simulated weather en-route contained significant areas of convective activity along the Rocky Mountain Front Range and strong Northerly winds that required a north departure out of KDEN before a circuitous route to the west and then south to KABQ. This same planned route of flight was used for the entire two days of data collection. Weather and visibility were designed to affect any diversion decisions [5].

A live controller and pseudo-pilot(s) were tied into the simulation radio in real-time to simulate Air Traffic Control (ATC) and some proximate traffic to promote realism

and maintain realistic pilot workload levels. A confederate also served as dispatcher in the Airline Operations Center and provided communications as necessary and appropriate when radioed.

2.4 Training

No additional training was conducted for the crews as they were qualified and current B737-800 pilots and the simulator was Level D-certified.

The crews were briefed on the purpose of the experiment and received the dispatch paperwork. The crews were instructed to use their company's standard operating procedures and checklists for the entire test, including any company dispatch calls and cabin crew communications.

Prior to boarding the aircraft, the crew reviewed the paperwork and discussed the flight plan and flight conduct. Once they boarded the aircraft, the crew did a familiarization check and reviewed the simulator safety briefing. Known simulator-isms and aircraft differences were identified and discussed with the crew prior to run initiation.

The aircraft initial condition was in the hold-short of Runway 35L at KDEN with the engines running, parking brake set. The Flight Management System (FMS) was pre-loaded with the planned flight routing and the crews were asked to double check the entries. After review and confirmation of the cockpit switches/set-up and completing their normal checklists, the crew called KDEN tower for departure.

Following clearance from ATC, the crew flew an entire nominal flight from KDEN to KABQ following the planned route of flight. The nominal flight served as a baseline for 'normal' airline two-crew operations (i.e., nominal data) to which the non-normal runs flown in the RCO and SPO configurations would be compared. The nominal flight also promoted familiarity for the two-person crew interaction during the approximately 1.3 h of flight time required for completion. This nominal flight was flown as the first run on Day 1 of data collection for each crew.

3 Results

The results shown here describe the major findings of only one of the Category B failure conditions, a System A hydraulics failure.

The 737-800 has three hydraulic systems (A, B, and Standby) that operate independently at 3000 psi. Each has a reservoir, pumps and filters. Either A or B Hydraulic System can power all flight controls with no decrease in airplane controllability. System A also provides hydraulics for landing gear, ground spoilers, alternate brakes, Engine 1 thrust reverser, Autopilot A, normal nose wheel steering, and power transfer unit. System B provides hydraulics for leading edge flaps and slats, normal brakes, Engine 2 thrust reverser, Autopilot B, alternate nose wheel steering, landing gear transfer unit, autoslats, yaw damper, and trailing edge flaps. The Standby System provides a third source for the rudder control system and a second source for thrust reversers and leading edge flaps and slats.

The hydraulic leak failure was modeled as a large leak, at a rate of 10 gallons per minute, in the System A reservoir. When the reservoir quantity dropped to less than

18.7% full, the System A hydraulics failure was annunciated to the flight crew through illumination of the: (a) ENG 1 (engine-driven pump) and ELEC 2 (electric-motor-driven pump) LOW PRESSURE lights on the forward overhead hydraulic panel; and, (b) Left and Right side MASTER CAUTION lights and HYD system annunciator light on the glare shield annunciation panel. Approximately 30 seconds later, the System A Flight Controls LOW PRESSURE light illuminated on the forward overhead flight control panel and the MASTER CAUTION lights and FLT system annunciator light illuminated on the glare shield. Additionally, if the A side autopilot was engaged, it automatically disconnected and the autopilot disconnect horn would sound. The B side autopilot was still available after the failure. If not shut down in time, approximately one minute after loss of hydraulic fluid cooling, the electric hydraulic pump OVERHEAT light would illuminate. This light would remain illuminated even after the pump had been shut down until the pump cooled down.

3.1 Failure Handling and Flight Path Control

For the hydraulic leak scenario, the failure occurred approximately 5 minutes prior to the top of descent during the cruise phase of flight at 36,000 ft mean sea level (MSL) while heading south toward KABQ.

Once the failure occurred, 13 out of 18 pilots/crews declared an emergency with ATC. 11 pilots/crews requested special handling (holding pattern, vectors for descent from cruise altitude to a lower altitude, and vectors for long straight-in approach) from ATC.

There were two single-pilot runs where diagnosing and attending to the System A hydraulic failure significantly affected the pilot's airplane state awareness. These two single-pilot runs appear to be the only ones in the 18 hydraulic failure runs flown where dealing with the failure directly influenced the pilot's aircraft awareness as summarized below:

- SPO-Captain configuration run: Before the hydraulic leak occurred, the pilot was cleared to descend via the SANDIA3 arrival and a preselected altitude of 9,000 feet was set to allow the aircraft to descend at the calculated top of descent. When the failure occurred, the aircraft had not yet reached top of descent and the pilot declared an emergency and requested a hold to have time to troubleshoot the problem before descending. The pilot was given vectors to a hold and engaged heading select and then flight level change, which placed the automation in autothrottle arm mode and pitch on speed for the vertical mode. The pilot had intended to select altitude hold. The pilot asked for and was given a clearance to slow down for the hold from 0.78 to 0.7 Mach and the speed reduction caused the aircraft to start a descent. After losing almost 2000 feet of altitude over 20 seconds, the pilot realized his error and requested an altitude to hold from ATC. The pilot verbalized the altitude deviation during the run, but incorrectly thought this deviation was caused by putting the autothrottles in speed mode. Post-run, the pilot recognized his altitude excursion and said "I was convinced I was in a different mode and I descended through my altitude and didn't start my hold because I was in heading select". The pilot also commented post-run "I'm a little upset with myself here for that descent. I blew an altitude and even though I'm an emergency aircraft

and all that. I really like to stay out of other people's airspace. I thought that was an objectionable excursion."

- SPO-First Officer configuration run: ATC cleared the pilot to descend via the SANDIA3 arrival to 9000 feet approximately 40 seconds after the pilot was alerted to the failure of the hydraulic system and the failed side autopilot disconnect horn sounded. The pilot silenced the horn, acknowledged the clearance, declared an emergency, and set selected altitude to 9000 feet. Autothrottles were engaged and the flight directors were still providing lateral and vertical corrections to the FMS-programmed path as the pilot proceeded to troubleshoot the failure. While handling the checklists and troubleshooting, the pilot was not actively hand-flying the aircraft, but it was well trimmed and continued to follow the flight path. When top of descent was reached, the automation retarded the throttle and the aircraft started a descent at the trimmed airspeed. This action did not violate the clearance but was not what the pilot was expecting as he did not actively engage a descent or follow the flight director commands once the automation began the descent. After almost ten minutes and a loss of 12,000 feet of altitude, the pilot finally re-engaged the autopilot to the non-failed side once he reached the note in the Loss of System A checklist that the autopilot B was available. Post-run, the pilot said "I was single pilot, had to unstrap myself, trying to put the gear down. Had nobody to fly the airplane or monitor the systems. I have to totally rely on the autopilot while I'm messing around with the checklists, trying to program. It's a good thing I had at least one autopilot or it would have been a different situation".

3.2 Checklist Usage

Time-to-first correct checklist was used as a metric for quick and proper troubleshooting of equipment problems. For the System A Hydraulic failure, it was an alerted failure with annunciation on the flight deck that had a direct entry in the Quick Reference Handbook (QRH) with the Loss of System A checklist.

Crew complement was significant ($F(2,15) = 8.66$, $p = 0.003$) for time-to-start Loss Of System A checklist. Crews flying in the SPO or RCO configurations took approximately three times longer to start the correct checklist than those crews in the Two-Crew configuration (see Fig. 1). [The boxplots show the median ratings, with the 25^{th} and 75^{th} percentile spread in the data; the maximum and minimum values; and mean ratings (connected by a line).] There were no significant differences between the SPO and RCO configurations for time-to-start Loss of System A checklist. Only in one of the six RCO crews who experienced this failure did the pilot flying (PF) start the Loss of System A checklist before the resting pilot returned to the flight deck. Recall that a fixed delay of two minutes was implemented in this experiment before the resting pilot could return to the flight deck once summoned by the PF. Taking that delay into account the average time-to-start the loss of System A checklist was similar for the Two-Crew and RCO configurations. In the SPO case, getting access to the QRH and locating the correct checklist was significantly delayed compared to the Two-Crew configuration. Initially, the SPO crew had to hand-fly the aircraft because the failure disconnected the autopilot.

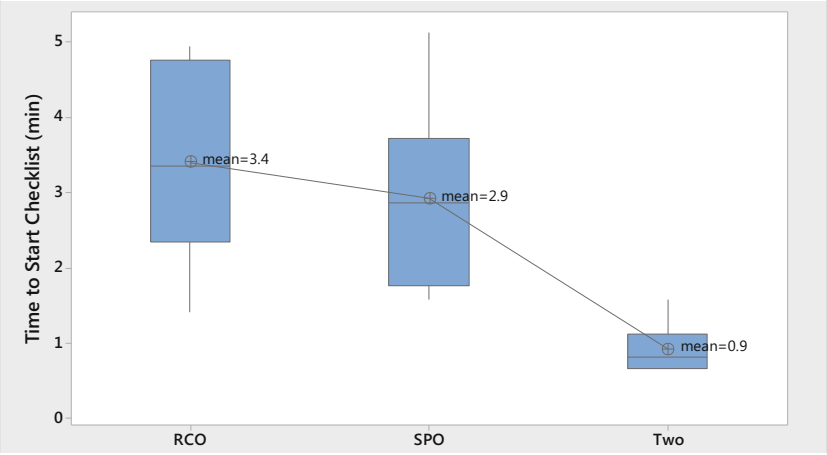


Fig. 1. Boxplot of time to start the loss of system a checklist after hydraulic failure

Time-to-complete the checklist was considered another indicator for safely handling failures. There were two checklists, the Loss of System A and Manual Gear Extension, to execute for the System A hydraulic failure. The time-to-complete metric for the Loss of System A checklist included time to execute all checklist items up to the deferred items in the Descent, Approach, Manual Gear Extension, and Before Landing checklists. The items in the descent, approach and before landing checklists varied by airline carrier, but were similar in their content. The manual gear extension checklist was the same for all carriers; so, the time-to-complete metric included all of the items in this checklist.

Loss of System A checklist: Crew complement was not significant ($F(2,15) = 0.95$, $p = 0.410$) for time-to-complete the Loss of System A checklist. The SPO configuration had the most variation for this measure as the single pilot had to simultaneously maintain aircraft control, read/execute the checklist items, communicate with ATC/dispatch, and gather weather information (See Fig. 2). The overall mean time for the crews to complete the checklist was 3.3 minutes.

Manual Gear Extension Checklist. Crew complement was not significant ($F(2,15) = 0.77$, $p = 0.482$) for time-to-complete the Manual Gear Extension checklist. The SPO configuration had the most variation for this measure as the single pilot had to simultaneusly maintain aicraft control, read/execture checklist items, and manually lower the landing gear (See Fig. 3). The overall mean time for the pilots to complete the checklist was 1.8 minutes.

The number of checklist items missed was another measure of failure handling. In 18 hydraulic failure runs, only 2 crews missed checklist items for the Loss of System A checklist. One RCO crew (with Captain resting during cruise) missed two items (turning Standby Rudder On and turning Off Hydraulic Pumps) because the First Officer did not indicate where in the checklist the resting pilot (Captain) should begin. The First Officer had the checklist out in his lap and the Captain concurred it was a

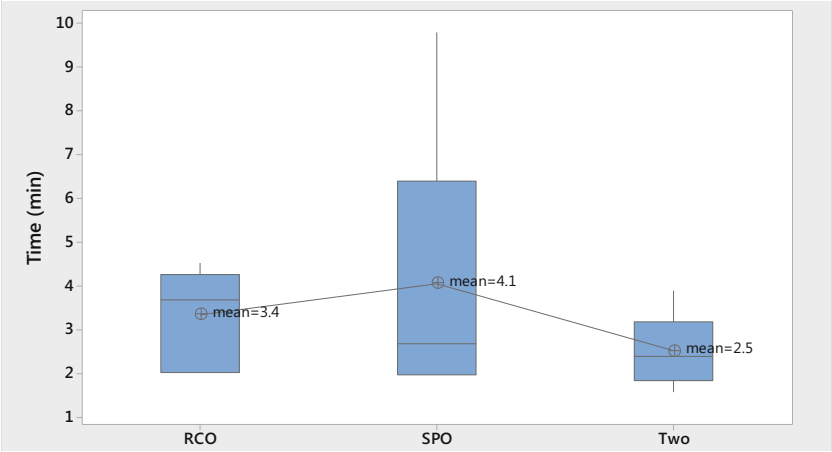


Fig. 2. Boxplot of time required to complete loss of system a checklist

Loss of System A hydraulic failure and started with the inoperative items in the checklist. The consequence of not turning the pumps off was the electric pump would continue to overheat and either burn the motor out or start a fire. Not turning standby rudder on would mean that only System B would be providing hydraulic power to the rudder. One First Officer flying in the SPO configuration missed the checklist item of turning the nose wheel steering switch to alternate. Additionally, one Captain flying in the SPO configuration missed one item, putting the landing gear lever in the down position, while completing the Manual Gear Extension checklist. This omission caused a GPWS aural warning during landing.

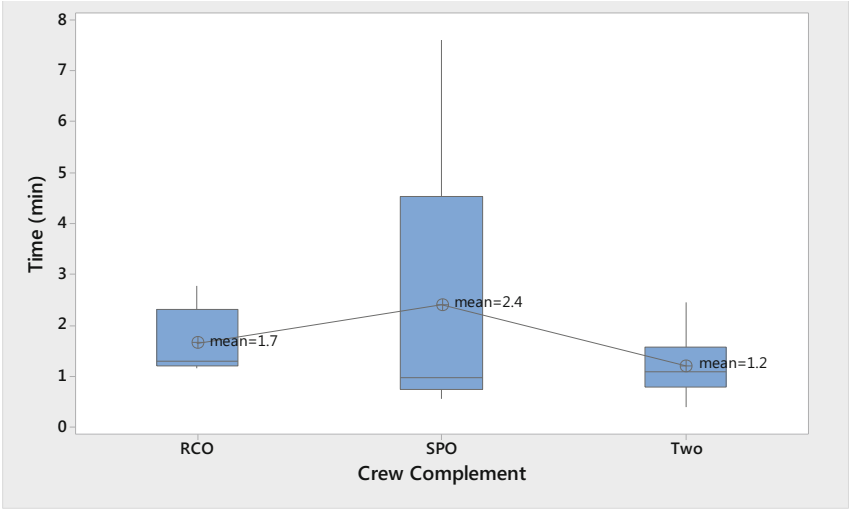


Fig. 3. Boxplot of time required to complete manual gear extension checklist

3.3 Diversion Decision

The test was staged to evaluate decision-making by the flight crew. A diversion decision after a failure was part of this decision-making test which tasked the pilots to consider distance to fly with the failure, the weather at each airport (KABQ and possible divert airports), and the time it took to troubleshoot the problem. Another factor that played specifically in the pilot's decision-making process for a System A hydraulic failure was that once the landing gear was manually lowered it could not be retracted which could make landing at an alternate airport impossible.

The reality is that when the hydraulic failure happened close to the top of descent, the best option was to continue a landing to the destination at KABQ. For the System A hydraulic failure, all crews, regardless of crew configuration, continued to the destination and landed safely. Santa Fe was the alternate airport for the flight but even though it was in the direct flight path, it had the same weather as Albuquerque, a shorter runway, and would require a steeper descent rate or ATC vectoring on a flight path that would be the same distance as going to the destination. Since Albuquerque, the destination airport, was only an additional 60 miles with better support facilities and the Flight Management System was already configured for flight to KABQ, Santa Fe was not considered a better alternative by any crews for the hydraulic failure runs.

3.4 Workload

The NASA TLX captured a subjective rating (0 [Low] to 100 [High]) of perceived task load. There are six subscales of workload represented in the NASA TLX: mental demand, physical demand, temporal demand, performance, effort, and frustration level [6]. The overall score results of this measure were examined to investigate task load variation.

Independent analyses revealed no significant ($p > 0.05$) differences between the nominal runs and hydraulic failure runs for either the PF or pilot monitoring (PM) TLX ratings. For the hydraulic failure runs, pilots rated their overall workload as being low to moderate, as reflected in the PF (median rating = 38) and PM (median rating = 26) TLX ratings.

There were no significant ($p > 0.05$) PF workload differences for crew complement. Single pilot operations were rated as having moderate workload, while crewed operations were rated as having low to moderate workload (see Fig. 4).

3.5 Safety-of-Flight

Perceived level of safety was self-assessed using a Likert type scale from 1–7, where 1 was completely acceptable and 7 was completely unacceptable.

An ANOVA revealed significant differences ($F(1,34) = 7.78$, $p = 0.009$) between the nominal runs and hydraulic failure runs for PF Perceived Safety of Flight ratings. Figure 5 illustrates these differences, where from an overall perceived level of safety for the PF compared to normal flight, this failure was difficult for some pilots as indicated by the large spread in data. Figure 6 shows the PF rating for each crew

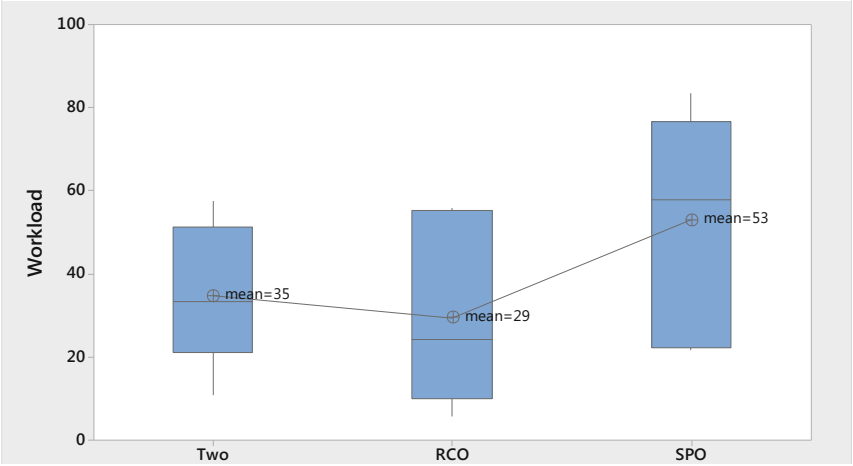


Fig. 4. Overall TLX ratings for PF hydraulic failure runs by crew configuration.

complement configuration. Pilots viewed the safety of this failure as unacceptable during single pilot operations (median rating = 5.0) where the pilot had to simultaneously maintain flightpath control, communicate with ATC/Dispatch, perform checklists, and eventually manually lower the landing gear. PF ratings indicated safety of flight was acceptable for this failure when there were two pilots to attend to it.

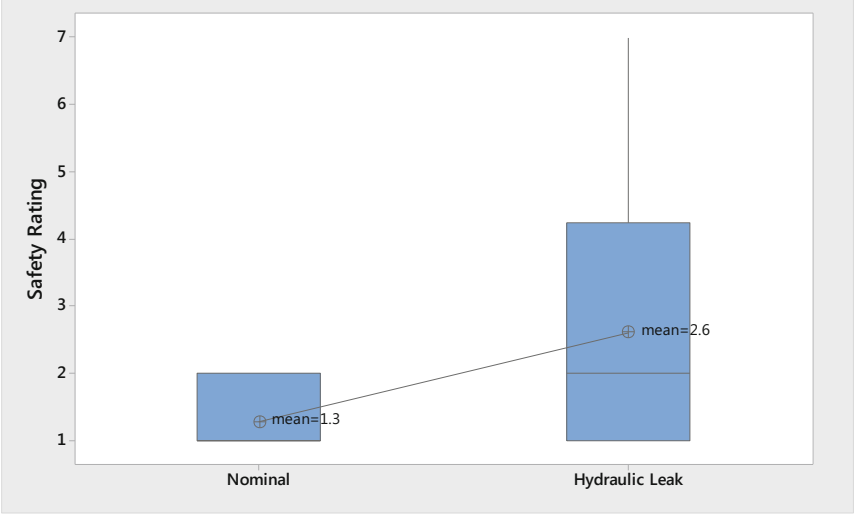


Fig. 5. Perceived safety of flight ratings for pilot flying nominal and hydraulic system failure runs collapsed across crew configuration.

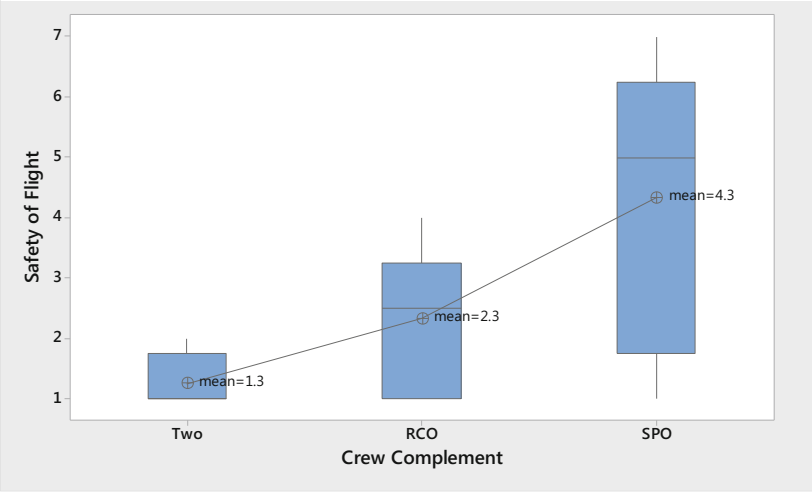


Fig. 6. Perceived safety of flight ratings for pilot-flying hydraulic system failure runs by crew configuration

4 Conclusions

This paper reflects the analysis for one non-normal scenario out of six evaluated, a hydraulic system failure, and it supports the conclusion that anything less than two crew members will require significant redesign of automation and increased levels of automation support. Time for a single pilot to troubleshoot and attend to failure of a primary hydraulic system increased three-fold and safety of flight was compromised. Workload increased to moderate levels during single pilot operations. The hydraulic failure occurred right before top of descent and initially caused the autopilot to disconnect. Two crews while flying single pilot lost airplane state awareness during the initial high workload phase of communicating the problem to ATC, finding the pertinent pages in the quick reference handbook, and trying to hand fly while configuring the aircraft. While not critical for this failure, loss of altitude awareness can quickly become catastrophic without the cross-check of a second pilot.

A potential latent failure was observed in the resting pilot re-engagement. The shared event of watching the failure develop, when cautions and warnings are observed, and the fact that the cautions are already reset before the resting pilot returns to the flight deck limits the shared knowledge of the resting pilot. Additionally, if the pilot flying has already opened the checklists and does not explicitly state what has already been done, the shared knowledge is further degraded. A resting Captain assumed that the First Officer had completed the checklist and concentrated on inoperative items without verifying correct aircraft configuration and missed a number of checklist items. Although mostly benign in this failure, the safety margins were decreased for the rest of the flight since some of the flight control redundancy was missing.

This failure occurred near top-of-descent; diversion to an alternate airport was not necessary. Workload and performance may have been optimistic considering the location of the failure.

Data analysis of the nominal runs and six failure runs is being used to establish quantitative baseline levels of performance and flight safety during nominal two-crew operations. This nominal data are being used to assess performance and safety decrement in reduced crew or single pilot operations using current-day flight deck design and certification. The nominal data are also being employed to identify and develop new applications and technology requirements for increasingly autonomous systems to assist pilots during dynamic and unplanned situations and perhaps future operations with two-crew, reduced crew, or possibly commercial single pilot operations.

Acknowledgments. This work was sponsored by NASA's Airspace Operations and Safety Program, Safe Autonomous Systems Operations project, led by Parimal Kopardekar, Project Manager. This work was conducted as part of an Interagency Agreement, No. IA1-20800, between the National Aeronautics and Space Administration (NASA), Langley Research Center, and the U.S. Department Of Transportation, FAA. This work could not have been possible without the hard work, diligence, assistance and support of Mr. Harry Hodges (AFS-440), Mr. Doug Rodzon, and the entire AFS-440 Flight Simulation Branch who were instrumental to this research. The support of Mr. Bruce DeCleene (AFS-400) was also critical and is gratefully acknowledged. The technical assistance of Dr. Robert "Buck" Joslin - the FAA Chief Scientist and Technical Advisor for Flight Deck Technology Integration - was critical to the simulation planning. Further, the help of Mr. Bryan Burks, Alaska Airlines, was instrumental in the development of our test scenarios.

References

1. Li, G., Baker, S.P., Grabowski, J.G., Rebok, G.W.: Factors associated with pilot error in aviation crashes, Aviation, space, and environmental medicine, vol. 72, No. 1, January 2001
2. Air Line Pilots Association: Air Line Pilots Association White Paper on Unmanned Aircraft Systems, April 2011
3. Federal Aviation Administration Advisory Circular: Aviation Safety Action Program, AC-120-66B, November 2002
4. Emergency and Abnormal Situations in Aviation Symposium: LOSA data and Emergency and Abnormal Situations, 10–11, June 2003. http://human-factors.arc.nasa.gov/flightcognition/download/EAS_Symposium_Presentations/LOSA.pdf
5. Etherington, T.J., Kramer, L.J., Kennedy, K.D., Bailey, R.E., Stephens, C.L.: Quantifying pilot contribution to flight safety for normal and non-normal airline operations. In: Paper Presented at AIAA Digital Avionics Systems Conference, October 2016
6. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, pp. 139–183. North-Holland, Amsterdam (1988)

Perceptions and Affective Responses to Alternative Risk-Based Airport Security

Kenneth D. Nguyen(✉) and Richard S. John

Department of Psychology, University of Southern California,
Los Angeles, CA, USA
{hoangdun,richardj}@usc.edu

Abstract. The current research examines how American air travelers perceive various risk-based airport security screening policies that vary in terms of selection procedure and agency. Respondents were randomly assigned to one of six experimental conditions that differ in terms of procedures for selecting passengers for enhanced screening. Respondents were presented with a conventional security option and a risk-based option, and were asked to rate these policies in terms equity, safety, and convenience. They also indicated anticipated feelings if selected for an enhanced screening in the risk-based procedure. Results suggest that the conventional approach was perceived as safer and more equitable but less convenient. Importantly, while different passenger selection procedures for enhanced screening led to distinct perceptions and feelings, respondents were indifferent between an equivalent selection procedures conducted by humans versus by computer.

Keywords: Airport security policy · Passenger screening · Risk based security · Randomized security · Behavioral detection · Profiling · Human–computer agency

1 Introduction

Conventional wisdom suggests that security should be omnipresent. That is, an effective airport security measure should be pervasive and exhaustive, such that most, if not all travelers and their luggage are screened and checked thoroughly. Yet, security resources are limited. The U.S airports service more than 895.5 million of passengers in 2015 [1], and it is challenging to screen each of these travelers thoroughly. *Risk-based security* emerges as a promising alternative [2]. In this approach, scare security resources are efficiently assigned to areas in which the level of security threat is high and to passengers who are associated with a high threat level. Selective security screening is an example of this approach. For instance, the current TSA Pre identifies and requires a subgroup of passengers to undergo a more vigorous screening at security checkpoints.

Although risk-based security promises to enhance security efficiency, it presumes that individual travelers understand and support this approach. Yet, public perceptions are not always aligned with policy makers' visions. For example, one of the concerns with risk-based policy is that the public may view it as less safe than a conventional security policy. Indeed, a risk-based screening that probabilistically selects air travelers

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_3

for security checks is perceived as being riskier than a traditional security that searches everyone [3]. Such misperceptions can lead to unfortunate consequences. For example, travelers may opt for riskier transportation modes when they do not feel that the risk-based approach adequately protects them from security threats [4, 5]. Given the importance of understanding how the general public perceives and reacts to risk-based security, we investigate how a sample of Americans perceive various versions of risk-based security screening procedures. In particular, we explore how American air travelers perceive different risk-based screening policies that vary in the method of selection of passengers for enhanced screening and differ in terms of the intelligent agent, i.e. human versus computer, that decides who will be selected.

2 The Effects of Selection Procedures on Perceptions of Equity and Safety

Risk-based security is not without controversy. One key concern is that the TSA has relied on factors whose validity has not yet been established to classify high-risk air travelers [6]. Three examples of risk-based screening that have generated a great deal of public discussion are selection based on behavioral profiling, individual characteristics, and randomization. Although the explicit use of profiling is prohibited by law, demographic variables such as age and sex have been used in some screening policies. Indeed, the current policy allows passengers 75 years of age or older to keep their (light) jackets, belts, and shoes on when undergoing a security screening. Behavioral screening is another selection procedure that could be applied [6]. The rationale behind this technique is that potential terrorists who disguise themselves as regular air travelers may exhibit certain emotional responses such as anxiety, fear, and distress. Through extensive training, behavioral detection officers are believed to be able to detect micro changes in the facial expression of these potential terrorists, and apprehend these suspects before they can execute an attack. A third selection procedure is randomized security screening. A limited version of randomized screening has been applied for passengers who currently participate in the Trusted Travel Program. Recall that passengers who sign up for this program undergo an expedited screening. However, TSA officers still randomly select some of these travelers for additional screening.

Although the three screening policies result in the same inequitable outcome, i.e. only a selective group of travelers will undergo an enhanced screening, the three selection strategies are distinct in the procedures employed to select travelers for an enhanced screening. One interesting question is whether travelers perceive these selection strategies as equally effective in identifying high-risk passengers. A second question is whether travelers perceive these selection strategies as equally fair to all passengers. The latter question was indirectly addressed in a recent study in which air travelers were asked to make trade-offs for equity when the levels of equity were manipulated [7]. Respondents in this study were presented with two airline options: The first employs a conventional search-all policy whereas the second adopts a risk-based screening policy. The latter one is less fair than the former one in the sense that only some passengers will incur the cost of additional security searchers but it is also less costly, more convenient, and safer than the conventional security. Importantly,

the selection strategy used in the risk-based screening was manipulated. Results indicated that respondents were willing to pay more, wait longer in a security line, and give up some degree of safety and convenience to avoid a risk-based screening that selects travelers based on their individual characteristics such as sex, age, and nationalities compared to an enhanced security risk-based screening procedure that randomly selects travelers.

Even though these results imply that respondents perceived the selective screening based on individual characteristics as unfair, this implication is limited because perception of equity was not measured. The current research remedies this limitation by exploring how American air travelers perceive the level of equity in alternative risk-based screening policies that vary in their selection strategies. A possible result is that air travelers perceive selection for enhanced screening based on individual characteristics as worse (in terms of equity) than randomized screening, a finding consistent with the previously discussed study. Yet, people may still perceive biased patterns in a perfectly random process [8, 9], which suggests that individuals may believe and perceive a random selection process as equally or even more biased than a selection process based on individual characteristics. Likewise, it is unclear how air travelers perceive the level of risk or safety across different risk-based screening options. A possible result is that travelers may view the individual profiling as more effective in identifying terrorists than a randomized screening. This hypothesis, indeed, is the exactly the argument that has been used to justify a temporary travel ban against travelers from specific Muslim-majority countries [10]. Yet, an alternative and equally plausible pattern is that randomization is perceived as a superior strategy. One rationale for this hypothesis is that the randomized search process will deter adversaries from attacking the system because additional uncertainties make it difficult for attackers to observe weakness in the system and identify a preferred attack paths [11].

Given the conflicting nature of these hypotheses and the limited research on the topic, rather than testing a priori hypothesis, the first experiment aimed to explore how people judge the levels of equity and safety in different risk-based screening procedures that differ in their selection strategy for enhanced screening. The first experiment also explores how a sample of American travelers evaluates the convenience level for different selection procedures for enhanced screening.

3 The Effect of Security Automation on Perceptions of Safety and Equity

Because intentionality plays a key role in forming fairness attribution [12, 13], we expect that perception of equity is also a function of perceived intention. Because intentionality is judged differently depending on the type of intelligent agents conducting the action, it is expected that perception of equity also depends on the type of intelligent agent conducting the security checks. Although traditional security screening is implemented by humans, recent prospective security screenings are built around the concept man-machine system. Indeed, the TSA houses its own human factor division to study, examine, and develop integrated security prototypes that facilitates collaboration between human security personnel and computerized systems [14].

Table 1. Descriptions of the experimental contexts

Agent	Screening procedures	Coastal airline	Pacific Airline
Computer	Individual profiling	…some passengers will be selected for a more thorough screening based on their characteristics…The selection process is implemented **by a computer**…	…all passengers and their carry-on baggage are screened thoroughly.
	Behavioral selection	…some passengers will be selected for a more thorough screening based on behaviors that indicate malicious intent…… The behavioral selection is implemented **by a computer**…	…all passengers and their carry-on baggage are screened thoroughly.
	Randomization	…some passengers will be selected for a more thorough screening based on randomization…The randomized selection process is implemented **by a computer**…	…all passengers and their carry-on baggage are screened thoroughly.
Human	Individual profiling	… (Same as above)…The selection process is implemented by **a security officer**…	…all passengers and their carry-on baggage are screened thoroughly.
	Behavioral selection	… (Same as above) … The behavioral selection is implemented by **a security officer**…	…all passengers and their carry-on baggage are screened thoroughly.
	Randomization	… (Same as above)…The randomized selection process is implemented by **a security officer**…	…all passengers and their carry-on baggage are screened thoroughly.

Given the advances in autonomous security, the critical question is how air travelers perceive autonomous or semi-autonomous security screening systems. In particular, how do perceptions of safety and equity vary when the automation feature is combined with various selection strategies? Thus, the second aim of this research is to further explore the role of computer automation in airport security screening. The distinction between humans and computers is an important one. While humans are more prone to error and bias and tend to apply predefined rules in a haphazard way, computers have the superior capabilities to apply rules in a consistent manner. This means that a computerized procedure may be perceived as being more consistent than human [15]. Because consistent rule application forms the basis of fairness, a computerized security procedure may be perceived as being more equitable and less biased than one based one based on human judgment.

The computer versus human agent effect on traveler perceptions of safety and risk is less clear, however. On the one hand, an automated computerized system would be expected to have superior performance over humans because of its ability to apply the same set of presumably valid screening criteria consistently over time, without being prone to bias and without subject to fatigue. Thus, air travelers may perceive a computerized security system as being more effective and safer. On the other hand, an equally plausible hypothesis is that travelers view a security procedure implemented by security personnel as being safer than an autonomous computerized security system. This could be due to the lay belief that humans are uniquely capable of detecting micro facial and behavioral expressions, e.g. fear, anxiety, and stress, which a terrorist may express when carrying out his mission. Because these human behaviors are subtle, a computerized security program may easily overlook relevant cues and allow an attacker to pass through the security checkpoint. Furthermore, while anautonomous system is pre-programmed, human detectors may be perceived as more flexible and able to detect new and emerging risk factors that have not been programmed in the computer system. This is particularly important because security threats are constantly evolving over time. Terrorist adversaries can observe a screening infrastructures put into place by a defender, and adaptively modify an existing attack plan, or devices new attack plan [11]. Despite these conflicting hypotheses, there has been little research to address this issue.

In summary, this study was designed to explore the effects of different risk-based screening procedures and the influence of security automation on perceptions of equity and safety (and convenience). In addition, the first study also explored the role of affective response. This is because affective response, in addition to subjective perception of safety and fairness, are potentially powerful determinants of behavior. Furthermore, learning how travelers react to different alternative risk-based security screening strategies provides valuable information regarding public preferences and support for alternative selection procedures.

4 Method

4.1 Procedure

Six-hundred respondents were recruited from Amazon Mechanical Turk, and they were each paid 0.25 cents for their participation. This study is a 2 Intelligent Agent (human vs. computer)-by-3 Screening Procedure (behavior vs. profile vs. randomization) between-subject factorial design. Respondents were randomly assigned to one of the six experimental conditions. Respondents were told that they were taking part in a study to help researchers understand how travelers perceive different screening procedures at airports. They were also told to imagine that they were searching for an airplane ticket to a vacation spot within the United States, and they have narrowed down their search to two options: Pacific and Coastal Airlines. The two airlines are identical in most respects except for their security screening policies. Pacific Airline employs a conventional screening process under which all passengers are screened thoroughly whereas Coastal Airline utilizes a risk-based screening process under which

all passengers simply walk through a screening device but some passengers will be selected for additional screening. Respondents were asked to contrast the screening procedures of the two airlines in terms of safety, (security) risk, bias, consistency, convenience, and equity. Note that the comparison of two options is similar to actual decisions in which people compare and contrast multiple travel options. The final sample included 527 respondents[1]; 47% of were male, the median age was 33 years, and 79% self identified as Caucasian. Sixty-seven percent of the respondents had flown at least once in the past 12 months.

4.2 Experimental Manipulation

The key manipulation is the description of the Coastal Airline's screening policies. Respondents in each condition received a different version of the screening policy, in which the type of intelligent agents and the type of screening procedures were varied. Table 1 is the summary of the exact wording for each version. Respondents in all conditions were asked to compare the two airlines on the following attributes: equity, safety, risk, convenience, consistency, bias, and risk. Respondents rated the relative standing of the two airlines on each of these six attributes on a bipolar scale, where seven means the attribute is extremely high in the conventional screening policy (Pacific Airline) and one means the attribute is extremely high in the risk-based screening policy (Coastal Airline). The order of the attributes was randomized to avoid order effects. Once ratings were completed, respondents were asked to choose which airline they prefer to fly.

4.3 Emotion Elicitation

To elicit emotional responses, respondents were told to imagine how they would react when they are singled out for additional, enhanced security screening when flying on the airline that has a risk-based screening policy. Respondents were asked to provide three to five words describing their feelings, and were asked to complete a brief version of the Positive Affect and Negative Affect Scale [16].

4.4 Other Dependent Measures

In addition to the attribute and PANAS ratings, respondents completed several demographic questions, including sex, age, and ethnicity. They were also asked about their awareness of several actual security programs initiated by the TSA, and their perceptions of the current screening policy in practice at U.S. airports.

[1] Seventy-three responses were removed because respondents failed an attention check question or did not indicate their Mechanical Turk IDs—an indication that the respondents did not take the survey seriously because without the IDs, we could not assign their credits or assign penalties for careless responses.

5 Results

5.1 Analytical Approach

Attribute ratings were recoded into a bipolar scale from −3 to 3, symmetric around 0. Negative ratings indicate that the risk-based screening is stronger in the attribute under study than the conventional screening whereas positive ratings indicate otherwise. A value of zero indicates that the two screening policies achieve the same level of the attribute. Examinations of reliability coefficients, Chronbach's Alpha, suggest that ratings of several attributes could be combined. Specifically, ratings for safety and risk (reverse coded) attributes were combined to create a safety index ($\alpha = 0.87$). Ratings for equity, bias (reverse coded), and consistency were combined to form an equity index ($\alpha = 0.71$). Convenience ratings remained as a separate attribute scale.

In addition, respondents' qualitative responses during the emotion elicitation were also coded. Two specific emotions, anger and shame were the two most commonly expressed, and were coded according to whether they were present or absent for each respondent. A code of 1 was given when a respondent expressed at least an anger-laden word (e.g. frustrated, irritated, and annoyed) and 0 otherwise. Similarly, a code of 1 was given when a respondent expressed at least a shame-laden word (e.g. ashamed, judged, and embarrassed).

A series of ANOVAs were conducted to examine the effects of the 3 × 2 experimental manipulation of the risk-based screening procedures on the attribute scores and affective responses. Binary logistic regressions were conducted to predict the expression of specific emotions, as well as to predict their choices between the risk-based vs. the conventional policy.

Table 2 presents the Pearson correlations among the dependent variables as well as summary statistics (on the diagonal). It is particularly noteworthy that equity is significantly correlated with safety, Spearman's $r = 0.49$, $p < 0.001$, and safety is significantly correlated with convenience, $r = -0.22$, $p < 0.001$.

5.2 The Experimental Effects on Security Attributes and Affective Responses

Series of ANOVAs were conducted with the experimental groups as the independent variables and equity, safety, and convenience, negative and positive affect, and behavioral intention as dependent variables. The tests revealed that there were main effects of Screening Procedures on perception of equity, negative affect, and positive affect, $F(2, 521) = 8.329$, $p < 0.001$; $F(2, 521) = 11.256$, $p < 0.001$; $F(2, 521) = 3.855$, $p = 0.022$, respectively. Surprisingly, there was no significant effect of agency (computer vs. human) on screening attributes and affective responses.

Follow-up t-tests with Bonferroni's correction for family-wise error rate (FWER) were conducted to explore the nature of the effects of screening procedures on equity, negative affect, and positive affect. Figure 1 illustrates these significant effects. Results indicate that the risk-based screening that applies to random selection was perceived as being more equitable than the risk-based screening that selects travelers based on their

individual characteristics, $t(348) = 3.978$, $p < 0.001$. Similarly, the randomized screening was perceived as being fairer than the risk-based behavioral screening, $t(348) = 2.711$, $p = 0.007$.

Results also indicated that being randomly selected for an enhanced screening led respondents to feel less negative than being selected based on individual characteristic, $t(333) = 3.404$, $p < 0.001$, or being selected based on behavioral and emotional expressions, $t(333) = 3.480$, $p < 0.001$. Unexpectedly, respondents felt more positive when being selected based on their individual characteristics than when being selected for enhanced screening based on their behavioral and emotional expressions, $t(347) = -2.826$, $p = 0.004$.

Table 2. Correlations among dependent variables

	Equity	Safety	Convenience	Positive	Negative	Anger	Shame	Decision
Equity	1.545 [1.249]							
Safety	0.47*	0.637 [0.760]						
Convenience	-0.09**	-0.17*	-1.347 [1.823]					
Positive affect	-0.06	0.01	0.08	2.169 [0.981]				
Negative affect	0.12*	0.06	0.16*	-0.28*	2.118 [1.159]			
Anger (1 = Presence)	0.14*	-0.02	-0.04	-0.13*	0.12*	67% [0.471]		
Shame (1 = presence)	0.11*	0.11*	0	-0.04	0.1**	0.06	0.12 [0.327]	
Decision (1 = convention)								63% [0.48]

Note: * p < .05; means and standard deviations are provided in the diagonal

5.3 Predicting Shame, Anger, and Decision

Binary Logistic Regression (BLR) was used to predict the emotions of shame (0 = absence or 1 = presence) and anger (0 = absence or 1 = presence) and the choice between airlines using risk-based screening (coded 1) versus conventional screening (coded 0). The followings were included as predictors in BLR models: the experimental group (random selection is the reference group), race (White vs. non-White), sex (males coded 1), and age. In the anger BLR model, only age significantly predicted the presence of anger, OR (odds ratio) = 0.97, 95% CI [0.96–0.99], $p < 0.001$. A one-year increase in age was significantly associated with 0.973 decreases in the odds of reporting feelings of anger. In other words, older respondents were less likely to experience anger than were younger respondents when being selected for an additional security search. For the BLR model predicting shame responses, the contrast between randomization versus profiling was significant, OR = 2.88, 95% CI [1.18–7.80], $p < .01$. Hence, the odds for

respondents in the individual profiling group reporting feelings of shame were 2.88 times higher than the odds for respondents in the randomization group reporting the same emotion. Note that the 95% confidence was wide, which was partially due to the low number of respondents experiencing shame (12%).

In the airline choioce BLR model, both sex and the contrast between Whites vs. non-Whites significantly predicted respondents' choice of the screening policy. Compared to male respondents, the odds for female respondents choosing the risk-based screening were 0.49 times smaller with a 95% CI [0.34–0.71], $p < 0.001$. In addition, the odds for non-White respondents to choose the risk-based screening were 0.49 times smaller than the odds for White respondents to choose the same screening with a 95% CI [0.30–0.79], $p = 0.004$.

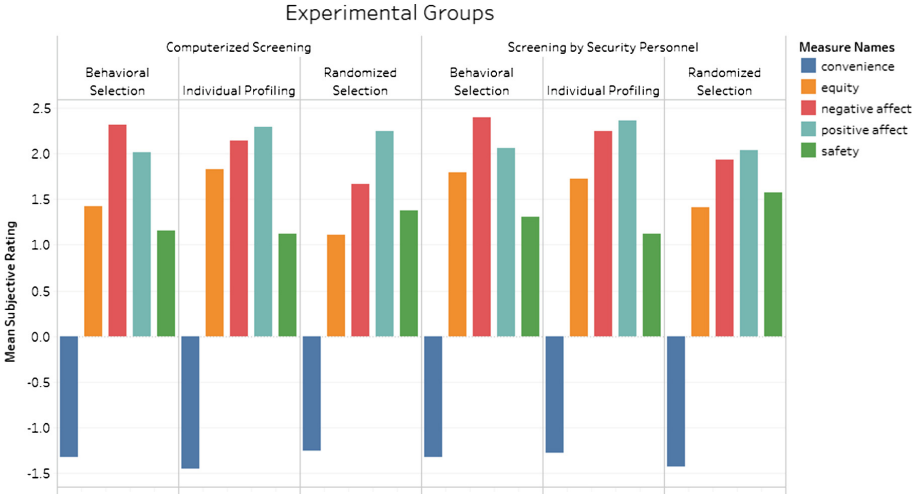


Fig. 1. Subjective ratings of attributes and affect. A higher mean rating indicates a higher level of an attribute. Negative ratings indicate the risk- based screening is stronger in an attribute compared to the conventional screening and positive screening indicates otherwise

6 Conclusion

This study investigated how people perceive various versions of risk-based airport security screening. Interestingly, while respondents perceive the three distinctive selection procedures in risk-based screening as being equally effective in enhancing airport security, they viewed the methods of selecting passengers for additional security checks based on individual profiling and behavioral screening as being less fair than a randomized selection procedure. Respondents also indicated that they would react more negatively and would be more likely to feel ashamed if they were selected for an additional search based on their personal characteristics such as sex, age, and nationality compared to a randomized selection. Most respondents indicated a preference for the conventional screening procedure over a risk-based procedure.

One of the surprising results in the first experiment is the absence of an agency effect. Although our results suggest that travelers perceive no safety or equity differences between computerized and human-based screening selection, this result may be due to the fact that power may have been limited. In particular, the experimental stimuli were somewhat subtle: The only difference between computer and human group manipulation was a single word, i.e. "computer" versus "human". Had the experimental manipulation of agency been more potent, we would have had a better chance of detecting an agency effect on perceptions of equity and safety. We are currently experiments with more potent manipulations of agency.

Public support and cooperation for passenger screening is essential for keeping air travel safe. Although the TSA has proposed and piloted a number of security initiatives, little research has been conducted to examine public support for these policies and procedures. The present research bridges this gap by showing that the traveling public does not hold a positive view of risk-based security compared to the conventional security approach. However, there may be an opportunity to enhance perceived fairness of airport security fairer by expanding on the use of randomized selection procedures. Future research that explores how various aspects of airport security measures contribute to travelers' affect, risk perceptions, and consumer choice behaviors will contribute to making air traveling safer, fairer, and more cost-effective.

References

1. Bureau of Transportation Statistics https://www.rita.dot.gov/bts/press_releases/bts018_16
2. Morral, A.R., Price, C.C., Ortiz, D.S., Wilson, B., LaTourrette, T., Mobley, B.W., Willis, H. H.: Modeling terrorism risk to the air transportation system: An independent assessment of TSA's risk management analysis tool and associated methods. Technical report, RAND Corporation (2012)
3. Scurich, N., John, R.S.: Perceptions of randomized security schedules. Risk Anal. **34**(4), 765–770 (2014)
4. Gigerenzer, G.: Dread risk, September 11, and fatal traffic accidents. Psychol. Sci. **15**(4), 286–287 (2004)
5. Gigerenzer, G.: Out of the frying pan into the fire: behavioral reactions to terrorist attacks. Risk Anal. **26**(2), 347–35 (2006)
6. U.S. Government Accountability Office. Aviation security: TSA should limit future funding for behavior detection activities: Report to congressional requesters. Technical report, United States Government Accountability Office (2013)
7. Nguyen, D.K., Rosoff, H., John, R.S.: Valuing Equal Protection in Aviation Security Screening. Risk Anal. (2017). doi:10.1111/risa.12814
8. Lopes, L.: Doing the impossible: a note on induction and the experience of randomness. J. Exp. Psychol. Learn. Mem. Cogn. **8**(6), 626–636 (1982)
9. Lopes, L.L., Oden, G.C.: Distinguishing between random and nonrandom events. J. Exp. Psychol. Learn. Mem. Cogn. **13**(3), 392–400 (1987)
10. National Public Radio. http://www.npr.org/sections/thetwo-way/2017/02/09/514384549/appeals-court-rejects-bid-to-reinstate-trumps-travel-ban

11. Jain, M., Tsai, J., Pita, J., Kiekintveld, C., Rathi, S., Tambe, M., Ordóñez, F.: Software assistants for randomized patrol planning for the LAX airport police and the federal air marshal service. Interfaces **40**(4), 267–290 (2010)
12. Falk, A., Fehr, E., Fischbacher, U.: Testing theories of fairness—Intentions matter. Games Econ. Behav. **62**(1), 287–303 (2008)
13. Nelson, W.R.: Equity or intention: It is the thought that counts. J. Econ. Behav. Organ. **48** (4), 423–430 (2002)
14. Greene, F., Kudrick, B., Muse, K.: Human factors engineering at the Transportation Security Administration. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Chicago (2014)
15. Madhavan, P., Wiegmann, D.A.: Effects of information source, pedigree, and reliability on operator interaction with decision support systems. Hum. Factors **49**(5), 773–785 (2007)
16. Thompson, E.R.: Development and validation of an internationally reliable short-form of the positive and negative affect schedule PANASJ. Cross Cult. Psychol. **38**(2), 227–242 (2007)

Autonomous Stall Recovery Dynamics as a Prevention Tool for General Aviation Loss of Control

Tiziano Bernard(✉), Lucas Stephane, and Guy A. Boy

Florida Institute of Technology, Human-Centered Design Institute, Melbourne, FL 32901, USA
tbernard2011@my.fit.edu, {lstephane,gboy}@fit.edu

Abstract. Loss of control in general aviation is responsible for 48% of reported accidents. The United States government has encouraged ulterior training in the field of stall recovery and understanding of the various factors that cause loss of control events. Although recommendations have been implemented, loss of control related accidents continue to mark a high percentage among aviation accidents. This paper proposes an enhanced autonomous loss of control recovery avionic, developed upon the Trajectory Recovery System (TRS). The newly developed system (TRS Mk.2) is described in this document with regards to its design characteristics and high-level requirements. The system will provide autonomous stall recovery, backup information to avoid workload spikes in case of system failure, and a parameter summary to inform the pilot of the recovery's performance. The system is also proposed as a means to increase pilot training, as a byproduct of recovering from real LOC events.

Keywords: LOC-S · ILOC · LOC · TRS · Stalls · Autonomous flight dynamics · General aviation · Stall prevention · Stall training · Trajectory recovery system · Human-systems integration · Aviation safety · Human-centered design

1 Introduction: Loss of Control in General Aviation

Loss of Control (LOC), or Inflight Loss of Control (ILOC), is a topic of increased attention in the realm of general aviation (GA). A continuously high level of accidents in the United States led to the National Transportation Safety Board's (NTSB) selection of *Prevent Loss of Control in Flight in General Aviation* as one of its "Most Wanted" elements on their 2016/2017 list. According to the NTSB, ILOC accounts for 48% of all general aviation accidents, and lead to 1,194 fatalities [1].

Loss of control effectively involves an aircraft abandoning controlled flight. Although there can be a myriad of causes (often vague), the resulting event is quite often a stall. A stall, in aerodynamics, is a situation in which a wing exceeds its critical angle of attack (dependent upon airfoil), causing flow separation on the suction surface of the wing (top surface) [2]. One of the challenges involved with stalls is the non-linear nature of the event. Although the lift generated by a wing behaves in a linear

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_4

manner as angle of attack increases, it enters a non-linear domain once the critical angle of attack is exceeded, rendering aircraft modeling challenging.

Stalls, as well as many other aspects of flying, fall under the definition of a complex event. Although complexity analysis is intended for operational or natural systems, a generalization could be made with regards to events. Complexity, which is a highly studied, and perhaps still not fully defined discipline, also deals with chaos theory. A chaotic system is one of which uncertainties of the initial conditions might have enormous ramifications [3]. Recall the famous butterfly example: "A butterfly flaps its wings in Africa and a tornado forms in Texas". This butterfly was, in reality, a mathematical visualization for a non-linear dynamical equation that Edward Lorenz worked was investigating [4]. Although the nature of chaos involves the "unknown", mathematical studies have proved that there is certain predictability with chaos. These predictions, often in the form of a series of consequences from a non-linear dynamical system, are called attractors [5].

An attractor in the domain of loss of control is the aerodynamic stall. A simplified LOC attractor model is shown below in Fig. 1.

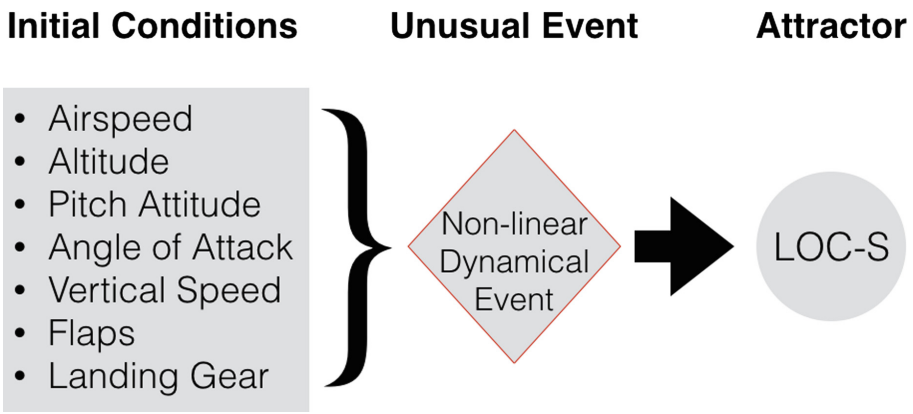


Fig. 1. The model describes how stalls are an attractor to a non-linear dynamical event whose initial conditions are many and often imprecise. The initial conditions listed above are only a few of the many that characterize an aircraft entering an unusual attitude.

The model shown above lists a short series of possible initial conditions, or variables. These variables, or rather their combination, will determine how the aircraft will results in an unusual event, called unusual attitude in aviation, where the aircraft is in an off-nominal attitude. The combination of the initial conditions will dictate exactly how the non-linear dynamical event will present itself. What can be predicted from this chaotic situation is that a typical consequence is a stall. What this means is that on a certain percentage of events, the attractor, or predictable consequence is a stall. This attractor, "loss of control due to stalls: LOC-S" is a term coined at Florida Institute of Technology to identify all those LOC situations that require stall recovery procedures. Studies in this field led to the development of the Trajectory Recovery System (TRS) [6–8].

The Trajectory Recovery System is a socio-cognitive system that provides a recovery trajectory to the pilot during a stall event. The TRS provides an alternate PFD visualization to reduce pilot workload, increase situation awareness, and optimize maneuvering time by providing a target for the pilot to follow. The system was tested and verified on a Boeing 737-800 high fidelity flight simulator [9].

2 Safety Enhancement via Training and System Knowledge

The most common approach to safety enhancement in general aviation has been training. GA is a very particular aspect of aviation, since it includes a multitude of pilots with difference training background, different flying routines, and different knowledge. Any sample population will differ among each individual component. It is therefore logical for the U.S. government to reinforce an element common to all pilots: training. This view was adopted by the European Aviation Safety Agency (EASA) by proposing enhanced stall and spin recovery training [10]. Stall training is very well covered as required by the Pilot Training Standards (PTS) [11]. The issue with this population is that many pilots will finish training and not practice recovery maneuvers for an extended period of times. The biannual flight reviews are the only required refresher for most pilots. However, a short review will not be sufficient to ensure understanding of such a vast field that is LOC-S.

The most useful performance measure during stall recovery and prevention is angle of attack. It has been long discussed how angle of attack indicators could provide sufficient information to the pilot in reducing LOC-S events [12, 13]. On a short survey sent and analyzed by the authors with nine general aviation pilots, only 3 (33.3%) recommended the use of angle of attack indicators to improve situation awareness during stall recovery. This result could be attributed to unfamiliarity with an angle of attack indicator. Pilots are acquainted with the concept of angle of attack, but are not accustomed to using it as an indicator in the cockpit. Moreover, angle of attack is difficult to associate with a visual reference. On the same survey, only 11.1% of the pilots referred to angle of attack when asked about the cues used when learning stall

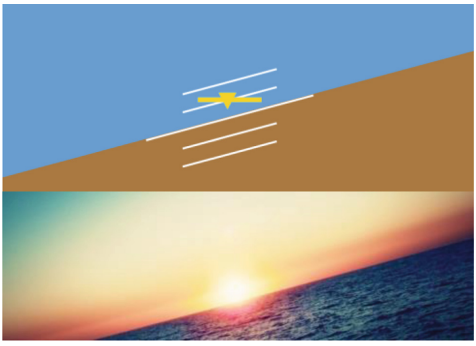


Fig. 2. The image describes how pitch attitude can easily be crosschecked with directly available visual references. In this example, the horizon is clearly visible [3]

recovery. Most answers explicitly state the lowering of pitch attitude (or lowering the nose). Only one pilot mentioned that the decrease in pitch attitude would also cause the angle of attack to decrease. This difficult association to the real world can be demonstrated in Fig. 2.

A pilot observing the horizon from the cockpit will have an idea of its attitude (pitching up, down, baking left, right, etc.). The instrument will quantify what the pilot observes (10º pitch up, 15º right bank, etc.). Angle of attack does not work the same way. Showing angle of attack on a dial or display will not necessarily represent what the pilot visualizes from looking outside. For instance, and aircraft could be flying at a 0º pitch attitude yet be at an angle of attack of 10º. There is a clear cognitive incongruence between observation and reasoning. Thus, angle of attack indicators will require extensive training in their use, and extensive training in balancing the attention dedicated to the equipment.

The NTSB's Most Wanted List, when referring to ILOC, clearly states that pilots should also be better prepared in using the available equipment [1]. This refers to causes such as weather and similar factors. However, the addition of ulterior technologies on board GA aircraft could also be a solution. Adding an angle of attack indicator will require cabin modifications, however adding a digital trajectory recovery display on an already present Primary Flight Display (PFD) could also be a viable solution (many GA airplanes built presently are equipped with advanced glass displays). It could be assumed that the application of a TRS will not be sufficient if used manually. Rarity of use and specific training requirements could render the system difficult to implement. During the final stages of the TRS development, it was considered to render the system autonomous. This would be a challenge involving complexity in automation.

3 Design Considerations in Aeronautical Autonomous Systems

The design of autonomous systems in aviation requires extensive analysis like any life-critical system. A survey conducted by Cummings et al. shows that pilots are often bothered with the question of why the automation is behaving in a certain way, or why there appears to be little to no logic regarding what is and what is not automated. Moreover, whenever an anomaly is observed, the crew is required to perform manual operations, which often results in an increase in workload [14].

The same paper by Cummings et al. presented a list of features that characterize a "good" and "bad" copilot. Although general aviation often involves single-pilot operations, knowing what a good copilot is can be useful in designing systems that should act as copilots and promote multi-agent teaming under a crew resource management (CRM) modality [15]. Among the good features were various elements that could summarized into: ubiquitous nature, context awareness, procedural knowledge, workload anticipations (Note that these are not all exact words but rather a summary compiled by the authors of this paper). Among the negative is the need to respect authority and know when to follow the "captain's orders" (disengagement criteria). These factors have been explicitly transformed into high level design requirements for the autonomous TRS system as listed in Table 1.

Table 1 provides basic high-level requirements for automating a system that was designed to be performed in a manual mode. The design requirements shown above are

Table 1. Co-Pilot feature in co-pilots transformed into high level design requirements

Co-pilot feature	Design requirement
Ubiquitous nature	The system should always be online and appear to the user only when needed
Situation awareness	The system should be knowledgeable of the aircraft's current state through data feed and analysis from airplane systems. In this way it will know when to engage
Procedural knowledge	The system should possess the procedural knowledge for the required situation
Workload anticipation	The system should provide, if needed, guidance to the user in alternate ways if the system fails, in order to avoid automation surprises
Disengagement criteria	The system should respect the pilot in command's (PIC) decisions and should be equipped with a disengagement system to allow the PIC to perform whatever he/she decides

straight forward, although some require some creative and critical thinking. The first element, ubiquitous nature, requires the system to be constantly available to the pilot when needed. There should be no "loading" or "acquiring data" time involved due to the time-critical nature of stall recovery. Situation awareness of the system is a determined by the continuous data feed from the airplane itself. This is important to assure that the system understands the current situation and provides correct information to the pilot. The content must portray the situation accurately. Procedural knowledge is key in the internal algorithmic design of the system, and is rather an aeronautical engineering problem. Workload anticipation involves creative thinking. Often, when automation fails, the pilots are required to turn to manual control that involves a spike in cockpit workload. Moreover, automation surprises often result as a consequence [16]. In the case of automation failure during stall recovery, the system should provide a "Failure Mode" in which it facilitates pilot intervention and readjustment to manual control. The failure mode visualization shown in this paper is designed to facilitate the return of the pilot into "checklist" mode. When the automation takes over recovery, the pilot does not necessary follow all the minimum changes that the aircraft is undergoing. The pilot's situation awareness becomes disembodied [16]. The challenge involves optimizing the time required for the pilot to re-embody the aircraft. The last criterion, disengagement, involves the respect of authority in the cockpit. The pilot in command must have the option to disengage the system whenever he/she pleases. The disengagement method has been idealized to be the "autopilot disengage button" on the yoke. This was chosen because it becomes a cognitive function for the pilot to "make the airplane stop doing things automatically" by pressing that specific button.

3.1 Training in Observation

The author, on a routine flight from Melbourne International Airport (KMLB) to Space Coast Airport (KTIX) during the late evening, encountered a pre-stall situation. The aircraft, a Piper Archer III (PA28-181) had three souls on board: pilot, copilot, and passenger. This aircraft is certified for single-pilot operations, however in this paper the second pilot on board will be referred to as the copilot. The aircraft took off from runway 9R and entered a climb for cruise altitude while turning northbound. The aircraft was flying at an approximate 79 knots recommended airspeed. The aircraft was still under 1000 feet altitude. Both pilots had a PFD shown on both PFD and MFD (backup display mode on G-1000 TAA glass cockpit). The PIC performed a radio call and right after the stall alarm engaged. The immediate action of the PIC was to push on the yoke and apply power (which was already at maximum given the climb and his hand constantly on it). The cause of this airspeed change is unknown. It could have been attributed to distraction (with both pilots observing a PFD is unlikely). Nevertheless the immediate PIC action resulted in an immediate recovery. The success of the event was a result of muscle memory. The PIC's flight instructor was proficient in training the PIC in responsive stall recovery actions that resulted in muscle memory. Such immediate action resulted in two consequences: the successful recovery from a stall and the lack of situation understanding.

Training is no substitute for experience. Experience in real unplanned stalls is highly dangerous, and there is a reason why stall training is so extensive. However, events that occur in a stressful situation will often be better memorized [18]. Needless to say, the author remembers the event extremely well.

There are two benefits from an autonomous recovery system: the system can recover the situation if operations are not followed optimally, and could provide summary data to make the pilot understand what failed defenses occurred for the stall to occur. The system should be able to perform this task since it is situation aware, and will provide meaning to a situation that is often unclear. Therefore, the system becomes a training device as well.

4 Autonomous Stall Recovery System: TRS Mark 2

The TRS Mk. 2, which is an autonomous and enhanced version of the TRS, presents itself in four modalities.

1. **Normal Mode:** The PFD displays standard information as during normal operations. The pilot has all information available to him.
2. **TRS Mode:** The PFD is replaced by the TRS when a stall is sensed and no pilot intervention is detected (or is found to be non-sufficient). The design rationale for this mode was already described by Kasdaglis et al. [Ref TRS].

3. **TRS Failure Mode:** The TRS fails to function, and the pilot is required to quickly take over. The system provides a stall recovery checklist on the PFD to quickly tell the pilot what to do.
4. **Recovery Summary Mode:** The normal PFD returns and summarizes recovery performance in terms of airspeed, angle of attack, and altitude loss. It also provides the trigger that caused the TRS to engage, showing the pilot the cause fo system engagement.

The visualizations are provided in a simplified PFD design in the figures below.

In Fig. 3 it is visible that the TRS mode engages by eliminating certain data from the PFD. The words "TRS MODE" appear vertically on the screen, and a text at the bottom of the screen instructs the PIC on how to disengage the system (if the PIC desires so).

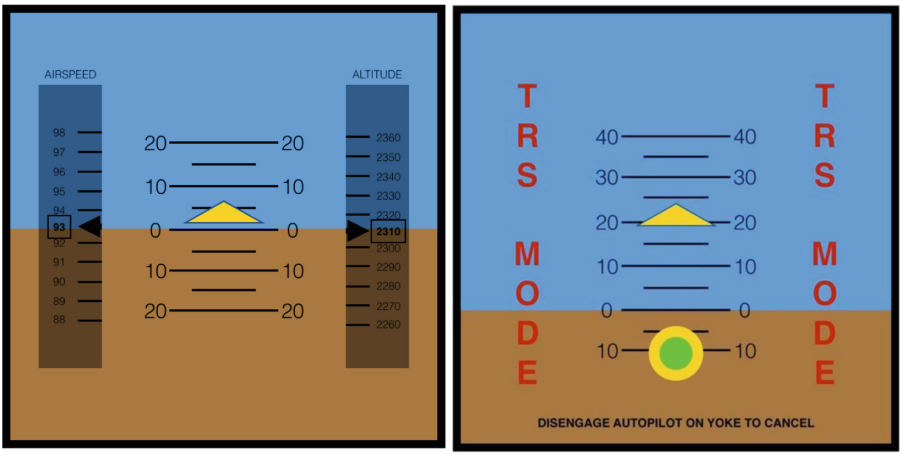


Fig. 3. PFD normal mode (left) and TRS mode (right) are part of the original TRS design. The system in normal mode represents the standard operation of the airplane's PFD. The data shown in the simplified design represents that which is usually available for a pilot during normal operations. The TRS mode eliminates all data but pitch attitude and throttle level (in the target).

In Fig. 4 the enhanced features of the TRS are shown. The TRS Failure Mode (left) occurs in the event that the TRS algorithm malfunctions and the system fails. The PFD reappears but also provides a clear checklist to show where the TRS left off and where the pilot needs to continue. This will reduce the automation surprise effect and reduce the consequent surprise workload. When the situation is resolved, the PFD will include summary data to show the pilot some performance parameters (right). These are:

1. Minimum Airspeed Achieved
2. Maximum Angle of Attack Achieved
3. Total Altitude Loss
4. Reason for TRS taking over the situation

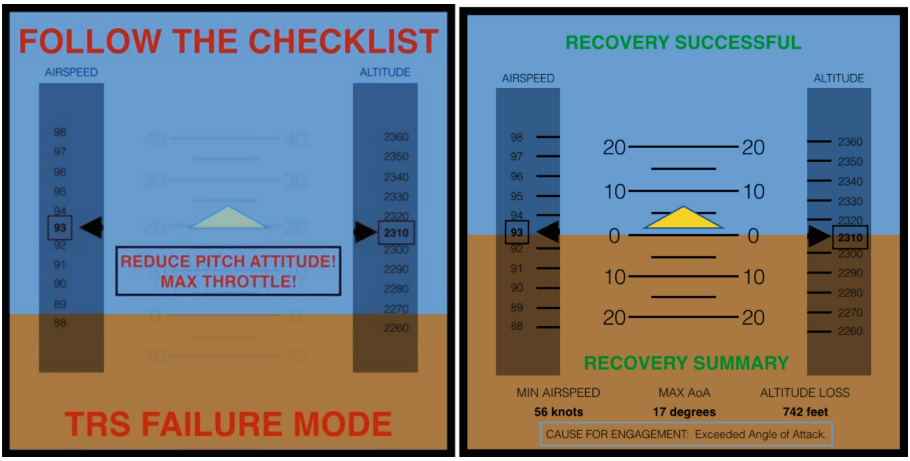


Fig. 4. PFD TRS failure mode (left) and PFD in recovery summary mode (right) are supplements to the TRS given the autonomous nature of the new version

The system also displays the words "Recovery Successful" and "Recovery Summary". Eventually the system returns to normal by displaying the PFD exactly as it should be during normal operations.

4.1 Ground Station Data Transfer for Training Purposes

The use of TRS Mk. 2 as a training device should also allow for post flight analysis. In this way, the pilot can visualize flight paths, angles, trajectories, and power levels in a stress-free manner. It is believed that providing too much performance data post-stall on board will decrease attention from normal operations and focus the pilot's attention on the full understanding of what has happened. This data can be streamed to a certain ground station that could be used by multiple entities. If the aircraft belongs to an FBO or flight school, the company will have records of how events occurred inflight. If the aircraft is privately owned, the data could be automatically downloaded on some personal device such as a tablet to be used as reference by the owner.

Of course, the data streaming is useful as it provides a record in case an accident is involved. Moreover, a LOC-S event could also be due to various factors, including medical issues. Ground stations could support in those manners, and will not require dedicated personnel. Usually a dispatcher will have extensive flight experience, and an extra monitor at flight operations could be sufficient. After all, data is not required to be always available. It could help though when something anomalous is occurring.

5 Conclusive Remarks

The Trajectory Recovery System Mark 2 is designed to provide socio-cognitive stability. Only through this type of stability will a user truly reach the optimal level of workload, situation awareness, and activity execution. Because the flight deck is already in existence, the TRS Mk. 2 is designed for being integrated via the systems-of-systems (SoS) approach.

This paper described how the design criteria reflect high situation awareness, low workload, appropriate information management and an effective operational use. This system is created as a direct result of life-critical needs and requirements. Moreover, for an excellent level of human-systems integration, the human-centered design process is carried throughout the future development, integration, and testing phase of the product lifecycle.

The TRS Mk. 2 will enhance aviation safety through the use of intelligent autonomy and carefully considered human-machine interaction. The system is intended specifically for the general aviation world, since so many aviation related accidents occur due to loss of control. This system will not only allow authority management under the pilot's control, but also engage when the situation is deteriorating past a certain level.

Finally, the system is designed to provide summary performance data, to enhance real-life scenario awareness, a sort of training under non-training circumstances.

References

1. National Transportation Safety Board: NTSB 2017-2018 most wanted list of transportation safety improvements. In: National Transportation Safety Board. https://www.ntsb.gov/safety/mwl/Documents/2017-18/MWL-Brochure2017-18.pdf. Accessed 2016
2. Milne-Thomson, L.: Theoretical Aerodynamics, 1st edn. Dover Publications, Inc., New York (1973). (All rights reserved under Pan American and International Copyright Conventions)
3. Mitchell, M.: Complexity A Guided Tour 1. Oxford University Press Inc., New York (2011)
4. Lorenz, E.: Deterministic nonperiodic flow. J. Atmos. Sci. **20**, 130–141 (1963)
5. Boy, G.: Orchestrating Human-Centered Design, 1st edn. Springer, London (2013)
6. Kasdaglis, N., Bernard, T., Stephane, L., Boy, G.: Affordant Guidance for In-Flight Loss of Control: The Trajectory Recovery System (TRS). HCI Aero, Paris (2016)
7. Kasdaglis, N.: Trajectory recovery system: angle of attack guidance for inflight loss of control. In: Harris, D. (ed.) Conference Proceedings on Human-Computer Interaction International - Engineering Psychology and Cognitive Ergonomics. Lecture Notes in Artificial Intelligence, Toronto, pp. 397–408 (2016)
8. Bernard, T., Kasdaglis, N., Rolins, A., Troshchenko, A., Stephane, L.: LOC-S: Improved model and control algorithm for a stall recovery on-board avionics system. In: IEEE Aerospace Conference, Big Sky (2017, in press)
9. Troshchenko, A., Kasdaglis, N., Bernard, T., Stephane, L.: Development of an OpenGL stall recovery system in a restricted-resource being 737 simulator with external data feed for flight testing. In: IEEE Aerospace Conference, Big Sky (2017, in press)

10. European Aviation Safety Agency: Notice of Proposed Amendment 2015-13: Loss of Control Prevention and Recovery Training. Proposed Amendment, EASA (2015)
11. Flight Standard Services: Private Pilot For Airplane Single-Engine Land and Sea Practical Test Standards. Aviation Supplies & Academics Inc., Washington (2011)
12. Forrest, F.: Angle of attack presentation in pilot training. Final Report, Department of Transportation Federal Aviation Administration (1969)
13. Pope, S.: Flying angle of attack - flying with the technology will improve safety. In: Flying Magazine. http://www.flyingmag.com/technique/tip-week/flying-angle-attack. Accessed 11 Feb 2014
14. Cummings, M., Stimpson, A., Clamann, M.: Functional requirements for onboard intelligent automation in single pilot operations. In: AIAA Infotech @ Aerospace, San Diego (2016)
15. Taylor, A., Cotter, T.: Human opinion counts-making decisions in critical situations when working with highly automated systems. In: Long, S., Ng, D.C., Nepal, B. (eds.) American Society for Engineering Management (2016)
16. Taylor, A.: Human-machine interaction in the cockpit and applicable design changes towards better collaboration. In: AHFE 2016 International Conference on Human Factors and System Interactions, vol. 497, pp. 271–279. Springer, Orlando (2016)
17. Boy, G.: On the complexity of situation awareness. In: 19th Triennal Congress of the IEA, Melbourne, Australia (2015)
18. McIntyre, C., Roozendaal, B.: Adrenal stress hormones and enhanced memory for emotionally arousing experiences. In: Bermudez-Rattoni, F. (ed.) Neural Plasticity and Memory: From Genes to Brain Imaging, 1st edn. CRC Press/Taylor & Francis, Boca Raton (2007)

Planning of General Aviation Pilots Using Interviews

Elena Psyllou(✉), Arnab Majumdar, and Washington Ochieng

Centre for Transport Studies, Imperial College, London, UK
elena.psylloull@imperial.ac.uk

Abstract. This paper determines the manner in which general aviation (GA) pilots plan their flight in ordinary conditions, i.e. weather conditions suitable for visual flight rules. Twenty-seven semi-structured interviews were conducted with recreational pilots who flew a fixed-wing motor, glider and ultralight aircraft in Finland, Norway and United Kingdom, nations with major GA operations. Thematic analysis was used. Due to the careful design of the sample, distinctive differences were determined between pilots of the three aircraft designs in the following three aspects: the purpose of and the time that the pre-flight planning and the features that are most likely to be considered. Whilst most pilots of motorised aircraft started planning the flight the evening before departure, pilots that were familiar with the area or used emerging planning technologies started planning closer to take-off. The findings implicate that the diversity of GA must be considered in any safety and operational analysis.

Keywords: General aviation · Pre-flight planning · Interviews

1 Introduction

General aviation (GA) represents various flight operations, including recreational flights, for which a diverse aircraft fleet is used [1, 2]. Fixed-wing motorised aircraft dominates recreational flights [3]; however, pilots of aircraft of lighter weight and smaller size, known as ultralight aircraft, have also increased in numbers in recent years given their lower costs due to less rigorous regulations, e.g. a transponder is only optional on an ultralight aircraft in Europe. Typically, most GA pilots fly for recreation purposes at the weekends and when the weather conditions meet the requirements for a visual flight rule (VFR) flight.

The planning of the GA flight has been studied over the last 20 years; however, it focused primarily on adverse weather decision making. Planning undertaken in good weather conditions has not been excessively studied, e.g. the features that influence the flight-route decision as well as the time that the pre-flight planning starts in relation to the departure. Furthermore, pilots increasingly use emerging flight planning apps, e.g. SkyDemon [4], that might have changed the flight planning in a similar manner as that of the use of Global Positioning System (GPS) [5].

This paper, therefore, studies the manner in which GA pilots plan the flight in ordinary weather and determines any distinctive differences between pilots that fly the most common types of GA aircraft: fixed-wing motorised, glider and ultralight aircraft.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_5

In order to do this, a series of structured interviews were conducted with such pilots. This paper is structured as follows. In Sect. 2, the literature regarding a GA pilot's decision making in planning is reviewed regarding the data, method and key findings. In Sect. 3, the participants, the interview design and the method to analyse transcripts used in this paper are outlined. In Sect. 4, the results are presented concerning the purpose of the pre-flight planning, the time that pilots start the planning and the features that are most likely to be considered. The key features of planning are discussed highlighting key differences between the groups of pilots, before concluding.

2 Literature Review

Planning has been extensively studied from the decision-making perspective in order to determine how pilots make decisions pre-flight and in-flight in adverse weather conditions. Such studies have found that pilots utilise information differently pre-flight and in-flight in simulation flights [6]. In particular, pilots that consider more information are more likely to make dichotomous decisions pre-flight, e.g. Go/no go, whilst they are more likely to continue the flight into adverse weather. However, pilots who use less information pre-flight typically select an ambivalent option. A reason for this can be the insufficient information available. This shows that each pilot utilises the information differently and addresses the insufficient information differently as well. Such findings might be evident for any situation including ordinary weather conditions.

The use of GPS has had both positive and negative impacts on pilots [5]. Pilots that use the GPS can become overconfident of their abilities and proceed in adverse weather conditions, which is an abnormal situation for them. Such pilots are willing to make decisions beyond their comfortable zone. It is evident that pilots trust the GPS to provide them with further information about the position of the aircraft and thus, they believe that they can maintain their situational awareness. Similar to the GPS, planning apps, e.g. SkyDemon, can provide further information to pilots pre- and in-flight and similar implications should be expected from the use of the emerging planning technologies. For instance, pilots make a decision in-flight rather than pre-flight when there is sufficient information, e.g. actual than forecast weather. The impact on planning of such technologies has not been extensively studied.

The sample size used in previous simulation studies and questionnaires was selected based on the pilot's licence (e.g. private and commercial), rating (e.g. VFR) and flying experience measured in flying hours. Whilst these are important, features of equal importance were not considered in these studies, e.g. purpose of flying (e.g. recreation or paid flight), the length of the flight (e.g. local, cross-country flight) and aircraft design. For the latter, for instance, gliders fly in a different style from motorised aircraft as the pilot solely depends on the wind to stay up in the area. Such differences were ignored due to the sample design.

In conclusion, planning by GA pilots has been extensively studied for flights in adverse weather conditions, e.g. weather features that GA pilots consider and for changes in the pre-flight planning due to the use of the GPS. How pilots plan a flight in ordinary weather conditions was not specifically studied. In such conditions, pilots are expected to consider aspects other than the weather and consequently their planning will be different

from the adverse weather conditions. There will also be differences in the planning between pilots of aircraft with different designs. In order to explore the planning undertaken by GA pilots, interviews of GA pilots were conducted and the selected participants represented the diversity of GA sector. The sample, the interview design and the method of analysis of the interviews are presented in the following section.

3 Method

Interviews of recreational GA pilots were conducted in Finland, Norway and United Kingdom (UK), three nations with major GA operations, between March and November in 2015. Interview duration was between 45 and 70 minutes and all were recorded. Participants were sampled directly from flying clubs in the UK and through the airspace navigation service provider and national aviation authority in Finland and Norway.

3.1 Participants

The participants were selected based on six criteria in order to capture the diversity of GA as shown in Table 1.

Table 1. Selection criteria of the sample

Criterion	Selection choice
Country that pilots are based	• Finland, Norway and UK, three nations with major GA operations
Flying base of pilots	• 80% of participants are based in the capital of the country • 20% of participants are based in a location other than the capital
Purpose of flying	• Primarily fly for private purposes • The private pilot's licence was issued at least a month before the interview
Aircraft type	• 80% of the participants fly a fixed-wing motorised aircraft • 10% fly an ultralight aircraft • 10% fly a glider aircraft
Flying activity of pilots	• A VFR rating • The minimum total flying hours is 80 • Participants are active in the last three months on the date of the interview unless they fly in country with non-flyable weather in winter • 40% of participants fly cross-country flights
Pilot's language	• English

The sample consists of 27 GA recreational pilots as shown in Table 2. There were 20 pilots that flew a fixed-wing motor aircraft, three pilots that flew a glider and four pilots that flew an ultralight aircraft. Pilots who flew either an ultralight or a glider were difficult to find and thus, the minimum required number of these pilots was selected. Similarly, in Norway only participants that were based in the capital were found.

Table 2. Sample design

Criterion	Fixed-wing motor aircraft	Glider aircraft	Ultralight aircraft	Total
Country				
Finland (Helsinki)	6	1	1	10
Finland (Southern Finland)	2			
Norway (Oslo)	5	1	1	7
United Kingdom (Greater London)	5	1	2	10
United Kingdom (South England)	2			
Total	20	3	4	27
Total flying hours[a]	505.5 (1310.7)	500 (822.7)	100 (35)	N/A
Flying hours in the last three months*	14.5 (24.3)	8 (24)	11 (1.5)	N/A
Number of pilots who also flew cross-country flights	17	1	0	N/A

[a]Mean (standard deviation), N/A: not applicable

3.2 Interview Design

A semi-structured interview was designed to address research questions beyond the research question of this paper. The first three questions of the interview were related to flight planning and are of relevance to this paper. In the first question of the interview, participants were asked to describe the selection of a flight route for a desired origin and destination. For this flight route, pilots were asked to describe the difficulties that they expected to face. In the third question, they described the material and tools used for planning and navigation. The questions were open-ended and probe questions were asked, e.g. "will the temperature affect your flight route decision?"

3.3 Analysis of Interview

The interview transcripts were analysed using the phenomenological method of thematic analysis [7]. The transcripts were coded and the codes were subsequently grouped to develop the themes and their sub- themes that formed the basis of their findings of their analysis. The analysis itself followed the guidelines for the 'Publication of Qualitative Research Studies in Psychology and Related Fields' [8]. The qualitative data analysis and research software 'ATLAS.ti' was used to analyse the interviews. The results were validated by a subject matter expert, who had ten years of expertise in aviation safety and interview analysis and the agreement was 90%, which was higher than the threshold, i.e. 85%.

4 Results

Distinctive differences were determined between the pilots of the three aircraft designs in the following three aspects: the time that pre-flight planning starts with respect to departure, the purpose of the pre-flight and in-flight planning and the features that are most likely to be considered.

The main purpose of the pre-flight planning is to determine if the flight can be conducted mostly with respect to the weather and the choice of a destination. Whilst this is common between the three aircraft designs, the expectations from a pre-flight planning differs between motorised and unmortised aircraft, i.e. glider. Pre-flight, pilots of a glider aircraft study the weather in order to decide the departure time, the overall direction of the flight and the range of the desired flying altitude. The change of the aircraft's heading is decided in-flight on the basis of the weather conditions experienced, e.g. where there is a lift. On the other hand, pilots of motorised aircraft make specific decisions in order to estimate the flight duration and fuel consumption. Such pilots, pre-flight decide the exact flight route and correct the aircraft's heading given the weather report. It is worth noting that the pilots of ultralight aircraft plan their route in less detail given that they often fly in uncontrolled airspace and their aircraft design is such that such detailed calculations are not possible, e.g. three-axis ultralight aircraft. Furthermore, they often fly for shorter distances than heavier aircraft.

The time that pre-flight planning starts with respect to the departure differs between pilots and aircraft designs. Five planning strategies were identified as shown in Table 3. Half of the pilots of fixed-wing aircraft started planning their route during the evening before their planned departure day (P-1). However, pilots can start planning closer to the take-off in the situation where they are familiar with the area (P-2) and have already decided the destination and departure time (P-3). Pilots may not plan at all or might do a few calculations or use a previous plan as well. Planning a repeated leg of a flight can also be less detailed.

Table 3. Pre-flight planning strategies

No	Description	Fixed-wing motor aircraft	Ultralight aircraft
1	Planning starts the evening before the planned departure day	10	2
2	Planning starts just prior to take-off	2	1
3	Planning starts when the pilot is already on the aircraft	2	
4	Planning is not done and pilots might do a few calculations or use a previous plan	4	1
5	Planning while airborne given that planning was completed before take-off	2	

For local flights, pilots of the ultralight aircraft are familiar with the area and are less likely to plan the flight in detail when compared to a cross-country flight. Given that glider pilots search for the areas where there is sufficient lift for their aircraft while airborne, their pilots make more flight-route decisions in-flight than pre-flight.

Planning has changed due to the use of emerging planning and navigation technologies. Planning using apps significantly reduces the time needed for planning, from a few hours to a few minutes because all the calculations, including the aircraft's headings, are made automatically. Hence, pilots can select a route while already they are sitting on the aircraft ready for take-off. If internet connectivity is available in-flight, pilots will also modify their selected route given any weather updates. In addition, pilots might deliberately make more flight route decisions in-flight due to the ease of planning. For instance, the originally planned flight route might change in the situation where the request to fly in controlled airspace was rejected, i.e. the pilot re-plans the flight so that the aircraft flies in uncontrolled airspace rather controlled airspace.

Furthermore, the use of the GPS receiver independently of that installed in a mobile computer, e.g. tablet, has changed the manner by which pilots identify the aircraft's position in that pilots look for coordinates rather than landmarks. As a consequence, pilots use fewer landmarks during their flight route and this results in fast planning as pilots select fewer reporting points. This also encourages pilots to change the route in-flight.

The two features that GA pilots are most likely to consider in pre-flight planning are the weather and the airspace design. The former determines the extent to which the flight will be conducted and further determines the flight route chosen and the destination. For example, pilots that fly in an open-cockpit ultralight aircraft avoid flying when it rains. Whilst rain is not of a great concern for fixed-wing motorised pilots, they will not fly in areas where the cloud ceiling is lower than the minimum required for VFR flights. Wind is also important as it can stray the aircraft without the pilot's notice. It should be noted that the weather forecast is known only where weather stations, i.e. airports, are located. Thermals can also change an aircraft's heading and are not included in the weather report. Hence, pilots must rely on their knowledge and experience, e.g. in the summer.

The pilots of all aircraft studied the airspace design. Pilots of glider and ultralight aircraft selected a route in uncontrolled airspace because their aircraft were ill-equipped to fly in controlled airspace. On the other hand, pilots of fixed-wing motorised aircraft also select such a route because they either want to avoid contact with the ATCO or their request for entry into controlled airspace may be rejected. Such a situation often occurs at an altitude up to 6000ft. In the capital, all pilots study the altitude and shape of the horizontal boundary in order to select the best optimal altitude. Pilots also study the vertical boundary. They often select a route that is at an offset to the controlled airspace as shown in Fig. 1. Such a flight route might also be encouraged by the trust laced by the pilots in the navigation tools to accurately and precisely provide the aircraft's position.

Pilots also consider the traffic volumes and diversity in uncontrolled airspace. For instance, on days with good weather, especially at weekends, participants expect a high volume of GA traffic along their route and near airfields. These pilots are also concerned with glider and ultralight aircraft because these aircraft are difficult to see and are ill-equipped to be contacted via radio in the event of a conflict. Ultralight pilots are also concerned with all traffic as they depend solely on their eyes to see any aircraft on a collision path and are not equipped with tools, e.g. a radio, to resolve the problem. Such pilots fly around areas of high density and gliding areas whenever possible as shown in Fig. 1.

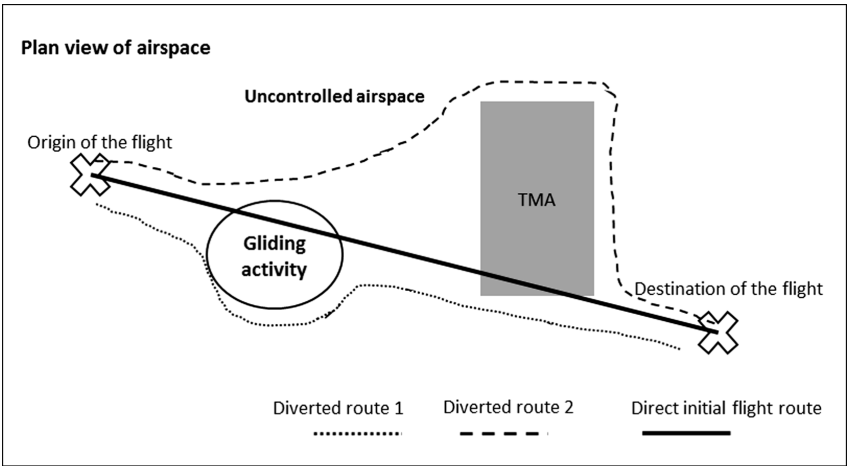


Fig. 1. Diverted flight route

5 Discussion

Due to the careful design of both the questions and the sample, distinctive differences were determined between the pilots of the three aircraft designs in the following three aspects: the purpose of the pre-flight, the time that pre-flight planning starts with respect to the departure and the features that are most likely to be considered.

Pre-flight, the pilots of fixed-wing motorised aircraft study the wind and correct the heading of the aircraft whereas the pilots of gliders study the weather to decide the overall direction of the flight, e.g. eastbound from departing airfield. The headings of the glider are estimated in-flight based on the lift. Pilots of fixed-wing aircraft are likely to start planning their flight the evening before the planned departure day. However, pilots can also select a flight route a few hours before departure or even while they are sitting on the aircraft ready for take-off. These two behaviours are evident in pilots who are familiar with the area they fly. Pilots also make more flight-route decisions in-flight due to the use of emerging planning and navigation technologies. Such technologies also contribute to ensuring that planning occurs close to the departure.

As for the features that studied by the pilots in planning, the two most common are the weather and airspace design. Weather is a key feature as it determines the extent to which the flight will be conducted, the destination and the flight route. For open-cockpit ultralight aircraft, strong wind or rain can degrade both a pilot's and the aircraft's performance. Furthermore, weather conditions beyond the VFR requirements result in a change of plans, e.g. cancelation of the flight, and in a diversion of the flight route and the destination airborne in order to accommodate the adverse weather. Whilst adverse weather is crucial, ordinary weather can also result in minor but essential deviation from the planned route. For example, thermals and a stronger than forecast wind can offset the flight route without the pilot's notice. The importance of weather has also been established in the literature.

Airspace design has found to be important in pre-flight planning for the first time. Pilots mostly fly in uncontrolled airspace. Amongst the reasons is that glider and ultralight are ill-equipped to fly in controlled airspace. However, fixed-wing aircraft meet the technological requirements and can consider flying in controlled airspace. However, there are pilots who prefer to fly in uncontrolled airspace. Pilots of all the aircraft designs are willing to modify their flight route to avoid congested areas. Pilots of fixed-wing motorised aircraft want to avoid flying in areas where glider and ultralight aircraft also fly because these aircraft are difficult to see. For the same reasons, ultralight pilots are also concerned with other ultralight and glider aircraft. Such diversions though are not always possible given the airspace design and thus, pilots have no choice but to fly in these areas.

Pilots that depart from an aerodrome in the capital city of a country are particularly concerned with the airspace design. The reason is that most airspace in which GA pilots would wish to fly is controlled, e.g. up to 6000 ft. In the airspace over the capital cities, pilots might be discouraged to plan a route through controlled airspace or their chances of receiving a clearance might be low in case they wished to fly through such airspace. In either situation, pilots select a route near this controlled airspace so that their actual flight route is as close as possible to their initial straight route that is supposed to be the quickest. Most pilots make this decision pre-flight; others make it in-flight once their request to enter controlled airspace is rejected. Furthermore, such a route might be selected to further ensure a safe flight, e.g. a greater distance from traffic in uncontrolled airspace and greater gliding distance. The key findings are summarised in Table 4.

Table 4. Key findings

Subject	Key findings
Purpose of flight planning	– Determine whether the flight will be conducted and select departure time – Pilots of fixed-wing motorised aircraft select the flight route and correct the aircraft's headings – Pilots of gliders decide the overall direction of the flight
Time that pre-flight planning starts	– Five planning strategies – Most pilots start the evening before departure – Pilots of fixed-wing motorised aircraft may start closer to the take-off if they are familiar with the area and use emerging planning and navigation technologies
Features that pilots consider in the pre-flight planning	Weather: – Open-cockpit ultralight aircraft are sensitive to rain – Minor though essential deviation of the flight route in ordinary weather conditions Airspace design: – Controlled airspace over the capital cities – Gliders are often ill-equipped to fly in controlled airspace – Pilots avoid congested areas whenever possible – Plan a route very close to controlled airspace

6 Conclusions

This paper, by means of semi-structured interviews conducted in Finland, Norway and UK and a carefully designed sample of recreational GA pilots based on six criteria, e.g. geographical location of the flying base, sheds lights on the planning undertaken by recreational GA pilots. The paper highlights the profound differences between the most common aircraft designs. The purpose of pre-flight planning differs between motorised and motorised aircraft. The use of the emerging planning technologies has encouraged pilots, particularly those who fly a fixed-wing motorised aircraft, to start their planning closer to take-off. Furthermore, pilots consider both the weather and airspace design of ordinary flights whilst the latter has not be considered much in the past.

Amongst the findings is that pilots who are familiar with the area they fly or use emerging planning technologies may plan the flight closer to departure rather than the day before. With such technologies, planning can be completed faster, e.g. in a few minutes, than using paper charts. Furthermore, airspace design, along with weather, is considered in the planning by all pilots for various reasons. For instance, ultralight pilots fly around controlled airspace due to their ill-fitted aircraft to fly in controlled airspace.

Such findings imply that the diversity of GA must be considered in any safety and operational analysis. An immediate application of these results is their incorporation into the safety analysis of a frequent type of occurrence involving GA in European airspace, the unauthorised entry of GA aircraft into controlled airspace.

Acknowledgments. The authors would like to thank Lloyds Register Foundation who funded this research. The authors would also like to thank the participants of this study as well as Per Julius Helweg, Heli Koivu, Armin Zuger and Bettina Bajaj who organised the interviews in Finland, Norway and UK.

References

1. Civil Aviation Authority United Kingdom. Strategic Review of General Aviation in the United Kingdom, United Kingdom, July 2006
2. International Civil Aviation Organization. Tenth session of the sattistics division, Review of the classification and definitions used for civil aviation activities. International Civil Aviation Organization. Report No.: STA/10-WP/7 (2009)
3. Aircraft Register Statistics: UK registered aircraft by class and weight group, multi-year 1985-2016 (Internet). Civil Aviation Authority United Kingdom; (2016) https://www.caa.co.uk/Data-and-analysis/Aircraft-and-airworthiness/Datasets/Aircraft-register-statistics/
4. SkyDemon [Internet].; 2015, updated 11 Sept (2015) http://www.skydemon.aero/
5. Wiggins, M.: Perceived GPS use and weather-related decision strategies among pilots. Int. J. Aviat. Psychol. **17**(3), 249 (2007)

6. Wiggins, M.W., Azar, D., Hawken, J., Loveday, T., Newman, D.: Cue-utilisation typologies and pilots' pre-flight and in-flight weather decision-making. Saf. Sci. **65**, 118–124 (2014)
7. Coyle, A., Lyons, E.: Analysing Qualitative Data in Psychology. SAGE Publications, Thousand Oaks (2007)
8. Elliott, R., Fischer, C.T., Rennie, D.L.: Evolving guidelines for publication of qualitative research studies in psychology and related fields. Br. J. Clin. Psychol. **38**(3), 21–29 (1999)

HMIM: A Method to Study the Information and Control Flow Exchange in the Flight Deck

Yiyuan Zheng[1(✉)], Yanyu Lu[2], Yuwen Jie[1], and Shan Fu[2]

[1] Shanghai Aircraft Airworthiness Certification Center of CAAC, Beijing, China
leodeisler@sjtu.edu.cn
[2] Shanghai Jiao Tong University, Minhang Qu, China

Abstract. According to cognitive process of operators, any undertaken task process could be decomposed into information flow and control flow. The amount of the flow reflects the difficulty of the task at any moment during the task process, and the change of the amount expresses the complexity of the task procedure. Human-machine interactive momentum (HMIM) was proposed in this paper to represent the exchange of information and control flow during the task process, and the related processing methods were also suggested to analyze the relationship between task and workload of individual operator. The manual approaching and landing process of a typical aircraft was used as a case study, and the results demonstrated a good correlation among HMIM, flight performance and workload variation. HMIM could be used as a human machine interaction analysis method.

Keywords: Information and control flow exchange · Human-machine interaction · HMIM

1 Introduction

Timeline analysis (TLA) is one of the effective methods to evaluate human-machine interaction [1], as it breaks down the ratio of time required, that is, execution time, to time available throughout a mission scenario [2]. To implement the TLA method, the missions and tasks should be decomposed into the detailed activities firstly. Mission is defined as in terms of a higher lever goal normally. Taking a normal flight mission from Shanghai Pu Dong Airport to Peking International Airport as an example, the mission is arriving at Peking safely. During the flight, it includes several tasks such as taxing, taking off, climbing, cruising, approaching and landing. TLA method has already been used in aviation and aerospace fields widely. Dixon and Wickens [3] implemented it in multiple UAV Flight Control, Muraoka [4] established a crew behavioral model through TLA, and Schlierf and Stambolian [5] applied it to improve the processing flow of the Orion spacecraft.

However, as definition, timeline analysis only could provide the information of time pressure, some other aspects could also influence the operator performance directly in the human-machine interaction process, for instance, the complexity characteristic in the process [6, 7]. The complexity of human-machine interaction contains two aspects. The first is the information and control interaction that provides the

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_6

resources exchange through flow transmission between the operator and the systems. Many studies have been concerned with this interaction. Arya and Glover studied the interaction between decision and control problems and the value of information [8]. Ryan summarized that the control feedback undermined intrinsic motivation relative to informational feedback [9]. Secondly, the probability of the changeable arrangement of the operating procedures also determines the complexity, and influences the operator's performance consequently. For instance, the Standard Operating Procedures (SOPs) provide typically a list of action items or criteria arranged in a systematic manner, allowing the operator to record the presence/absence of the individual items listed to ensure that all are considered or complete, in order to decrease human error based on the procedure level in aviation [10]. Nevertheless, although pilots should follow the rigorous procedures, the operations might be disturbed or even overlapped sometimes in real flight. These unexpected disturbances or overlaps are often the reasons of human errors because they lead to the resource channel blockage [11].

In order to study the process of information and control flow exchange in flight deck, in this paper, we proposed a novel concept which was named as Human Machine Interactive Momentum (HMIM) and stated its associated techniques. Furthermore, a case study of a manual landing was implemented to prove its validity and usability with performance and workload. In this study, we only considered the visual information from the displays preliminarily, and we used glide trajectories deviation and NASA-TLX [12] to verify HMIM.

2 Definitions

2.1 Definition of HMIM

Inspired by the concept of Momentum (P) in classical mechanics, which is the product of the mass and velocity of an object [13], we defined a novel value HMIM, which is equal to the multiplication of information quantity and control inputs, to describe the process of information and control flow exchange of the system. Information quantity and control inputs could be regarded as two characteristics of a closed control loop system, where the information is the inherent feature, and the control is the external characteristic. Meanwhile, either of them could affect the result of HMIM.

The computational formula of HMIM is defined as:

$$HMIM = I \cdot C \tag{1}$$

where I is the information quantity that represents the information acquisition and process, and C is the control inputs, digitizing as control value.

Information acquisition and process is based on the specific information quantities (I_q), and control value are determined by control sequences (C_s). Therefore, the small amount of HMIM is calculated as:

$$\Delta HMIM = I_q \cdot C_s. \tag{2}$$

The total HMIM is the sum of $\Delta HMIM$:

$$HMIM = I_{q1} \cdot C_{s1} + \ldots + I_{qn} \cdot C_{sn} \tag{3}$$

2.2 Determination of Information Quantity

In this study, I in Eq. (1) was deduced from following formula based on information theory [14].

Let X be a discrete random variable space with probability mass function $p(x)$, where $x \in X$. I(x) is the information quantity, defined as:

$$I(x) = \log\left(\frac{1}{p(x)}\right) = -\log(p(x)) \tag{4}$$

where $p(x) = \Pr(X = x)$ is the probability that message x is chosen from all possible choices in the space X. The logarithm is base 2, and the measure of information is expressed in units of bits.

With different display areas, the information quantity that transfers to the operator is various. In flight deck, the prime flight display presents a variety of vital information directly related to the aircraft configurations including airspeed, attitude and etc. Moreover, diverse visual forms have been utilized to present these parameters. For instance, airspeed and altitude displays use bar graph with numeric display to show actual and intended value, while attitude indicator is depicted as square graph. Therefore, when pilot gazes at different display areas, the information quantity he/she acquires and processes is different. The quantity of specific display is determined by the changing probability of the next state. Considering the display of airspeed, for example, the change possibility equals to 1/3, since the next state is either same as the current one or at a positive or negative value. Therefore, for one fixation on airspeed within the minimum fixation period, the information quantity is $\log_2 3$ bits.

2.3 Determination of Control Value

We defined two kinds of operations in the flight deck, continuous and discontinuous ones. For instance, the controls on steering wheel and throttles belonged to the first kind, and the operations on overhead panel, flaps and landing gear were the second.

Considering continuous operations, the control value equaled to the absolute normalized value of difference between the two adjacent time points. Furthermore, for discontinuous operations, the control value of the pilots' right hands were defined as '1'. Because these operations required the specific behaviors of the pilots' hands, such as pressing the buttons. If the two hands of the operator were not on the same control device at the same time point, the control value referred to the sum of the two hands operations.

3 Case Study

As mentioned above, as long as an operation occurred based on information acquisition and process, HMIM could characterize the flow exchange. We implemented a landing phase as a case study, and evaluated the validity and usability by analyzing HMIM results and comparing them with performance and workload.

3.1 Apparatus

The experiment was carried out in a Boeing 777-200ER flight simulator, which was build up based on flight gear. The simulator was established by three workstations (Intel(R) Core: trademark: i5-2310 Central Processing Unit @2.90 GHz, 4 GB DDR3 RAM, HP, China) and six 22-in liquid crystal display screens (LE2201W, HP, China). The outside view was displayed on a semi-spherical screen by three projectors, which presented a real immersed experience of the flight to the subjects.

Besides the flight simulator, an eye tracker (Smart Eye, Smart Eye AB, Sweden) was used to determine the fixation areas and corresponding durations of the subjects during the experiment. Otherwise, all the experiment procedures were recorded by two cameras which were fixed at the rear of flight simulator's ceiling.

3.2 Setting

Areas of Interest. We divided the displays interfaces into 7 areas of interest (AOI) based on the common information acquisition areas for the left seat in the fight deck. The prime flight display (PFD) was divided into 3 areas: the Airspeed part (No.1 area); the Attitude & Heading part (No.2 area) and the Altitude & Vertical Speed (VS) part (No.3 area). Navigation display that was under map mode was No.4 area. No.5 area was Engine-indicating and Crew-Alerting System (EICAS) with flaps and landing gear indicators. Out of the window (OTW) was No.6 area, where the pilot could capture the runway or any identifications in airport, like Visual Approach Slope Indicator (VASI). No. 7 area was Overhead Panel. The specific separation result was shown as Fig. 1.

Information Flow Quantization. The next step was to determine the information quantity acquired and processed by each fixation at different AOI. The information quantity was calculated though Eq. (4) based on following assumptions:

- Pilots acquire and process information through eye fixation on displays.
- The minimum fixation period for pilot to acquire and process information from displays was 200 ms [15].
- The information quantity of each fixation was based on the change possibility and the duration of the fixation. For example, if one fixation was on the airspeed area, then the information quantity in this fixation was $\log_2 3 * \Delta t$ bits, where Δt was the fixation duration. Table 1 was given out the information quantity of the above 7 areas in the minimum fixation period (200 ms):

Fig. 1. AOI of the left seat in the flight deck

Table 1. Information quantity of different areas in the minimum fixation period

Areas	Airspeed	Attitude and heading	Altitude and VS	Navigation display	EICAS	OTW	Overhead Panel
Information quantity	$\log_2$ 3bits	$\log_2$ 9bits	$\log_2$ 9bits	$\log_2$ 6bits	$\log_2$ 12bits	$\log_2$ 9bits	$\log_2$ 3bits

Control Flow Quantization. During the landing phase, the responsibility of pilot flying (PF) is to descent the aircraft toward the runway at a constant glide slope (normally the slope is −1/29). Therefore, the most used control devices are joystick and throttles. Besides, flaps and landing gear operations are also important. The control value was calculated in the landing phase based on the following assumptions:

The whole landing phase was accomplished by manual flight.

- When the subject operated the steering wheel by his/her two hand, the control value equaled to the absolute normalized value of the difference of the steering wheel comparing with the last time points.
- If the subject adjusted the throttle, flaps, landing gear or overhand panel by his/her right hand, then the control value should be calculated referring to the sum of the left and right hands operations.

3.3 Participants

Twelve college master students (all male) with normal vision from School of Aeronautics and Astronautics, Shanghai Jiao Tong University were participated in the experiment (average age = 22.68, SD = 2.3). All of them were familiar with aircraft design, and six of them had the simulator flight experience (mean flight hours = 126, SD = 76). All of them signed the subject consent form before the experiment which was approved by the Institutional Review Board of Shanghai Jiao Tong University.

3.4 Procedure

The experiment of the landing began at ten nautical miles to the airport KSFO 28R (San Francisco), and the height of aircraft was 2300 feet. The subjects were required to perform a final manual landing at the left seat. In this phase, the subjects should maintain about a 3 to 5 degree nose-up and a constant glide slope (−1/29) to descent, and need to hear the instructions from one experimenter, who sat at the right sear, to settle the flaps at the specific airspeeds and extract the landing gear, the airspeeds and corresponding flap configurations were shown in Table 2. After extracting the landing gear, the subjects were required to align the aircraft to the runway and land successfully. According to the above flight operations, we divided the experiment into two flight segment, flight before extracting landing gear and flight after extracting landing gear. After the experiment, each subject was required to two forms of NASA-TLX to evaluate subjective workload of the two flight segments.

Table 2. The relationship between flap configurations and airspeed

Flap configurations	Flap 5	Flap 15	Flap 25
Airspeed	220 knots	210 knots	190 knots

Every subject took part in flight trainings three times a week in one month before the experiment, and each training was about half an hour. They were required to achieve the qualified level that their successful landing rate was over 90% at the last training session. These trainings were necessary because they make the subjects be familiar with the operations and the whole experimental process. Ten of the twelve subjects had completed the training prior to the experiment. The other two subjects who failed in the training process, were not participated in the experiment.

4 Results

Ten subjects were all landed on the runway successfully. To express the validity and applicability of HMIM, results would consist of three aspects. At first, three typical flight HMIM sequences were selected to express the characteristics. Then the corresponding vertical deviations of the glide trajectories would be presented to indicate the performance of the three flights. At last, the experimental HMIM results were compared with NASA-TLX results.

4.1 Characteristics of HMIM Sequences

In Fig. 2, the three typical experimental HMIM results were shown, and the time interval in axis X was 200 ms.

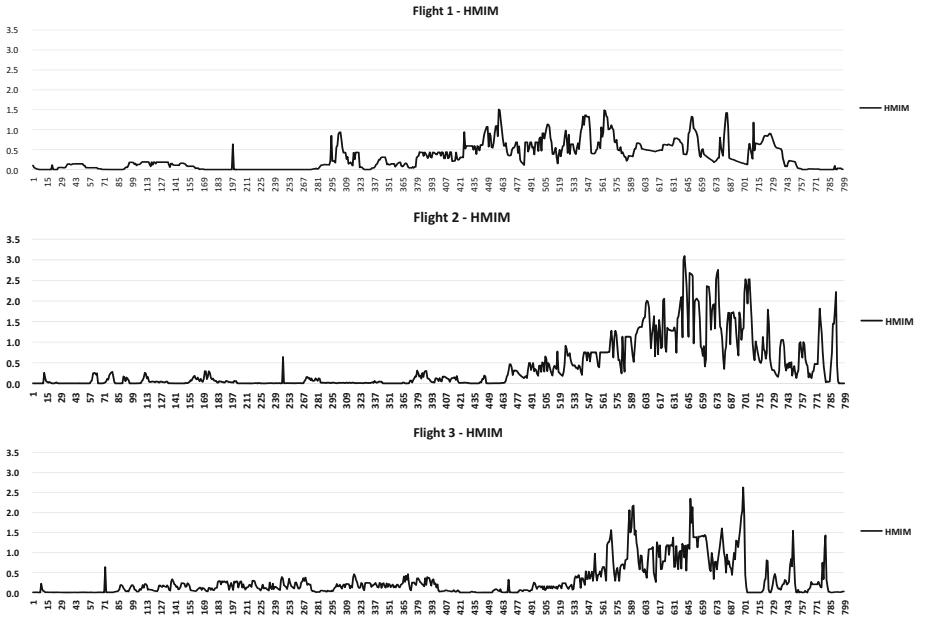


Fig. 2. Three experiments HMIM results sequences. The axis x was the number of time interval (200 ms per 1 interval), and axis y was the corresponding value of HMIM.

The characteristics of HMIM sequences were as follows:

- The first segment of sequences contained some small fluctuations and protuberances that reflected minor adjustments of aircraft and flaps and gear operations.
- The relatively large value of experimental HMIM began at almost second segment of the flight phase that from time interval of 421, 462 and 533 respectively.
- The second and third HMIM sequences were more similar than with the first one.

4.2 The Corresponding Vertical Deviations of the Glide Trajectories

The actual flight glide paths and the expected glide paths were shown as Fig. 3 corresponding to the above three different HMIM results sequences. The solid lines were the actual paths, and the dotted lines were the expected paths, which the glide slope were all −1/29 [16].

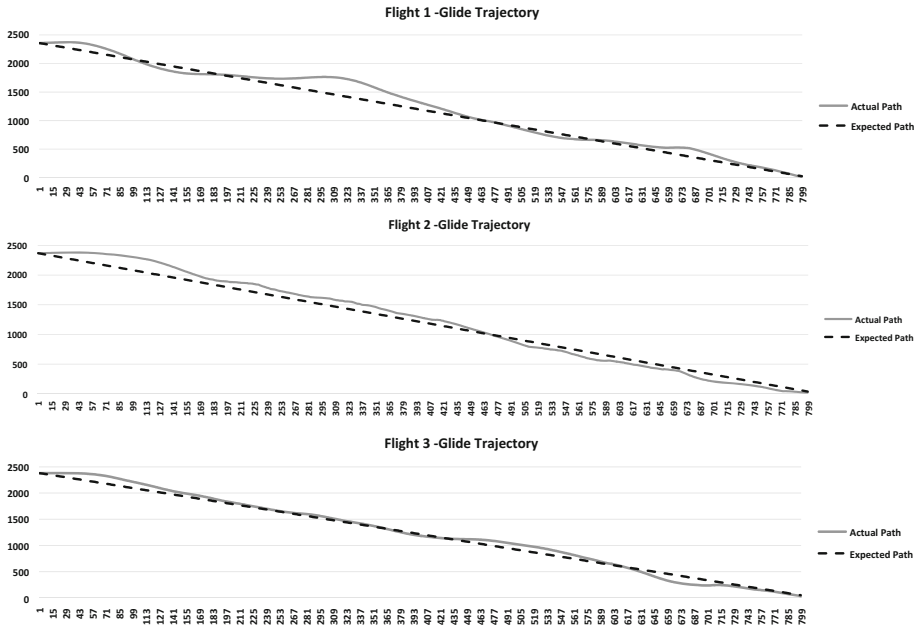


Fig. 3. Actual flight paths with expected flight paths, where solid lines were actual flight paths and dotted lines were expected descending paths that the slops were −1/29. The axis x was the number of time interval (200 ms per 1 interval), and axis y was the corresponding height in feet

In Fig. 3, the worst flight path was the flight 1 according to mean deviation and standard deviation shown in Table 3, and the best one was the flight 3. Table 3 presented the respective mean deviations and the standard deviations of these three flights.

Table 3. Average deviation and the standard deviation of the three flights

	Mean deviation (feet)	Standard deviation
Flight 1	202.88	134.82
Flight 2	107.46	83.96
Flight 3	57.68	43.92

4.3 HMIM with NAS-TLX

In order to compare with the scores of NASA-TLX, we also divided the HMIM sequences into two segments at the flight point of extracting the landing gear, and the mean value the two parts of the HMIM were calculated. The maximum, minimum, mean value and SD of two segments of HMIM and the corresponding values of NASA-TLX of the ten subjects were shown in Table 4

The NASA-TLX scores indicated that the two segment of landing phase had different complexity that result in that the workload difference was significant ($F(1, 18) = 11.520$, $p = 0.03$). Meanwhile, the difference of HIMI was also significant

Table 4. The subjects' results of two segments of NASA-TLX scores and HMIM

	First segment of HMIM	First segment of NASA-TLX	Second segment of HMIM	Second segment of NASA-TLX
Max.	0.645	63	1.481	69
Min.	0.088	20	0.732	28
Mean	0.25	34	1.04	53
SD	0.17	13	0.28	12

($F(1, 18) = 57.751$, $p < 0.001$). Furthermore, HMIM and NASA-TLX was not related in the first segment of flight ($R = 0.118$, $p = 0.745$), the in the second segment, HMIM was relevant to NASA-TLX ($R = 0.658$, $p = 0.039$). Combining the results of two segments together, there was a correlation between HMIM and NASA-TLX ($R = 0.668$, $p = 0.001$).

5 Discussion

Although the level of automation in the fight deck is pretty high, the aircraft still need to be operated and controlled by pilots, especially during the flight phases which required significant amount of pilot's physical and mental workload like takeoff and landing [17]. According to accident investigators, aviation accidents are usually the result of a chain of unexpected events culminating with the unsafe controls of operators [18]. In order to represent the human machine interaction and the information and control flow exchange process during the manual flight, we proposed an indicator HMIM, which combined pilot's information acquisition, process and control input together in this study. Furthermore, we described the characteristics of HMIM, and used glide trajectories deviation and NASA-TLX to evaluate it.

The results in Fig. 3 expressed that to the different operators, the information and control flow exchange processes were similar in the specific task to achieve a better performance. Taking the landing phase as an example, the operators only needed to focus on prime flight display (PFD) and made a few of adjustments to maintain descending in the first segment. Meanwhile, they were required to extend the flaps at the specific airspeeds. Therefore, the first segment of HMIM sequences were relatively low with some small fluctuations and protuberances. However, after extracting the landing gear and approaching to the destination, the operators needed more efforts to control the aircraft. Moreover, when interacting the glide path, they should distract some of their attentions to the outside window, align the aircraft with the runway and finish landing. Thus, the second segment of HMIM sequences were relatively large with more frequent information and control flow exchange processes.

Considering the performances of glide paths in Fig. 3, flight 2 and flight 3 were better than flight 1. Therefore, the patterns like the second and the third HMIM sequences in Fig. 3 were more likely to give rise to a better performance. This could be used as a standard for training evaluation, and this kind of HMIM pattern should be the training target for the novice. Furthermore, the inappropriate HMIM results could support post hoc fault diagnosis combining flight performances. For instance, the

deviations of glide path were quite large from time interval 285 to 323 in flight 1, and the corresponding HMIM results were also abnormal. Looking back to the experiment record, we discovered that the subject did not pay attention to PFD but OTW to maintain descending, and the manipulations were quite large consequently.

The overall correlation between HMIM and NASA-TLX showed that HMIM could be used as an indicator of the workload of the task seemingly. This was reasonable, for the information acquisition and process, and the control devices operations were the most crucial aspects of the workload. Considering the six aspects of NASA-TLX [19], the process of human machine interaction was directly related to physical workload, mental workload, and effort. Otherwise, time pressure, performance and frustration also had influences on information and control flow exchange processes. However, some other factors also influences human machine interaction, for instance, trust and situation awareness [20, 21]. Meanwhile, since NASA-TLX is a subjective measurement, the operator is sometimes an unreliable measurement instrument [22]. Therefore, there was no correlation between HMIM and NASA-TLX in the first segment. Nevertheless, the differences between two segments were all significant, both in HMIM and NASA-TLX. This might because of the different complexities of the two segments. The first segment only required the small adjustments and some memory events of the subjects, and the second segment needed immediate response to the voice instruction and more precise operations. As the human machine interaction was strongly related to systems complexity [23], HMIM could also represent the task complexity.

Tradition timeline analysis computers the ratio of time required to time available for each section of the task [24]. Comparing with the traditional method, HMIM has two improvements. Firstly, HMIM contains the complexities of different display areas presented and different operating characteristics that based on timeline, because it considers both the information quantity and the control input. Therefore, HMIM could represent the complexity of each time point in the task. Secondly, although SOPs should be followed during the flight, the changeable arrangement of operations are inevitable, especially when two tasks are overlapped. Timeline analysis is more applicable to linear task, and HMIM could express every operating consequence whatever the task is.

However, some limitations still should be resolved in further study. As HMIM quantitatively expresses the flow exchange, the threshold should be determined. Excessive and inappropriate HMIM might bring about human error. Therefore, human-machine function allocation analysis might be based on HMIM results. Effective and principled task-partitioning should represent a central strategy for the evolution of complex human–machine systems [25]. Moreover, the recent study assesses the information obtained by display, but the effect of inattentional blindness was not considered [26], and we used the fixation results from eye tracker directly. Otherwise, pilots also get information though communicating with Air Traffic Control in real flight. The further work would concentrate on these information factors. In addition, the current study case was selected as the landing phase, the further validation would extend to the some other flight scenarios that consist of the emergencies and system malfunctions.

References

1. Boy, G.A.: The Handbook of Human-machine Interaction: A Human-centered Design Approach. Ashgate Publishing, Ltd, Farnham (2012)
2. Miller, K.M., Timeline Analysis Program (TLA-1), final report, N.A.S. Administration, Editor. 1976, Langley Research Center
3. Dixon, S.R., Wickens, C.D., Chang. D.:Comparing quantitative model predictions to experimental data in multiple-UAV flight control. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications (2003)
4. Muraoka, K.: Crew Behavioral Model for Timeline Analysis. In: 10th CEAS Conference Free Flight. (1997)
5. Schlierf, R., Stambolian, D.B.: Human factors operability timeline analysis to improve the processing flow of the orion spacecraft. In: Aerospace Conference. 2011 IEEE.
6. Tan, H.-T., Ng, T.B.-P., Mak, B.W.-Y.: The effects of task complexity on auditors' performance: the impact of accountability and knowledge. Auditing: A Journal of Practice & Theory **21**(2), 81–95 (2002)
7. Campbell, D.J.: Task complexity: a review and analysis. Acad. Manag. Rev. **13**(1), 40–52 (1988)
8. Arya, A., Glover, J.C., Sivaramakrishnan, K.: The interaction between decision and control problems and the value of information. Account. Rev. **72**(4), 561–574 (1997)
9. Ryan, R.M.: Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. J. Pers. Soc. Psychol. **43**(3), 450 (1982)
10. Hales, B.M., Pronovost, P.J.: The checklist-tool for error management and performance improvement. J. Crit. Care **21**(3), 231–235 (2006)
11. Wickens, C.D.: Multiple resources and mental workload. Human Factors J Human Factors Ergonom. Soc. **50**(3), 449–455 (2008)
12. Roscoe, A.H., Ellis, G.A.: A subjective rating scale for assessing pilot workload in flight: a decade of practical use. DTIC Document (1990)
13. Feynman, R.P., Leighton, R.B., Sands, M.L.: The Feynman Lectures on Physics: Mainly Mechanics, Radiation, and Heat, vol. 1. Basic Books, New york (2008)
14. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (1991)
15. Liu, Z., Yuan, X.: Analysis on eye movement indices based on simulated flight task. China Safety Sci. J. **16**(2), 47–51 (2006)
16. Manual, F.O.: Flight Operations Manual. Emirates revision, 18 (2003)
17. Wilson, G.F.: An analysis of mental workload in pilots during flight using multiple psychophysiological measures. Int. J. Aviat. Psychol. **12**(1), 3–18 (2002)
18. Shappell, S., et al.: Human error and commercial aviation accidents: an analysis using the human factors analysis and classification system. Human Factors **49**(2), 227–242 (2007)
19. O'Donnell, R., Eggemeier, F.T.: Workload assessment methodology. Meas. Tech. **42**, 5 (1986)
20. Lee, J., Moray, N.: Trust, control strategies and allocation of function in human-machine systems. Ergonomics **35**(10), 1243–1270 (1992)
21. Hoc, J.-M.: From human-machine interaction to human-machine cooperation. Ergonomics **43**(7), 833–843 (2000)
22. Gopher, D., Donchin, E.: Workload: an examination of the concept. (1986)
23. Wilson, J.R.: Fundamentals of systems ergonomics/human factors. Appl. Ergon. **45**(1), 5–13 (2014)

24. Corwin, W.H., et al.: Assessment of Crew Workload Measurement Methods, Techniques and Procedures. Process, Methods and Results, Vol. 1, DTIC Document (1989)
25. Hancock, P.: Task partitioning effects in semi-automated human-machine system performance. Ergonomics **56**(9), 1387–1399 (2013)
26. Herslund, M.-B., Jørgensen, N.O.: Looked-but-failed-to-see-errors in traffic. Accid. Anal. Prev. **35**(6), 885–891 (2003)

Analyzing Positive and Negative Effects of Salience in Air Traffic Control Tasks

Haruka Yoshida[1(✉)], Hisae Aoyama[2], Satoru Inoue[2], Taro Kanno[1], and Kazuo Furuta[1]

[1] The University of Tokyo, 7-3-1, Hongo, Bunkyo-Ku, Tokyo, Japan
h-yoshida@cse.t.u-tokyo.ac.jp, kanno@sys.t.u-tokyo.ac.jp, furuta@rerc.t.u-tokyo.ac.jp

[2] Electronic Navigation Research Institute, 7-42-23, Jindaiji Higashimachi, Chofu, Tokyo, Japan
{aoyama,s-inoue}@mpat.go.jp

Abstract. We focus on visual attention as a high-impact perception in the accurate and efficient ATC tasks, which demand high cognitive workloads. Our goal is to develop a screen design policy considering salience of information. Salience is one of the user interface elements causing users' visual attention. We carried out an experiment to clarify relationships between the degree of salience gaps and the performance of ATC tasks and we will develop a screen design policy based on the relationships. As a result, the lager the salience gaps between important and unimportant airplanes were, the shorter the reaction time for replying situation awareness queries became. The larger the salience gaps among airplanes were, the better instruction timing was obtained. On the other hand, on the screen design with a large salience gap, the performance of sub task for the airplanes displayed with a low salience delayed than that without salience gaps. From these results, we conclude that our hypotheses were supported.

Keywords: Salience · Air traffic control · Situation awareness · Screen design

1 Introduction

The role of air traffic control (ATC) is keeping and facilitating safe and smooth flight for airplanes. It recently becomes more important with the increasing demand of air traffic. In ATC tasks, the controllers accurately understand many kinds of dynamic information on the radar screen. They are always aware of and comprehend the situation of airplanes and information about the environment so that the airplanes can arrive at their destinations safely and efficiently. If there are risky airplanes for safe flight, the controllers give instructions such as change of the altitude or direction to the airplanes using communication channels in order to avoid the risk. ATC tasks are multiple and complex tasks that require understanding many kinds of dynamic information and executing different tasks in parallel. Moreover, they are crucial tasks, because a small mistake may lead to major accidents. Thus, ATC tasks are very difficult task of a high cognitive demand [7]. They hereafter become more complicated and require more workloads for the controllers with the increasing demand of air traffic.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_7

To these problems, we suppose that screen design of the ATC systems considering cognitive characteristics can reduce controllers' workloads and improve task performance. Hence, we focus on visual attention as a high-impact perception in accurate and efficient ATC tasks. In this study, we aim to develop a screen design policy considering salience of information. According to existing researches, salience is one of the user interface elements causing users' visual attention [2]. We made the following hypotheses about salience effects on ATC tasks.

- Hypothesis 1: Controlled salience is helpful for situation awareness (SA) of air traffic.
- Hypothesis 2: Controlled salience is helpful to judge the target and timing of instruction.
- Hypothesis 3: Controlled salience according to importance of information in terms of main tasks gives negative effects on sub task performance.

This paper describes an experiment for verifying these hypotheses and the results of them. Especially, we study the effects for novice users in this experiment.

Previous researches by our group have developed a salience model that calculates salience value of objects from object color, size, and position on the screen. In this experiment, we use the model in order to control salience of information displayed according to the importance of instruction tasks (main tasks). We study positive and negative effects of salience on the ATC tasks in order to develop the screen design policy considering the salience effects.

There are many existing researches about salience [2, 6]. This study provides concrete examples of applying the salience model to designing screens for difficult task of a high cognitive demand. By this study, we can show the salience is an applicable user interface element to screen design of interactive tasks. Additionally, this study attempts to apply the cognitive psychology to human computer interaction design. It contributes to establishing a design methodology of user interface for improving task performance by applying achievements in the domain of cognitive psychology.

2 Method

We carried out an experiments with participants in order to verify the above three hypotheses. Experimental tasks were multiple tasks containing main tasks and sub tasks. The main tasks were giving instructions to airplanes and the sub tasks were hand in and hand off operation. The participants executed the experimental tasks on the ATC simulator screens [1].

2.1 Participants

Ten people from 20 to 30s of age participated in this experiment. All of them had little knowledge and experience about the ATC tasks. They usually use computers and used to operating computers with a mouse.

2.2 Experimental Condition

We prepared four patterns of screen design in terms of salience of displayed information and two kinds of task scenarios. We defined eight experimental conditions by combing the four screen patterns and two task scenarios. Each participant executed the eight experimental conditions in different orders.

Screen Design. We prepared four patterns of screen design for the experimental simulator in terms of salience of airplane displays. The four patterns include (a) no, (b) small, (c) middle, and (d) large gap of salience between important airplanes and the others. We controlled salience of each airplane display using the color salience model that we developed in our previous work [3–5]. The model quantifies total color salience of an object by a linier sum of the absolute color salience (feature salience) and the relative color salience against the surrounding colors (heterogeneous salience). We defined the feature salience of an object as a linier combination of saturation, hue and gap of lightness between the object and background color. On the other hand, we also defined the heterogeneous salience as a linier of saturation gaps, hue gaps, and lightness gaps between the screen average and the object color.

Table 1 shows the screen design conditions for color and salience values of airplane displays. Our model above mentioned calculates the salience values.

Table 1. The conditions for colors and salience in each screen design

Expression of airplane information	Screen design			
	(a) No gap	(b) Small gap	(c) Middle gap	(d) Large gap
(1) Airplanes to be instructed [most important]				
Color [RGB]	111,224,210	162,162,234	131,249,131	247,80,247
Salience	2.35	2.20	3.10	4.31
(2) Airplanes to be controlled [important0]				
Color [RGB]	111,224,210	179,215,179	179,215,179	131,249,131
Salience	2.35	2.00	1.51	3.05
(3) The other airplanes [unimportant]				
Color [RGB]	111,224,210	82,170,170	88,125,88	88,125,88
Salience	2.35	1.13	0.15	0.10
Salience gap between (1) and (3)	**0.00**	**1.07**	**2.95**	**4.21**
Salience gap between (2) and (3)	**0.00**	**0.87**	**1.36**	**2.95**

In this experiment, we classified airplanes displayed on the simulator into three categories, (1) airplanes to be instructed, (2) airplanes to be controlled, and (3) the other airplanes, considering the ATC instruction tasks. The order of importance for airplanes was the order of (1), (2) and (3). We controlled salience according to the order of importance in the screen design (b), (c) and (d). The airplanes (1), airplanes to be instructed, are the airplanes to be instructed immediately in the defined ATC plans. The colors of their information displayed are changed from the colors of Design (2) to Design (1) at 10 s before the moment of instruction according to the ATC plans, then, the colors are set back to Design (2) after the instruction has been issued by participants. The airplanes

to be controlled are those that cannot keep separation without any air traffic control. Namely, the airplanes to be instructed are a part of the airplanes to be controlled.

Figure 1 shows an example of the screen design (a)–(d) at the timing that the TEST 978 airplane has to be instructed. The airplane TEST 978 corresponds to the (1) airplanes to be instructed, the airplane TEST 451 and TEST 772 correspond to the (2) airplanes to be instructed, the TEST 217 and TEST 400 correspond to (3) the other airplanes.

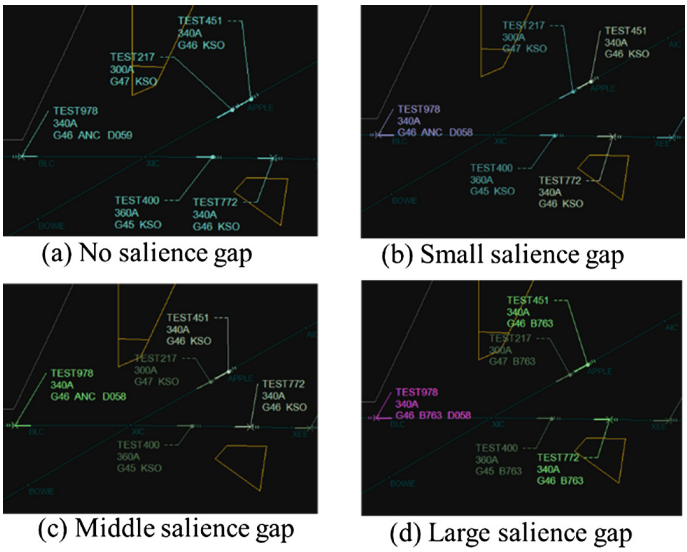


Fig. 1. A part of screen design example (a)–(d) displaying information of airplanes (1)–(3)

Scenario. We prepared two fictional scenarios simply imitating situations of air traffic and air traffic control plans (ATC plans) for each scenario. In both scenarios, we supposed that all airplanes were cruising at their own altitude and that the ATC plans included only instructions on heading for simplicity. There were about ten airplanes displayed on the screens, and three or four of them were the airplanes to be controlled that cannot keep separations without any ATC instructions. The ATC plans of this experiment provided that the controllers give from six to ten heading instructions to a part of the airplanes to be controlled.

In addition, there were airplanes going through the boundaries of sectors during execution of heading instructions. The participants have to execute hand in and hand off operations to the airplanes. Here, a hand in operation is an operation for receiving a new airplane from a neighboring sector to one's own sector. A hand off operation is an operation for handing an airplane from one's own sector to a neighboring sector. One of the authors who is a retired air traffic controller and has enough experience of educating novice controllers made these scenarios and ATC plans.

2.3 Procedure

The following list of steps shows the task flow of participants. In this experiment, we defined giving heading instructions as the main tasks, and hand in and hand off operations as the sub tasks. Duration of experimental session was about 10 min.

- Step 1: Explanation of a task scenario and the corresponding ATC plan by the experimenter
- Step 2: Starting the ATC simulator and giving heading instruction to target airplanes according to the ATC plan
- Step 3: Hand in and hand off operations at an appropriate timing in parallel
- Step 4: Answering oral questions about situation awareness (SA) [7] asked by the experimenter
- Step 5: Finishing the all planed ATC instructions

As for an experimental environment, participants executed all tasks on 22 inched displays whose resolutions were 1920 * 1080.

2.4 Measurement

In order to verify the hypotheses 1, 2, and 3, we made the following measurements.

Level of Situation Awareness (SA). To verify the hypothesis 1, we measured the level of SA using Online SA Probe method [8]. In this experiment, the experimenter asked the participants some oral questions about air traffic situations. We obtained the accuracy rate and the required time for answering (reaction time), and measured participants' SA level using the obtained data.

Instruction Timing. To verify the hypothesis 2, we obtained data of participants' instruction timing and measured gaps between participants' and the planed instruction timing. The smaller gap means the better timing of instructions.

Performance of Sub Tasks. To verify the hypothesis 3, we obtained time data of participants' hand in and hand off operations. Then, we measured the time gaps between airplanes' crossing the sector boundaries and participants' hand in and hand off operations. The time of boundary crossing is the time limit of the hand in or hand off operation. A smaller gap means a delay of the hand in or hand off operation.

3 Experimental Result

3.1 Level of Situation Awareness (SA)

Figure 2 shows the reaction time for oral questions about SA during task execution using each screen design. The graph A shows the reaction time for questions about SA of level 1 and level 2 respectively on the airplanes to be controlled. The graph B shows the reaction time for questions about SA of level 1 and 2 respectively on the other airplanes.

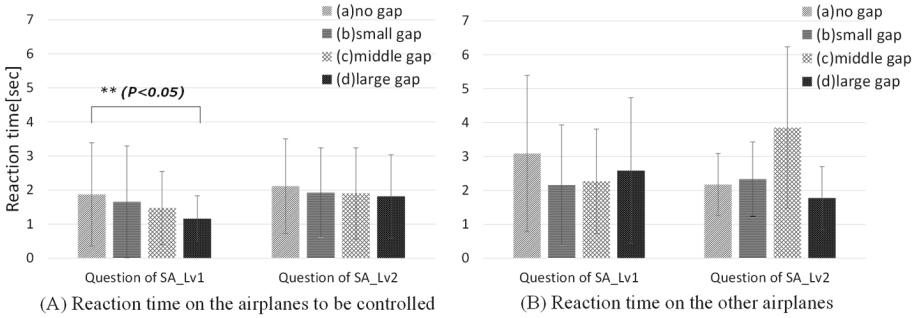


Fig. 2. Reaction time for oral questions about Situation Awareness (SA) – level 1 and 2 of important and unimportant airplanes

According to the graphs, we obtained the result that the reaction time for queries about SA-level 1 on the airplanes to be controlled is significantly shorter for the screen of a large salience gap. We observed this tendency only in queries about the airplane to be controlled. On the other hand, in queries about the other airplanes, there was no significant difference in the reaction time among the screen designs. Moreover, the lager the salience gap between the airplanes to be controlled and the others are, the shorter the reaction time about SA queries on the airplanes to be controlled becomes. From these results, we found that a large salience gap of information display could improve novices' SA.

3.2 Instruction Timing

Figure 3 shows the gap of instruction timing between the ATC plans and participants' executions for each screen design.

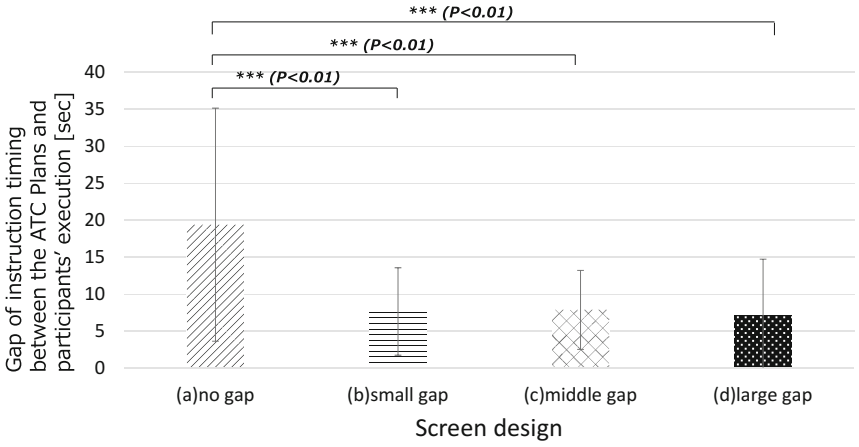


Fig. 3. The gap of instruction timing between the ATC plans and participants' executions

According to the graph, the gap of instruction timing between ATC plans and participants' executions are significantly smaller for the screen with small, middle and large gap of salience between important and unimportant airplanes. In addition, there was no significant difference of instruction timing among the screen design with a certain gap of salience, namely Design (b), (c) and (d). From the result, novices can execute instructions more timely if there is a salience gap among airplanes displayed.

3.3 Performance of Sub Tasks

Figure 4 shows the time from the moment of executing hand in and hand off operations to that of boundary crossing by the airplanes.

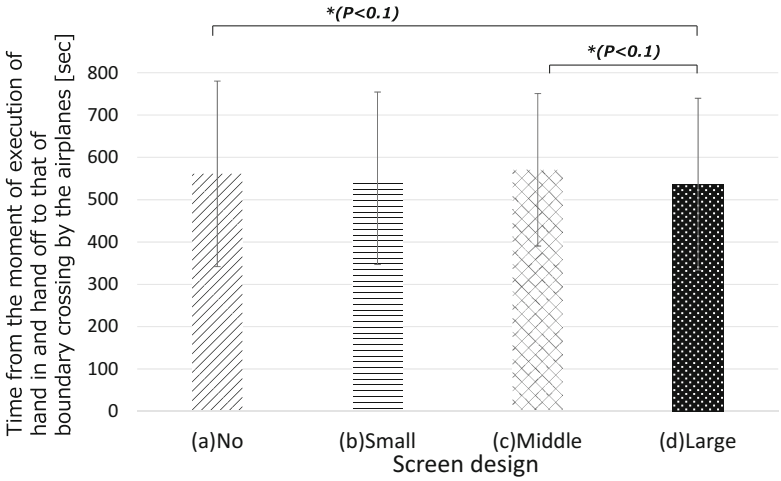


Fig. 4. Time from the moment of executing hand in/hand off to that of boundary crossing by the airplanes

In this measurement, a small value represents a delay of hand in or hand off operations, namely a low performance of sub tasks. According to the graph, on the screen design with a large salience gap, hand in and hand off operations to airplanes with a low salience gap tended to delay than that with a small or middle salience gap. From this result, we found that the performance of sub tasks tends to degrade on the screen design with a large salience gap.

4 Discussion

4.1 Verification of Hypothesis

From Fig. 2, we obtained the result for novices that the reaction time about SA queries became shorter on the screen with a larger salience gap between important and unimportant airplanes. Therefore, the hypothesis 1 was supported if a shorter reaction time

of SA queries corresponds to a better SA. Next, from Fig. 3, we obtained the result that novices can give instructions timelier if the salience of displayed information is controlled according to the importance of information and the salience gap is large. Therefore, the hypothesis 2 was supported. Finally, from Fig. 4, we obtained the result that the performance of sub tasks with a low salience gets worse if the salience is controlled according to the importance of information in terms of main tasks. Therefore, the hypothesis 3 was supported.

4.2 Effects of Salience for Novice Users on the ATC Tasks

In this experiment, all participants were novices for the ATC tasks. We verified the effects of salience on their SA, instruction tasks (main tasks), and hand in and hand off tasks (sub tasks). As for SA, controlled salience according to the importance of information can improve novices' SA. We supposed the reason that different salience of displayed information could help novices to recognize information according to the importance and to understand traffic situations. As for the instruction tasks in ATC, novices gave instructions more timely as the salience gap between the airplanes to be controlled and the others were large. We supposed the reason that novices allocated more attention to the airplanes to be instructed and could easily recognize instruction timing by a large salience gap. From these results, we found that controlled salience of displayed information gave positive effects on the ATC tasks for novices such that recognizing the importance of information has contributed to improving SA.

On the other hand, the performance of hand in and hand off operations, namely sub tasks tended to degrade if the salience of airplane displays is controlled according to the importance of airplanes in terms of the main tasks. We considered the reason that novices would allocate attention to the airplanes with a high salience and could not allocate attention to the airplanes with a low salience. From this result, salience controlled according to the importance in terms of main tasks would give negative effects to the performance of sub tasks.

5 Conclusion

Our goal is to construct a policy of screen design to support ATC tasks, which demands high cognitive workloads, by considering the salience of displayed information. We have developed a mathematical model to quantify the salience of display focusing on visual attention. In this paper, we carried out an experiment in order to verify the effects of salience gap in the displayed information on ATC tasks. The result shows that the relationships between the salience and ATC tasks as follows:

- The larger the difference of salience among displayed information becomes, the higher SA novices can obtain.
- The larger the difference of salience among displayed information becomes, the more timely novices can issue instructions to the air traffic.

- If the salience of displayed information is controlled according to the importance in main tasks, a large salience gap among the displayed information degrades the performance of sub tasks.

In the future, we will measure the impact of salience on tasks in terms of workloads by conducting a subjective assessment, and examine the effects for experts. From the experimental findings, we will develop screen design policies.

References

1. Karikawa, D., et al.: Study on performance visualization for educational purpose in air traffic control area. Trans. Hum. Interface Soc. **11**(3), 1–10 (2009). (In Japanese)
2. Itti, L., et al.: A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Res. **40**, 1489–1506 (2000)
3. Yoshida, H., et al.: Quantifying visual attention - effect of object size. In: Advances in Ergonomics Modeling, Usability & Special Populations. Advances in Intelligent Systems and Computing, vol. 486, pp. 173–187 (2016)
4. Yoshida, H., et al.: Quantification of visual attention (2) - modeling effect of size, layout, and integrated elements. In: Human Interface Symposium 2013 (2013). (In Japanese)
5. Yokomizo, A., et al.: Quantification of visual attention (1) - modeling effect of color. In: Human Interface Symposium 2013 (2013). (In Japanese)
6. Tanaka, S., et al.: An attractiveness evaluation model based on the physical features of image regions. IEICE A JPN edition, J83-A(5), 576–588 (2000). (In Japanese)
7. Endsley, M.R.: Situation awareness. In: Aviation Systems, Handbook of Aviation Human Factors, p. 257. Lawrence Erlbaum Associates, Mahwah (1999)
8. Endsley, M., Sollenberger, R., Nakata, A., Stein. E.S.: Situation awareness in air traffic control: enhanced displays for advanced operations (No. DOT/FAA/CT-TN00/01). Federal Aviation Administration Technical Center Atlantic City, NJ (2000)

Engine Failure Induced Task Load Transient for Simulation Based Certification Aiding for Aircraft

Jens Alfredson[1,2(✉)], Luís G. Trabasso[1,3], Niklas Blomstrand[1], Maria Eckerberg[1], Linda Klamer[1], Johanna Ledin[1], Jasmine Tarander[1], and Magnus Bang[1]

[1] Department of Computer and Information Science, IDA, Linköping University, Linköping, Sweden
{Jens.Alfredson, luis.gonzaga.trabasso, Niklas.Blomstrand, Maria.Eckerberg, Linda.Klamer, JohannaLedin, Jasmine.Tarander, Magnus.Bang}@liu.se
[2] Saab – Aeronautics, Bröderna Ugglas Gata, Linköping, Sweden
Jens.Alfredson@Saabgroup.com
[3] Department of Mechanical Engineering, DCTA, ITA, São José Dos Campos, Brazil
gonzaga@ita.br

Abstract. This study is one of a series of studies, researching various aspects that all aim at enhanced simulation based certification aiding for aircraft. An experimental within-group design study was performed with 10 participants (5 male, and 5 female). The results showed a significant difference, $F(2,16) = 5.11$, $p = 0.019$, in mental workload between an engine failure condition and an normal condition for eye blink frequency. No effect of speed at the engine failure event on mental workload was found.

Keywords: Task load · EEG · Eye data · Aircraft · Certification

1 Introduction

This study is one of a series of studies, researching various aspects that all aim at enhanced simulation based certification aiding for aircraft. The research approach is to investigate the potential of pilot behavior models in simulation based certification. To be able to generate and evaluate pilot behavior models, it is necessary to understand the characteristics of key situations related to certification demands [1]. This study provides significant experimental results on the characteristics of an engine failure induced task load transient during simulated take off. When an aircraft is certified, it must be able to stay within 30 ft from the midline of the runway during an engine failure. If an engine failure occurs while the aircraft is taking off, the speed is a crucial component that contributes to whether the pilot decides to abort the take off. Hence, it is important to determine the minimum control speed, or Velocity of Minimum Control on the Ground (Vmcg), during takeoff. Vmcg is defined as the minimum speed, on the

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_8

ground, at which directional control can be maintained with one inoperable engine and full power applied to the other engine, using only the rudder for steering [2].

Piloting an aircraft is at times mentally demanding and at times, task load transients substantially increase the pilot's mental workload, especially during an emergency where decisions need to be made quickly. These kinds of situations, although rare, are important to understand since errors may have fatal consequences. Despite pilots being trained in handling extreme situations and aircraft having systems in place to help them, the most contributing factor to accidents is the behavior of the crew. In some incidents and accidents, poor situational awareness or high mental workload has often been a major factor in the sequence of events leading up to the accident [3].

One approach to measure brain activity is using electroencephalography (EEG) to measure electrical brain waves, which can be detected on the scalp through an elastic cap containing a set of surface electrodes. EEG is a well-tested technique that is related to attention, information processing and mental effort. It correlates with eye movement patterns and subjective measures of mental workload [4]. The high mental workload state is often characterized by increasing EEG power in the theta band and a decrease in alpha band power compared to normal activation [3].

By using an eye-tracking device, one can know where the eyes are focused and assess various mental states, such as attention, drowsiness and consciousness. A good way to measure mental workload is by using an eye-tracking device to capture eye movement. Previous studies have indicated that various cognitive tasks have shown a significant effect on the eye blinking rate [5]. Activities such as speaking increase the blinking rate while, for instance, sustained visual stimuli decreases the blinking rate [5]. During conditions of high demanding tasks, such as driving or flight tasks, the blink rate often decreases. Thus, while exposed to a task that requires higher visual demand, the blinking rate decreases due to an attempt to capture as much visual information as possible.

The aim of this work was to analyze pilots' mental workload when calculating Vmcg during an engine failure. The analysis was performed in a simulated environment. A simulation analysis is to be preferred due to the safety risks involved in a flight experiment. Mental workload in this work is assumed to be related to (1) an increase in EEG power in the theta band, (2) lower EEG power in the alpha band, (3) shorter eye blinks and (4) fewer eye blinks. The most mentally challenging task in the takeoff procedure would be the actual takeoff, thus the mental workload will increase with the proximity to the takeoff. Already [6] stated that mental workload emerges in the interaction between the task requirements, the situation where the task is performed and the operator's skills, behaviors and perceptions. The high task requirements interact with a short time frame in which the task must be performed. With a higher speed, the task of keeping track of all relevant details in the environment has a higher time demand. The first hypothesis is thus that mental workload increases when the speed at the time of engine failure increases. To measure this, three speeds for engine failure have been chosen: 40, 60 and 80 kt. An aircraft with higher speed than 100 kt is most likely going to result in continuing the takeoff [8]. All the speeds are below the recommended takeoff speed of 100 kt as the purpose of the study is to examine engine failure on the ground, not in the air. (1) Higher mental workload should be observed at higher speeds at the time of engine failure. The second hypothesis is that the mental workload will be higher during the engine failure condition than the normal start

condition. This is because the task of calculating Vmcg during an engine failure is mentally demanding, as stated in [7]. (2) Higher mental workload should be observed in the engine failure condition than the normal start condition.

2 Method

Task load transient was induced by simulated engine failure to study participant workload.

A pilot study was conducted prior to the experiments. The participants consisted of 1 male and 2 females. The participants were between the ages of 20 to 25, with a mean of 22 years. After the pilot test, the instructions given to the participants were adjusted to better increase the participants' understanding of the scenarios.

An experimental within-group design study was performed with 10 participants (5 male and 5 female), based on lessons learned from a pilot study with 3 participants. The participants' age ranged from 20 to 38 years, with a mean of 26 years.

Two computers were used to conduct the experiment, using Windows 10 and 7. The Windows 7 computer was used to run the simulation software X-Plane 10 and measuring eye movements. It had an SMI eye-tracker connected to it and software from SMI (iView X, Experiment Center 3.6 and BeGaze 3.6) was used. The simulation software was displayed on an additional screen with attached speakers. A Thrustmaster T.Flight HOTAS X joystick was also used for the simulation. The Windows 10 computer was used to collect EEG data. An EEG headband (Emotiv Epoc+) was connected via Bluetooth. Before using the headband, a multi-purpose solution was used to wet the nodes. Three Python scripts worked were used to monitor and collect EEG data in the form of.csv-files. To analyze the data, Microsoft Excel and Google Spreadsheets were used. Three documents were presented to the participants: a consent form, instructions on how to fly the aircraft in X-Plane 1, and instructions for the experiment scenario.

All participants were given a 5-minute practice session in the flight simulator. The experimenters ensured that the participant knew how to take off from the runway by letting them take off 3–6 times during the session.

The participants were instructed to sit as still as possible, look at the screen and stay silent during the entire test, this to minimize unnecessary noise in the data. The participants were also encouraged to ask any questions before the start of the test. Before starting each experiment, the eye-tracker was calibrated.

After written and oral instructions and after a practice session, the participants performed simulated take offs while EEG and eye data were recorded for assessing workload [3]. The main part of the experiment, the scenarios, consisted of 7 takeoffs from the runway. The participant had been instructed to reach 400 ft. Recommended takeoff speed was stated as 90 knots. There was a 30 s rest between the takeoffs where the participants looked at a blue Windows® home screen. Out of the 7 takeoffs, 3 were of the engine failure condition, meaning that the left engine was turned off at different speeds: 40 kt, 60 kt and 80 kt. Takeoff 3, 5 and 6 were always engine failures. The engine failure time was 20 s if the participant was unable to take off, otherwise it proceeded like the normal condition.

Analysis was performed to investigate if higher mental workload should be observed at higher speeds at the time of engine failure and if higher mental workload should be observed in the engine failure condition than the normal start condition. A within-group design was used to test the hypotheses. To test the influence of speed, four repeated measures ANOVAs were used to compare the three takeoff speeds in each failure condition to our four representations of high mental workload. The dependent variables were: EEG power in the theta-band, EEG power in the alpha band, the eye blink frequency and the eye blink duration. The independent variable was the speed: 40 kt, 60 kt and 80 kt. To test the influence of engine failure on mental workload, the engine failure condition was compared to the normal start condition using four repeated measures ANOVAs. The dependent variables were: EEG power in the theta-band, EEG power in the alpha band, the eye blink frequency and the eye blink duration. The independent variable was: engine failure condition, normal start condition and rest condition. To counteract possible individual differences in the EEG-data all values were standardized with Z-score scaling before analysis. The mean for each condition was then calculated from the Z-score values.

3 Results

Both hypotheses were tested using four ANOVAs. The first test was EEG activation in the theta band, the second EEG activation in the alpha band, the third the eye blink frequency and the fourth the eye blink duration.

First hypothesis: For the first test, the theta band, Mauchly's test indicated that the assumption of sphericity had not been violated, X 2 (2) = 6.25, p = 0.04. The repeated measures ANOVA showed no significant difference in EEG power in the theta band between the three engine failure conditions F(2, 10) = 1.32, p = 0.3. For the second test, the alpha band, Mauchly's test indicated that the assumption of sphericity had been violated, X 2 (2) = 4.34, p = 0.11, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = 0.6). The repeated measures ANOVA showed no significant difference in EEG power in the alpha band between the three engine failure conditions F(1.2, 6.0) = 0.01, p = 0.94. For eye blink frequency, Mauchly's test indicated that the assumption of sphericity had been violated, X 2 (2) = 8.59, p = 0.14, therefore the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = 0.58). The repeated measures ANOVA showed no significant difference in eye blink frequency between the three engine failure conditions F(1.17, 9.37) = 1.19, p = 0.31. For eye blink duration Mauchly's test indicated that the assumption of sphericity had not been violated, X 2(2) = 2.13, p = 0.34. The repeated measures ANOVA showed no significant differences in blink duration between the three engine failure conditions F(2,16) = 0.41, p = 0.66.

Second hypothesis: For the first test, Mauchly's test indicated that the assumption of sphericity had not been violated, X 2(2) = 0.17, p = 0.91. The repeated measures ANOVA showed a significant difference in EEG power in the theta band between the different conditions F(2,10) = 6.43, p = 0,016. However, post hoc Bonferroni showed no significant difference between each of the conditions, takeoff vs rest: p = 0.27, takeoff vs engine failure: p = 0.47, rest vs engine failure: p = 0.06 (See Fig. 1). In the

second test, Mauchly's test indicated that the assumption of sphericity had been violated X 2(2) = 8.73, p = 0.01, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = 0.53). The repeated measures ANOVA showed no significant difference in EEG power in the alpha band between the different conditions F(1.06, 5.29) = 2.24, p = 0.19. For eye blink frequency, Mauchly's test indicated that the assumption of sphericity had not been violated, X 2 (2) = 0.34, p = 0.82. The repeated measures ANOVA showed a significant difference in blink frequency between the different conditions F(2,16) = 5.11, p = 0.019. Post hoc Bonferroni showed a significant difference between the takeoff and engine failure conditions p = 0.046. However, there was no significant difference between the other conditions, takeoff vs rest: p = 1.0, engine failure vs rest: p = 0.06 (See Fig. 2). For eye blink duration, the Mauchly's test indicated that the assumption of sphericity had not been violated, X 2(2) = 2.11, p = 0.34. The repeated measures ANOVA showed no significant difference in blink duration between the different conditions F (2,16) = 1.00, p = 0.38.

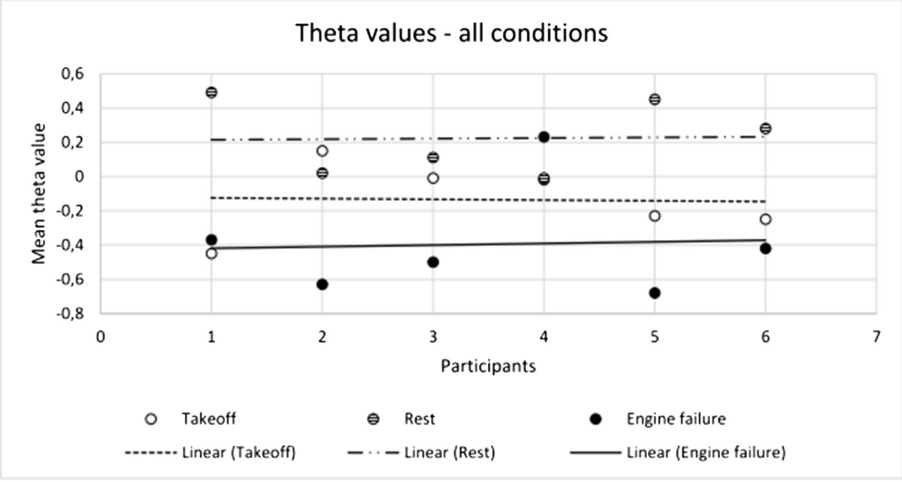


Fig. 1. Theta values – all conditions. The repeated measures ANOVA showed a significant difference in EEG power in the theta band between the different conditions F(2,10) = 6.43, p = 0,016.

Filtered EEG-data also revealed EEG-variations over time, where cycles of pauses and high workload were emergent. The main study was inconclusive with regards to EEG gamma, which is an indicator of cognitive workload. A problem was noise coming from excessive blinking and movement of subjects and problems with our approach to filtering data. However, using basic comparative data analysis, we were able to filter out and extract data from one subject that exemplifies possibilities with regards to workload classification. Figure 3 shows the inverse ratio between a baseline measurement where the subject was resting and the gamma signal during the three

phases of the experiment. The area above the red line indicates periods of decreased cognitive load during a stable measurement period in the middle of the experiment. This example indicates that it could be possible to categorize subjects as actively engaged and attentive based on EEG analysis. However, this requires subject baselines and stable sustained periods of EEG measurements to be able to calculate adequate threshold values.

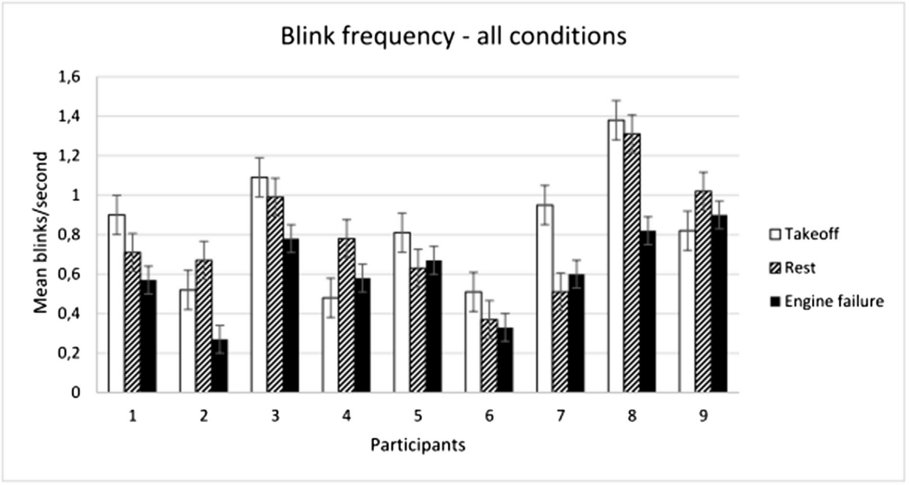


Fig. 2. Blink frequency – all conditions. The repeated measures ANOVA showed a significant difference in blink frequency between the different conditions $F(2,16) = 5.11$, $p = 0.019$.

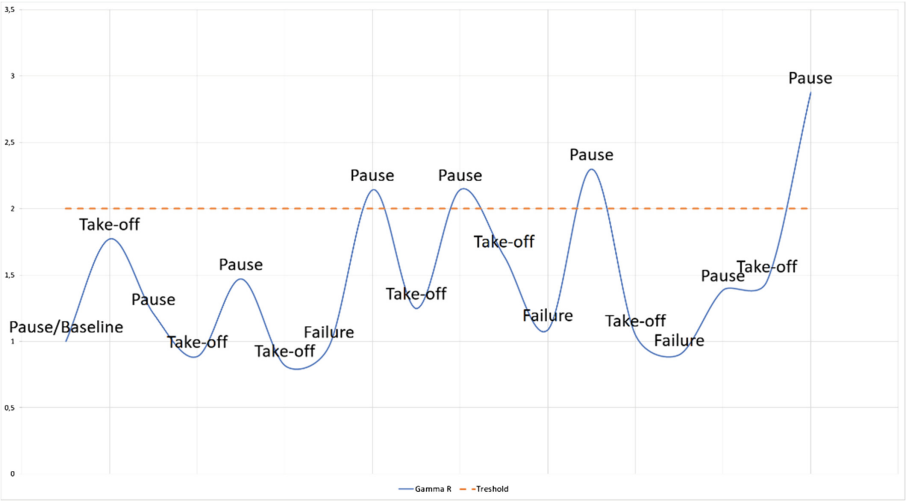


Fig. 3. EEG gamma power level proportional to a baseline during the different phases of an experiment.

4 Discussion

The purpose of the study was to investigate how the mental workload of pilots is affected during an engine failure. Two hypotheses were proposed: (i) higher mental workload should be observed at higher speeds at the time of engine failure, (ii) higher mental workload should be observed in the engine failure condition than the normal start condition. To test the hypotheses, a simulation experiment where the mental workload was measured using EEG and eye-tracking was set up. The first hypothesis was rejected. This could be because of the aircraft being easier to control at higher speeds in combination with mental workload increasing due to higher task and time demands, leading them to cancel each other out. The mental workload might decrease as the takeoff speed increases, for the engine failure case. This might be due to the fact that the aircraft is easier to control aerodynamically at higher speeds. There is support for the second hypothesis. However, only one representation of mental workload (eye blink frequency) showed a significant difference between the normal start condition and the engine failure, where the engine failure mean was 0.61 (SD = 0.07) and the normal start mean was 0.82 (SD = 0.09), following the literature's definition of mental workload in eye blink frequency. The activation of EEG power in the theta band was significantly different between the conditions, but the pairwise comparison showed no significant results. That the results showed a significant difference, $F(2,16) = 5.11$, $p = 0.019$, in mental workload between the engine failure condition and the normal condition for eye blink frequency. The activation of EEG power in the theta band was significantly different between the conditions, but the pairwise comparison showed no significant results. No effect of speed at the engine failure event on mental workload was found. That the EEG data showed very clear variations over time, where cycles of pauses and high workload were emergent are encouraging for future studies of EEG. Maybe a Faraday cage and a larger EEG net can be used to access more and even better data.

5 Conclusions and Future Work

The results of this study provide some guidance and support of future development of simulation based certification aiding for aircraft. The results showed a significant difference in mental workload between the engine failure condition and the normal condition for eye blink frequency. The activation of EEG power in the theta band was significantly different between the conditions, but the pairwise comparison showed no significant results. No effect of speed at the engine failure event on mental workload was found. That the EEG data showed very clear variations over time, where cycles of pauses and high workload were emergent are encouraging for future studies of EEG. A problem is that blinking interferes with the EEG measuring and filtering needs to be conducted.

Future work could verify the results during actual flight conditions. Also, future work could include applied models for aiding certification processes for aircraft.

By examining pilot's mental workload in various situations, an increased understanding of how these situations affect cognitive strain may be obtained. This information might be used to redesign training and alter procedures so that they become less

demanding for the crew. It may also be used to develop devices that alert the crew of their current levels of mental resources. Also, measuring mental workload in an aircraft context is relevant to the human factors field in general where future studies can investigate how mental workload influences performance in specific scenarios using aircraft simulators.

Acknowledgments. This work was in parts financially supported by the Swedish Foundation for Strategic Research.

References

1. Lone, M., Cooke, A.: Review of pilot models used in aircraft flight dynamics. Aerosp. Sci. Technol. **34**, 55–74 (2014)
2. U.S. Government Publishing Office: Aeronautics and Space - §25.149 Minimum Control Speed. http://www.ecfr.gov/cgi-bin/text-idx?SID=68ab268f8d6581e996c3f72924bec39e&mc=true&node=se14.1.25_1149&rgn=div8. Accessed 31 Oct 2016
3. Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., Babiloni, F.: Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci. Biobehav. Rev. **10**(44), 58–75 (2012)
4. Castor, M., Hanson, E., Svensson, E., Nählinder, S., LeBlaye, P., MacLeod, I., Wright, N., Alfredson, J., Agren, L., Berggren, P.: GARTEUR Handbook of Mental Workload Measurement. Grouo for Aeronatical Research and Technology in Europe, Flight Mechanics Action Group FM AG 13 (2003)
5. Oh, J., Jeong, S.-Y., Jeong, J.: The timing and temporal patterns of eye blinking are dynamically modulated by attention. Hum. Mov. Sci. **31**, 1353–1365 (2012)
6. Hart, B., Stavenland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, pp. 139–183. Elsevier, New York (1988)
7. Airbus: Flight Operations Briefing Notes (2004). http://www.airbus.com/fileadmin/media_gallery/files/safety_library_items/AirbusSafetyLib_-FLT_OPS-TOFF_DEP_SEQ07.pdf
8. Airbus: Revisiting the "Stop or Go" Decision, 1-11 (2005). http://www.airbus.com/fileadmin/media_gallery/files/safety_library_items/AirbusSafetyLib_-FLT_OPS-TOFF_DEP-SEQ04.pdf

An Analysis of Human Factor Aspects in Operational Fuel Saving

Daniel Vogel[1,2(✉)], Ivan Sikora[1], and Hans-Joachim Ruff-Stahl[2]

[1] City, University of London, SEMS, Tait Building – C128, Northampton Square London, EC1V 0HB London, UK
{Daniel.Vogel.1, Ivan.Sikora.1}@city.ac.uk
[2] College of Aeronautics, Embry Riddle Aeronautical University, Berlin Campus, Kurfuerstenstrasse 56, 10785 Berlin, Germany
Ruffha@erau.edu

Abstract. Over the last few years, the reduction of operational costs and control of pollutant emissions have become central issues for the commercial aviation industry, and as a result, airlines have been increasingly focusing their attention on operational fuel saving techniques. However, even though the practical implementation and economic potential of these techniques have been exemplified in a number of papers, little research has been dedicated to a systemic investigation of the effects of operational fuel saving on the human component of the system, i.e., the flight crew of an aircraft. This research examines this area, and investigates the human factors aspects in context with the application of operational fuel saving on the Airbus A 320 series aircraft. The study presents a detailed analysis of the flight crew´s performance and motivational factors related to the topic of interest, which were investigated by means of an online survey and a controlled simulator experiment. Results of the analysis revealed that the application of operational fuel saving imposes a number of latent performance impairments on the flight crew. Motivational factors were shown to be disrupted by the flight crew´s inability to achieve satisfaction from the application of operational fuel saving. The implications of these findings are wide-ranging, as they show, in essence, that the system´s safety and efficiency relies solely on the flight crew´s cognitive flexibility and workload compensation capability, while structured analyses and conceptual frameworks in regard to the human factors aspects of operational fuel saving are absent.

Keywords: Human factors · Workload analysis · Situational awareness analysis · Error management · Safety engineering

1 Introduction

Driven by unprecedented growth rates and constantly evolving markets, commercial aviation is one of the most competitive business fields worldwide. Characterized by intense market pressure and heavy overcapacities [1, 2], it is especially prone to macro- and microeconomic factors such as new business models, predatory competition or ruinous pricing [1–3]. While these factors are widely cited and broadly acknowledged in the literature, a more in depth look reveals one variable that seems constantly

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_9

re-appearing in context with the industry's profitability concerns: The cost of fuel. The "traumatic" years 1974 and 1975 [2], the "deep crises" between 1979 and 1983 and between 1990 and 1993, and the years following the 9/11 attacks, which "turned crises into disaster"[2, p.20], were all aggravated by fuel costs that "escalated at an alarming rate"[2, p.17]. To mitigate the impact of fuel price developments, aircraft manufacturers, aviation organizations and individual operators have begun, within the last few years, to foster operational programs to reduce fuel consumption [4–6]. However, even though these programs increased fuel efficiencies significantly, they could not completely offset the intensified cost pressure on the airlines. Since the mid-2000 s an unprecedented number of airlines collapsed.

To make the situation even more complex, the economic pressure on the industry is only "half of the story", as the necessity to increase efficiency is paralleled by the requirement to improve safety. Although recent global industry accident rates have been among the best ever in the history of commercial aviation [7, 8] there are some factors that might raise concern: First, the huge improvements in commercial aviation safety have slowed significantly within the last few years [9], suggesting that existing approaches have reached their limits. Second, contemporary research [10, 11] has identified an increasing number of fuel saving-related safety issues in commercial air transport, underlining that new safety challenges have emerged. Third, worldwide demand for commercial air transportation is expected to increase over the next two decades [12, 13], implying that the total number of aviation mishaps will likewise increase regardless of low accident rates.

In reflection of these considerations, it becomes clear that future risks to safety for the commercial aviation industry are as complex as mitigating them will be an ambitious task: These emerging challenges will make it even harder to balance efficiency and safety as the most fundamental needs to economically survive as an airline. While both requirements diametrically oppose each other [14], the aviation industry will have to find ways to increase operational efficiency and flight safety *at the same time*.

2 Methodology

The aim of this research was the investigation of the human factor aspects in operational fuel saving. The research methodology included an online survey and a controlled simulator experiment. The object of research for this study was the variation of the dependent variable human performance (as defined by subjective mental workload, situation awareness and human error rates) under the influence of the different fuel saving techniques and the measurement of the flight crews' motivational factors towards the application of operational fuel saving. The research design was of cross-sectional nature and consisted of a qualitative survey that intended to retrieve real world data from "the field" and a quantitative experimental part under laboratory conditions (the simulator experiment). This design was believed to be the most suitable for this research, as cross-sectional approaches enable the analysis of the effects under investigation at a particular point in time rather than over a specific period [15]. By this, confounding variables, such as learning - or "panel effects", were assumed to be minimal. Combining a laboratory experiment and a questionnaire was motivated by the

intention to achieve the highest possible internal and external validity by balancing the weaknesses of the respective design type with the strengths of the other.

2.1 Description of the Questionnaire

The online questionnaire served two main purposes; besides collecting real world data the aim was to investigate the flight crews' motivational factors towards operational fuel saving. The questionnaire was pre-tested based on contemporary questionnaire research [16–18] and embraced a "multi-method pre-test" design, i.e., a combination of several independent methodological components. In particular, it employed a dyadic procedure, coalescing a "conventional" or "standard" pre-test with a slightly modified derivative of the more contemporary "think aloud" technique and included twenty A320 rated pilots. The reliability of the instrument was evaluated by comparing the answers of the same respondent to the same question at two measurement points and revealed a high positive correlation (Pearson coefficients ranging from +0.64 to +0.81), which was considered to be a strong indicator of the instrument's reliability. The validity of the instrument was assessed with a variant of the "think aloud" methodology and a statistical analysis that intended to analyze patterns of correlation within specific questions of the survey. The statistical analysis of these patterns revealed average overall effects ranging from moderate (Person correlation coefficient +0.47) to highly significant (Person correlation coefficient +0.93). In combination with the findings of the ´think aloud` methodology, these results were considered to be supportive indicators of the validity of the instrument. The link to the online survey was disseminated to all identified commercial A320 operators in Europe via different distribution channels such as direct email, pilot forums (e.g., the British Airline Pilots Association forum), professional networks and promotion flyers. The promotion flyers were developed in English, German and Spanish. Based on a fleet list derived from www.ch-aviation.ch, 103 European commercial Airbus 320 series operators with a total fleet size of 1.749 aircraft were identified. Assuming an average crew factor of nine pilots per aircraft (on average 4.5 Captains and 4.5 First Officers), the total population size was therefore estimated to be approximately 15.800 pilots. The sample size of the research consisted of N = 1,076 subjects, whereas N = 854 (79.4%) were male and N = 222 (20.6%) female. The majority (N = 444 or 41.3%) of the subjects was at the time of their response holding the rank of a Commander, while 274 (25.5%) were holding the rank of a First Officer. 187 subjects (17.4%) were holding the rank of a Senior First Officer, while two subjects (0.2%) were Second Officers. 155 subjects (14.4%) were holding the rank of a Training Commanders (i.e. Type Rating Examiner, Type Rating Instructor or Line Training Commander), whereas 14 subjects (1.3%) were so called "Management Pilots", i.e., Fleet Chief, Chief Pilot or Director of Flight Operations.

2.2 Description of the Simulator Experiment

The simulator experiment was conducted in an Airbus A320 Level-D full flight simulator. The primary purpose of this experiment was to measure the flight crews'

performance (as operationalized through the constructs of subjective mental workload, situational awareness and human error) under the application of the respective fuel saving techniques in a line oriented flight scenario.

The conceptual set-up of the flight scenario followed a "within subject", repeated measures design, which means that each subject traversed each scenario in a predefined order. The decision for the employment of this design was based on the aim to balance out subject variance (such as age, background, experience or personality) and to create a highly sensitive and well-controllable environment. Besides this, its utilization allowed to derive 48 answers from 24 subjects (as each subject performed under both conditions of the experiment). In order to evaluate the flight crews' performance within a realistic scenario, the experiment encompassed a two leg flight scenario from Hamburg (ICAO Code EDDH) to Berlin Tegel (ICAO Code EDDT) and back with an Airbus A320-214, engine configuration CFM 56-5B4, under standard atmospheric conditions (i.e. standard pressure and temperature distribution).

The aircraft's ramp mass was set to 64.200 kg for both flights, which corresponded to a planned take-off mass of 64.000 kg (rounded taxi fuel 200 kg).

The flight from Hamburg to Berlin constituted the experimental part of the scenario and included the manipulation of the independent variables, i.e., application of the respective fuel saving techniques. Twenty-four Airbus A320 type rated pilots (23 Male, 1 Female), all of them employed by European airlines, participated in the simulator experiment. Twelve of the participants were holding the rank of a Commander at the time of the study, whereas 12 pilots were holding the rank of a First Officer. The Commanders' total flight experience ranged from 3.200 to 20.000 h (M = 11.558; SD = 4.326) with an average time on type of 6.200 h (SD = 3.220). Commanders were aged between 37 and 51 years (M = 44.5; SD = 5.4). Of the 12 Commanders, five were additionally rated as instructor pilot (i.e. Type Rating Instructor or Line Training Commander). Notably, seven out of 12 Commanders had a total flight time of 12.000 h or more. First Officers were aged between 22 and 37 years (M = 28.0; SD = 5.0) with a mean total flight time of 2.450 h (Min = 400; Max = 5.000; SD = 1.433) and a mean time on type of 1.725 h (Min = 100; Max = 4.500; SD = 1.354). One of the First Officers was additionally rated as Training First Officer (Synthetic Flight Instructor). Four of the First Officers had a total flight time of 3.000 h or more. All flight crews were trained according to, and familiar with, the Standard Operating Procedures of the aircraft including the respective fuel saving techniques under investigation.

All odd numbered flight crews began with the experimental scenario, i.e., the flight from Hamburg to Berlin, whereas all even numbered flight crews started with the control scenario, i.e., the flight from Berlin to Hamburg. This alternating order design was chosen to control (through randomization) the possible occurrence of "order effects", which are believed to occur in context with repeated measures designs [19]. The flight from Hamburg to Berlin was performed according to the Standard Operating Procedures of the aircraft including the application of the following supplementary fuel saving techniques: Single Engine Taxi Out, Packs Off for Take-Off, Flaps 1 Take-Off (instead of Flaps 3), Reduction of Acceleration Altitude from 1.500 feet to 800 feet, Visual Approach, Flaps 3 Landing (instead of Flaps Full), Idle Reverse (instead of Full Reverse), Single Engine Taxi In. In order to shape the scenario as close to reality as possible, the role distribution between pilot flying and pilot monitoring was left at the

flight crew's decision. However, all subjects were briefed that they must change roles between pilot flying and pilot monitoring after the first leg, so that each subject performed one leg as pilot flying and one leg as pilot monitoring.

At predefined stages of the flight, the subjects' mental workload and situational awareness were measured with computer applications of the NASA TLX test by Hart and Staveland [20] and the Situational Awareness Global Assessment Technique (SAGAT) put forth by SA Technologies [21]. Besides this, the test persons' number and types of error were collected and subjectively classified by the researcher against the backdrop of the Human Error Template (HET) External Error Mode (EEM) taxonomy provided by Marshall et al. [22].

3 Results

3.1 Findings of the Online Survey

Although the general existence of operational fuel saving was broadly acknowledged by the great majority of participants, the individual techniques' real world application frequencies varied significantly. Specifically, single engine ground operations, short and/or visual approaches and landings with reduced flap settings were reported to be applied at least, whereas magnitudes of application were testified to be highest for fuel saving techniques targeting the take-off/climb and cruise phase of the flight.

Besides this, more than 90 per cent of the subjects were of the opinion that operational fuel saving should be assessed in terms of flight safety relevance prior to implementation, while more than 40 per cent of the test persons reported that they have experienced flight safety relevant issues and/or degradations of personal performance during its application. However, only two per cent of the respondents conveyed involvement in a serious incident or accident in which the application of operational fuel saving was identified as a primary cause or contributing factor. The rank-specific analysis of answers revealed, in parts, significant disparities between the individual pilot groups, whereas the answers of Management Pilots were in some cases directly opposite to those of all other ranks.

The investigation of the flight crews' motivational factors revealed a general acceptance of the necessity to conserve fuel, as reflected in a weighted mean agreement value of 3.82 on a scale ranging from 1 to 5. Consistent with these findings, 64 per cent of the respondents acknowledged the requirement for their organization to increase their current fuel saving efforts, while more than 70 per cent of the subjects considered the application of operational fuel saving in general, and themselves in particular, as the "right instrument" to achieve the overall goal of cost reduction. The great majority of test persons considered themselves to be suitably qualified for the application of operational fuel saving in daily flight operations, while more than 64 per cent believed its implementation demonstrated professionalism by the flight crew. However, despite these degrees of perceived abilities and individual role perception, the magnitude of agreement dropped significantly when subjects were asked whether they associated operational fuel saving with an improvement in flight safety.

The analysis of extrinsic rewards revealed a rather low degree of perceived positive reinforcement, as more than half of the respondents did not feel their fuel saving efforts to be appreciated by their company. Beyond that, more than 40 per cent of the test persons reported perceived pressure to conserve fuel.

The investigation of subjects' perceived "company framework conditions" revealed low magnitudes of perceived framework suitability (38 per cent) and employee co-determination (weighted mean value of 1.81). Finally, the rank-specific breakdown of answers revealed again a significantly different, in parts even diametrically opposite, perspective on the topic on the side of the Management Pilots.

3.2 Results of the Simulator Experiment

The human factors-related findings of the simulator experiment conveyed a highly insightful picture of operational fuel conservation, as they disclosed a number of potentially hazardous and largely unrecognised effects associated with the topic of interest. These effects can be summarized as follows: First, an increase of subjective mental workload, which provides reasonable evidence to believe that operational fuel conservation basically imposes additional cognitive demands on the flight crew. In particular, single engine ground operations were found to impose significant workload increases on the flight crew, whereas perceived workload increase was moderate for the take-off and climb and approach and landing phase of the flight.

Simulator data further revealed that flight crews are basically able to deal with additional cognitive demands imposed by the application of the respective fuel saving techniques, but only up to a certain extent, as clearly evidenced in a statistically significant, strong correlation between increased mental workload and human error for all flight phases that included the application of operational fuel saving techniques. Besides, the degree to which increased demands imposed by a particular fuel saving technique will result in decreased performance was found to be not constant, but to vary significantly. In particular, single engine taxi operations and visual approaches with reduced landing flap setting were found to be especially error-prone and characterized by a mean workload value that was higher than that for a dual-engine go-around maneuver, which was used as a reference. Significantly, it was exactly those fuel saving techniques that were reported to be applied least in the field. In contrast, the additional demands of fuel saving techniques during the take-off/climb and cruise phase of the flight could - at least under the given circumstances - be compensated for to the greatest extent by the flight crews, as indicated in far lesser extents of flight safety relevant performance impairments. As these fuel saving techniques were reported to be applied most often in the field, it can be hypothesized that flight crews are cognizant of the relationship between their own metal capabilities and the demands imposed by a particular fuel saving technique in a particular situation, which means that they realize "what can be done safely and what cannot be done safely" and thus avoid potentially dangerous or demanding situations.

In consequence, the human factors-related impacts of operational fuel saving should not be viewed as an immediate and obvious threat for aviation safety, but as a latent, hidden condition that has the potential of impairing the safety of flight.

4 Discussion and Implications

The implications of these findings are far reaching. They exemplify, in essence, that it is currently solely the flight crew, or more precisely, the combination of their cognitive flexibility, workload compensation capability and the careful selection of when and when not to apply a specific fuel saving technique that determines the safety and efficiency of the "system component" operational fuel conservation.

With regard to aviation safety, this implies that a more in-depth understanding of the flight crew's cognitive flexibility and decision making is required, as these factors constitute currently the last (and only) barrier, which is protecting the "ultra-safe" commercial aviation system from the consequences of active failures of a flight crew in a state of mental overload.

However, since modelling the entire bandwidth of possible scenarios would be difficult if not impossible, working with different "levels of difficulty" (i.e. frameworks representing a particular level of difficulty) is recommended. In this light, prospective research endeavors could adopt the simulator set-up employed in this study and modify it to investigate the impacts imposed by the respective fuel saving techniques in "typical" low, low to medium, medium, medium to high and high complexity environmental scenarios. To achieve a workable basis for this environmental classification, subject matter experts (e.g. safety practitioners from airlines, accident investigators and human factors researchers) could be consulted via more "in-depth" instruments, such as standardized questionnaires, structured interviews and discussion groups. By moving these investigations forward, further light could not only be shed on the reliability of the existing "protective barriers", but also on the general relationship between mental workload and human performance and the extent to which a specific variable, such as visibility or airport complexity, contributes to the total workload perceived.

In this context, it appears as well that additional focus should be given to the "role" of the aircraft, or, more precisely, to its "operational philosophy", which seems - at least in parts - to be "incongruent" with the economic reality surrounding the airlines. For example, taxi out and taxi in procedures are designed, trained and performed on the presupposition that the aircraft is taxied with both engines operating, whereas the intensified cost pressure acting on the airline would actually require single engine taxi operations to be "the norm".

Next, the findings of this dissertation suggest that fuel saving education should be integrated into the earliest stages of pilot instruction, such as contemporary Multi-Crew Pilot License (MPL) training concepts, as the revealed correlation between experience and situation awareness indicates that pilots can rely upon existing mental models to deal with additional cognitive demands. Besides this, it appears that the investigations of this study should be extended to other twin, wing-mounted engine aircraft of similar size, mass and operating profiles, such as the Boeing 737 NG series or the Tupolew Tu-204 /214. Although it is recognized that the technical specifications and operating procedures of these aircraft vary significantly from those of the Airbus A320, their inclusion could nevertheless unveil additional insights into the "role" of the individual aircraft, or more precisely, into its "operational suitability" for the application of operational fuel saving techniques.

Besides this, the observed selective application strategies of the flight crew indicate the lack of organizations to provide clear, standardized structures and procedures for when to apply a specific fuel saving technique in a specific scenario. As a result of this "shift of responsibility", pilots are required to weigh protective versus productive goals on an individual basis in an environment often characterized by ill-structured information settings and the requirement to make decisions under time pressure. Yet, as decision making under these circumstances is most prone to potentially dangerous misjudgments, it becomes clear that possible effects arising from the implementation of operational fuel saving must be recognized and captured at the organizational level, i.e., before they reach the flight deck of the aircraft.

In order to provide this recognition, and to achieve structured and effective data for the assessment of individual fuel saving techniques within the realm of a company's Safety Management System (SMS), it is therefore recommended that existing and operationally verified risk assessment frameworks, such as the Operational Risk Assessment (ORA) methodology proposed by the Airline Risk Management Solutions (ARMS) Working Group (2010), be utilized for this endeavor. The ORA methodology is considered by its designers to be an "end to end Risk Assessment process, starting from Hazard Identification and leading to Safety Actions" [12, 23] and is sub-divided into two parts - an Event Risk Classification (ERC - employed for the assessment of incoming data from already happened safety events) and a Safety Issue Analysis (SIRA - employed for the analysis of data coming from databases or Safety Assessments). Due to this dyadic design, the methodology seems promising for an evaluation of past fuel saving-related safety events as well as existent and planned fuel saving techniques in an environment that is expected to grow continuously [24].

References

1. Clark, P.: Stormy Skies: Airlines in Crisis. Ashgate Publishing Ltd, Farnham (2010)
2. Doganis, R.: Flying Off Course: The Economics of International Airlines. Taylor & Francis, London (2009)
3. Holloway, S.: Straight and Level: Practical Airline Economics, 3rd edn. Ashgate Publishing Ltd, Aldershot (2008)
4. Airbus.: Getting to Grips with A320 Family Performance Retention and Fuel savings. Blagnac, France: Airbus Flight Operations Support & Services (2008)
5. International Civil Aviation Organization (ICAO): Operational Opportunities to Minimize Fuel Use and Reduce Emissions. ICAO, Montreal (2003)
6. International Air Transport Association (IATA): Guidance Material and Best Practices for Fuel and Environmental Management, 3rd edn. IATA Press, Montreal (2008)
7. International Air Transport Association (IATA): 2012 Aviation Safety Performance. IATA, Montreal (2014)
8. International Civil Aviation Organization (ICAO).: 2013 Safety Report. Montreal, Canada (2013)
9. The Boeing Company.: Statistical Summary of Commercial Jet Airplane Accidents: Worldwide Operations 1959–2012. Seattle, WA: The Boeing Company. (2013) http://www.boeing.com/news/techissues/pdf/statsum.pdf. Accessed 07 April 2014

10. Beer, P.W.: Fuel & Airmanship. Paper presented at the IFALPA Accident Analysis & Prevention Committee (AAP) meeting, 23–24 September 2010, Copenhagen, Denmark. Montreal: International Federation of Airline Pilots´ Associations (IFALPA) (2010)
11. International Federation of Air Line Pilots 'Associations (IFALPA): Airmanship and the Fuel Tank – the results of the IFALPA fuel use survey. IFALPA, Montreal (2011)
12. Airbus.: Airbus Global Market Forecast 2012-2032. Navigating the future. [Online] Blagnac, France: Airbus. (2012) http://www.airbus.com/presscentre/corporate-information/key-documents/?eID=dam_frontend_push&docID=25773. Accessed 05 Oct 2013
13. The Boeing Company.: Current Market Outlook 2013-2032. [Online] Seattle, WA: The Boeing Company. (2013) http://www.boeing.com/assets/pdf/commercial/cmo/pdf/Boeing_Current_Market_Outlook_2013.pdf. Accessed 04 Jan 2014
14. Reason, J.: Managing the Risks of Organizational Accidents. Ashgate Publishing Ltd, Farnham (1997)
15. Rindfleisch, A., Malter, A.J., Ganesan, S., Moorman, C.: Cross-Sectional versus Longitudinal Survey Research: Concepts, Findings, and Guidelines. Journal of Marketing Research, XLV, pp. 261–279 (2008)
16. Oksenberg, L., Cannell, C., Kalton, G.: New Strategies for Pretesting Survey Questions. J. Off. Stat. **7**, 349–365 (1991)
17. Porst, R.: Fragebogen: Ein Arbeitsbuch. Springer, Wiesbaden (2011)
18. Presser, S., Blair, J.: Survey pretesting: do different methods produce different results? Sociol. Methodol. **24**, 73–104 (1994)
19. Miller, S.: Experimental Design and Statistics. Routledge, London (1984)
20. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, pp. 239–250. North Holland Press, Amsterdam (1988)
21. Technologies, S.A.: Super SAGAT User Manual. SA Technologies Inc, Marietta, GA (2004)
22. Marshall, A., Stanton, N. A., Young, M., Salmon, P. M., Harris, D., Demagalski, J., Waldmann, T., Dekker, S.: Development of the Human Error Template – a new methodology for assessing design induced errors on aircraft flight decks (Final Report of the ERRORPRED Project E!1970). London: Department of Trade and IndustrySA Technologies (2004) Super SAGAT User Manual. Marietta, GA: SA Technologies Inc. (2003)
23. Airline Risk Management Solutions (ARMS). (2010) The ARMS Methodology for Operational Risk Assessment in Aviation Organisations. http://www.easa.eu.int/essi/documents/Methodology.pdf. Accessed 05 Feb 2013
24. International Air Transport Association (IATA) (2014) Passenger demand growth stays strong [Online] Montreal, Canada: IATA. http://www.iata.org/pressroom/pr/Pages/2013-10-02-01.aspx. Accessed 02 Feb 2014

Transition from Conventionally to Remotely Piloted Aircraft – Investigation of Possible Impacts on Function Allocation and Information Accessibility Using Cognitive Work Analysis Methods

Max Friedrich(✉), Anne Papenfuß, and Andreas Hasselberg

German Aerospace Center (DLR), Lilienthalplatz 7, 38108 Brunswick, Germany
max.friedrich@dlr.de

Abstract. Methods from cognitive work analysis were applied to investigate impacts of transitioning from conventional to remote aircraft control on function allocation and information accessibility. An abstraction hierarchy for an Airbus A320 was constructed, followed by an analysis of work functions and function allocation between pilot flying, pilot monitoring, and automation during approach and landing phases. Next, it was analyzed how removing the pilots from the flight deck to a remote location might affect the accessibility of important information that conventional pilots use for decision-making. The results revealed function specific information remote pilots have limited access to, which however seem to be important for decision making (such as visual cues from the environment). Furthermore, based on a supervisory control concept, a function reallocation between one pilot and automation was developed. The analyses contribute to deriving information requirements for remote pilot stations designed for remote single pilot operations in controlled airspace.

Keywords: Cognitive work analysis · Remotely Piloted Aircraft Systems

1 Introduction

Currently much effort is devoted to developing concepts for the integration of remotely piloted aircraft systems (RPAS) into controlled airspace. The FAA's roadmap for the integration of RPAS into the controlled airspace system [1] as well as the ICAO Doc 10019 [2] require RPAS i.a. to comply with instrumental flight rules and to be able to react to Air Traffic Control (ATC) instructions. Autonomous operation is not permitted; consequently a human pilot has to control the aircraft remotely. Further, in the scientific literature it has been suggested to lower the pilot-aircraft ratio from currently 2:1 to 1:1 or even further to 1:n, leaving one pilot in command of one or several aircraft simultaneously. Due to constantly advancing technology, the shift from remote single to remote multiple aircraft control will eventually become possible in the near future. In the following, implications and challenges resulting from transitioning from dual to single pilot operations (i.e. pilot-aircraft ratio of 1:1) as well as transitioning from

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_10

conventional to remote aircraft control will be discussed for medium altitude long endurance (MALE) RPAS.

The transition from dual to single pilot operations has several implications, one being the need of defining novel function allocations. Functions that are currently distributed between pilot flying (PF), pilot monitoring (PM) and automation need to be reallocated between two actors, namely the single pilot and the automation. Clearly defined allocation of functions is particularly important for defining information requirements for the future remote pilot station (RPS), since it defines the scope of functions that need to be executed with the devices integrated in the RPS.

Moving the pilot from the flight deck on board the aircraft to a remote location on the ground has implications on the accessibility of information. Due to the physical distances between the actual aircraft and the ground station, delays in control inputs and updates of system parameters are expected, as well as a lack of environmental sensory information including the visual window view, acoustic information and angular accelerations. Williams [3] lists a set of issues that are related to the use of such sensory information, including the detection of turbulence, unusual attitude and stall.

Possible negative impacts on the quality of the pilot's decisions due to the limited accessibility of information need to be avoided. As such, detailed analyses of information sources are required, aiming to identify function and situation specific information that is important for making decisions. If the analyses reveal that certain information that is critical for the safe execution of functions is not accessible or only accessible with difficulties (e.g. due to a critical time delay), this has to be considered in the development of operational concepts for the integration of RPAS into controlled airspace and in the design of the RPS.

The aim of the current study was to investigate impacts of lowering the pilot-aircraft ratio to 1:1 and transitioning from conventional to remote aircraft control on function allocation and information accessibility. The results contribute to deriving information requirements for RPS designed for remote single pilot operations in controlled airspace. First, the work domain of conventional pilots of an Airbus A320 was analyzed. Next, an analysis of work functions and function allocation between pilot flying, pilot monitoring, and the aircraft's automation during approach and landing phases was conducted. It was then analyzed how removing the pilot from the flight deck to a remote location affects the accessibility of important information that conventional pilots use for decision making. Further, a function reallocation, based on a supervisory control concept between one pilot and automation was developed.

2 Method

The methods included in the analysis were the abstraction hierarchy [4], the social organization and cooperation analysis [4] combined with a contextual activity template [5, 6] and the decision ladder [5, 7]. Each method will briefly be introduced in the following subsections followed by a short description of how it was applied. Figure 1 visualizes the methods used (left column) and their outputs (other columns).

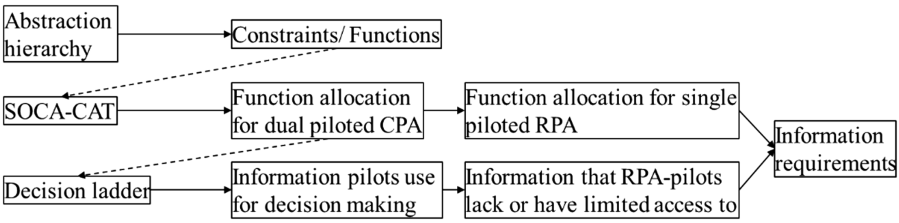


Fig. 1. Illustration of the used methods and desired outputs. CPA = Conventionally Piloted Aircraft; SOCA-CAT = Social Organization & Cooperation Analysis-Contextual Activity Template

2.1 Work Domain Analysis (Abstraction Hierarchy)

In order to analyze a work domain, it is important to get an understanding about the possible scope of action that is limited by the physical boundaries and constraints that are present in the work system. One method for this purpose is the development of an abstraction hierarchy. An abstraction hierarchy visualizes the functions of a work system on five levels of abstraction, ranging from functional purposes to physical forms. These functions in turn define the scope of action, i.e. the constraints of the work system (for further details please refer to [4]).

Since the purposes of conventionally piloted aircraft (CPA) and future RPAS are identical, i.e. maneuvering the aircraft through controlled airspace while complying with the same rules, the work domain of a conventional pilot was used as a basis for exploring the work domain of a remote pilot. Accordingly, an abstraction hierarchy for an Airbus A320 aircraft flying in controlled airspace was developed.

2.2 Activity and Function Allocation Analysis (SOCA-CAT)

Activity within a work system can be investigated by further analyzing the functions that were included in the abstraction hierarchy in terms of their respective contribution in specific work situations. For this purpose, the functions are mapped onto a so called contextual activity template (CAT, [5]). A CAT is a matrix with the horizontal axis visualizing situations and the vertical axis visualizing functions. Solid bars indicate situations in which the function should be active in order to fulfill its purpose. Dotted boxes indicate that the respective function can be active, as opposed to must be active.

Analysis of function allocation between different actors (e.g. pilot flying, pilot monitoring or automation) can be accomplished by conducting a social organization and cooperation analysis (SOCA, [4]). One way to visualize the results of the analysis is to fill the bars in the CAT with different colors, each representing a different actor. A CAT, combined with a SOCA might be referred to as SOCA-CAT (e.g. [6]).

In the current study, a SOCA-CAT, including the pilot flying, pilot monitoring and aircraft automation, was developed for the approach and landing phases of an A320. In a next step, the same analysis was carried out for supervisory single pilot aircraft control with actors including the single pilot and aircraft automation.

2.3 Analysis of Decision Making and Classification of Information Sources (Decision Ladder)

The decision ladder [5, 7] is a method for analyzing decision making processes. A decision ladder was developed for the approach and landing phases of an A320. In a next step, information sources that were identified in the decision ladder were classified into the categories *system bound information* and *environmental sensory information* according to the respective source. The aim of the classification was to analyze what information a remote pilot potentially lacks due to the physical distance to the aircraft in comparison to a conventional pilot sitting in the flight deck.

3 Results

In the following, the results of each method introduced in Sect. 2 will be presented in separate sections. Finally, in Sect. 3.4 the developed function allocation for supervisory single pilot control of an RPA will be presented.

3.1 Abstraction Hierarchy

The abstraction hierarchy is presented in Fig. 2. On the first level, the functional purposes level, two major purposes of an A320 are shown, namely *safe and efficient flight operation in shared, controlled airspace*. On the abstract functions level, conditions that need to be met in order to fulfill *safe, efficient flight in shared, controlled airspace* are shown. These conditions are defined as keeping the *flight parameters within the aircraft's flight envelope*, insurance of *separation to traffic, terrain, obstacles and meteorological risks*, proper *energy management* and *health and functionality of the aircraft systems*.

Functions that are necessary for fulfilling the purposes of *safe and efficient flight in shared, controlled airspace* were defined as *aviation*, *navigation*, *communication* and *protection & supply for systems* and are represented on the generalized functions level. For the construction of the physical functions level, selected aircraft systems as classified by the ATA 100 chapters for classification of aircraft systems were analyzed. Aircraft systems were classified according to their respective functional contribution to the four generalized functions. The selection and classification of relevant aircraft systems was carried out in cooperation with a retired Airbus A320 pilot. For example, the aircraft system flight controls (ATA Chap. 27) has the function of controlling the *attitude* of an aircraft and was thus classified as contributing to the generalized function *aviation*. Hence, the function *attitude* has to be executed in order to (partly) fulfill the function *aviate* and was labelled as one of the physical functions. The same procedure was repeated for all aircraft systems that were considered in the analysis.

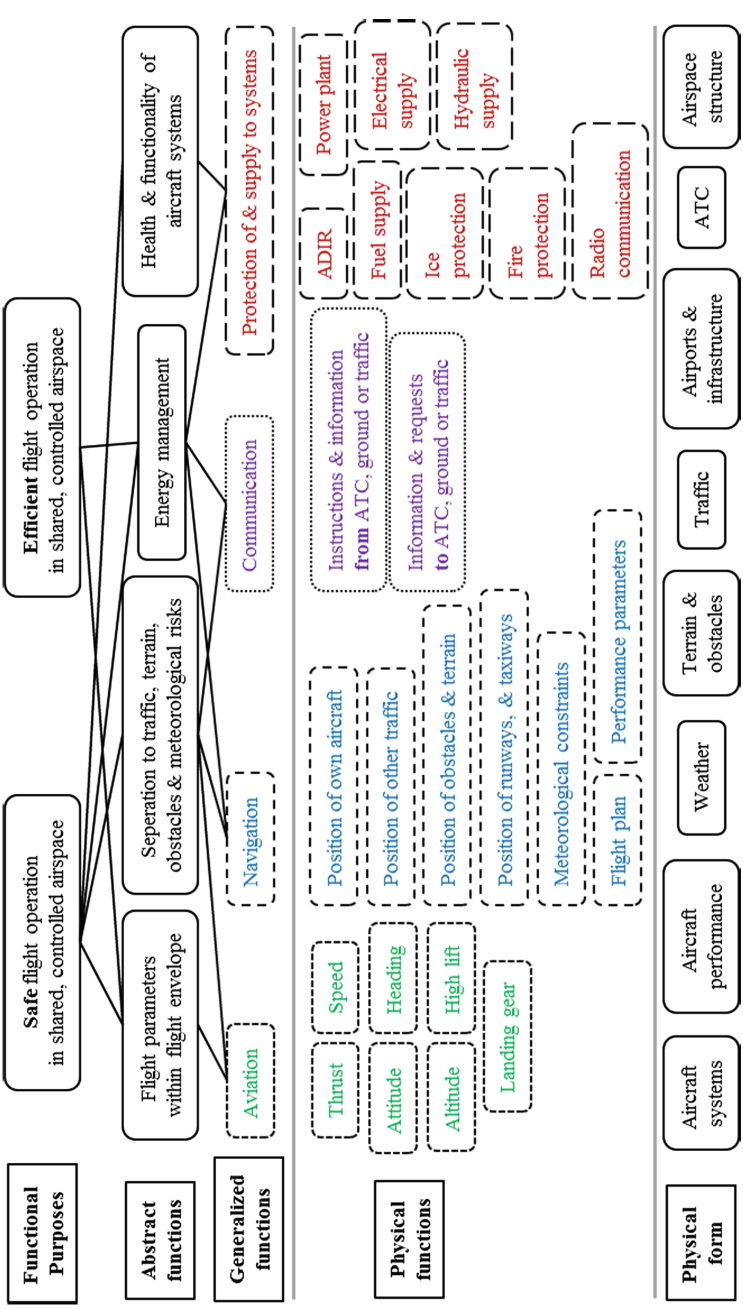


Fig. 2. Abstraction hierarchy for an Airbus A320 flying in controlled airspace. ADIR = Air Data & Inertial Reference

On the physical forms level, the actual physical objects of the aircraft itself (i.e. the aircraft systems) and of the environment (i.e. terrain, obstacles, traffic and airports) as well as constraints regarding aircraft performance, weather conditions, airport infrastructure, ATC and airspace structure are represented. Due to the limited space, the aircraft systems that were included in the analysis are not shown individually.

3.2 Results of the Analysis of Function Allocation in an Airbus A320 Aircraft

To illustrate the results, Fig. 3 shows the allocation of selected functions, which were considered as being primarily important during approach and landing phases of an Airbus A320 under normal procedure conditions. The functions represent the physical functions level of the abstraction hierarchy. The coding "Optional automation" means that the pilots can execute the function without automation, but automation can be beneficial and in cases of pilot's error even essential for the safe execution of the function. For example, during the approach phases, altitude can be controlled by either the PF alone (manual control), automation alone (e.g. during automated landings) or the automation assists the PF by e.g. indicating the glide slope on the primary flight display. However, automation must still be monitored. As such, for the altitude control function, the PF is in charge throughout all the approach phases until touch down, even when automation is engaged. Thus, automation may be engaged to assist or even take over control but is still optional in the sense that the pilot would still be able to control the aircraft's altitude manually. The same logic can be applied to every other function in Fig. 3.

3.3 Results of the Analysis on Decision Making and Classification of Information Sources

In a next step, the decision making process during approach and landing with regard to the selected functions in Fig. 3 was described using a decision ladder. The results are shown in Table 1. Information that pilots gather for identifying current system states were defined with regard to the goals pilots have during approach and landing under normal procedure conditions. Goals were defined as *Comply with ALT & SPD constraints*, *Criteria for stabilized approach*, *Stay clear of traffic, obstacles & terrain*, *Stay clear of hazardous meteorological conditions* and *Comply with ATC instructions*. In order to achieve these goals, pilots gather information from the system and the environment (shown in the row titled as *Information*). The target states were defined with respect to the goals (e.g. for the goal *Comply with ALT & SPD constraints* the target system state would be *ALT & SPD within constraints*). Information was categorized into system bound information and environmental sensory information. The results of the categorization are visualized in Fig. 4. The figure shows that sensory information is present in every function, except for the functions *Communicate with ATC* and *Manage radio communication systems*.

Fig. 3. SOCA-CAT for an Airbus A320 during approach and landing phases

3.4 Function Reallocation for a Single Piloted RPA

Current control concepts of RPA (e.g. Global Hawk) assume that automation takes over *Aviation* and (parts of) the *Navigation* functions (as defined on the physical functions level of the abstraction hierarchy shown in Fig. 2), leaving the remote pilot in charge of supervising the automation and configuring systems. Based on such a supervisory control concept, a SOCA was conducted with one remote pilot and

Table 1. Decision ladder for approach and landing phases of Airbus A320 pilots

State of knowledge	Value
Goals	• Comply with ALT & SPD constraints • Criteria for stabilized approach • Stay clear of traffic, obstacles & terrain • Stay clear of hazardous meteorological conditions • Comply with ATC instructions
Activation	
Alert	Passing intermediate fix point (IF)
Observe information	
Information	**System** • Barometric & radio altitude (ALT) • Vertical speed/sink rate (V/S) • Indicated airspeed (IAS) • Heading (HDG) • Roll index, pitch angle, turn indicator • Localizer (LOC) & glide slope (G/S) • Flap indication, flap lever position • Gear indication & gear lever position • Primary engine parameters, thrust lever position • ATC instructions via radio **Environment/Sensory information** • Window view • Angular accelerations
Identify state	
System state	Identified by PF & PM
Predict consequences	
Options	Evaluated by PF & PM
Evaluate performance	
Chosen goal	Evaluated by PF & PM
Predict consequences	
Target state	• ALT & SPD within constraints • IAS & V/S within limits for stabilized approach • Minimum deviations from RWY-HDG & G/S • Correct power setting engaged • Aircraft in correct landing configuration • No traffic, obstacles or terrain on flight path • No meteorological risks on flight path • ALT, SPD & HDG as instructed by ATC
Define task	
Task	Defined by PF & PM (if System state≠Target state)
Formulate procedure	
Procedure	As defined by the aircraft's operating manual

System Environment/ Sensory information

Flight phase / **Function**	Interm. Approach	Final Approach	Touch down	Landing Roll
Control altitude	ALT, V/S, G/S			
	Window view			
	Angular acceleration			
Control speed	IAS			
	Window view			
	Angular acceleration			
Control thrust	THR, N1, EGT			
	Thrust lever position			
	Angular acceleration			
Control heading	HDG, LOC			
	Window view			
Control attitude	Roll index, pitch angle, turn indicator			
	Side stick deflection			
	Window view			
	Angular acceleration			
Manage high lift devices	Flaps indication			
	Flaps lever position			
	Angular acceleration			
Manage landing gear	Gear indication			
	Gear lever position			
	Angular acceleration			
Determine position of own aircraft	Indication on ND			
	Window view			
Determine position of other traffic	Indication on ND			
	Window view			
Determine position of obstacles & terrain	Indication on ND			
	Window view			
Determine meteorological constraints	Indication on ND			
	Window view			
	Angular acceleration			
	Vibration			
Determine position of runways & taxiways	LOC & G/S			
	Window view			
Communicate with ATC	Radio transmitted voice			
Manage radio communication systems	Inidication of radio frequencies			

Fig. 4. Information sources for conventional pilots of an Airbus A320 during approach and landing phases

automation for the same functions that were included in the SOCA-CAT of an A320 (Fig. 3). For situations in which functions are allocated to automation, the concept assumes that the functions are executed by automation while the remote pilot supervises. When the remote pilot receives instructions from ATC concerning altitude, heading or speed changes, s/he configures the systems by sending the required commands to the aircraft while automation executes the required functions (e.g. reduces thrust to reduce speed). This is labelled as "Optional remote pilot". Further, the remote pilot should intervene in cases of abnormal attitudes and false thrust settings. Figure 5 illustrates the results.

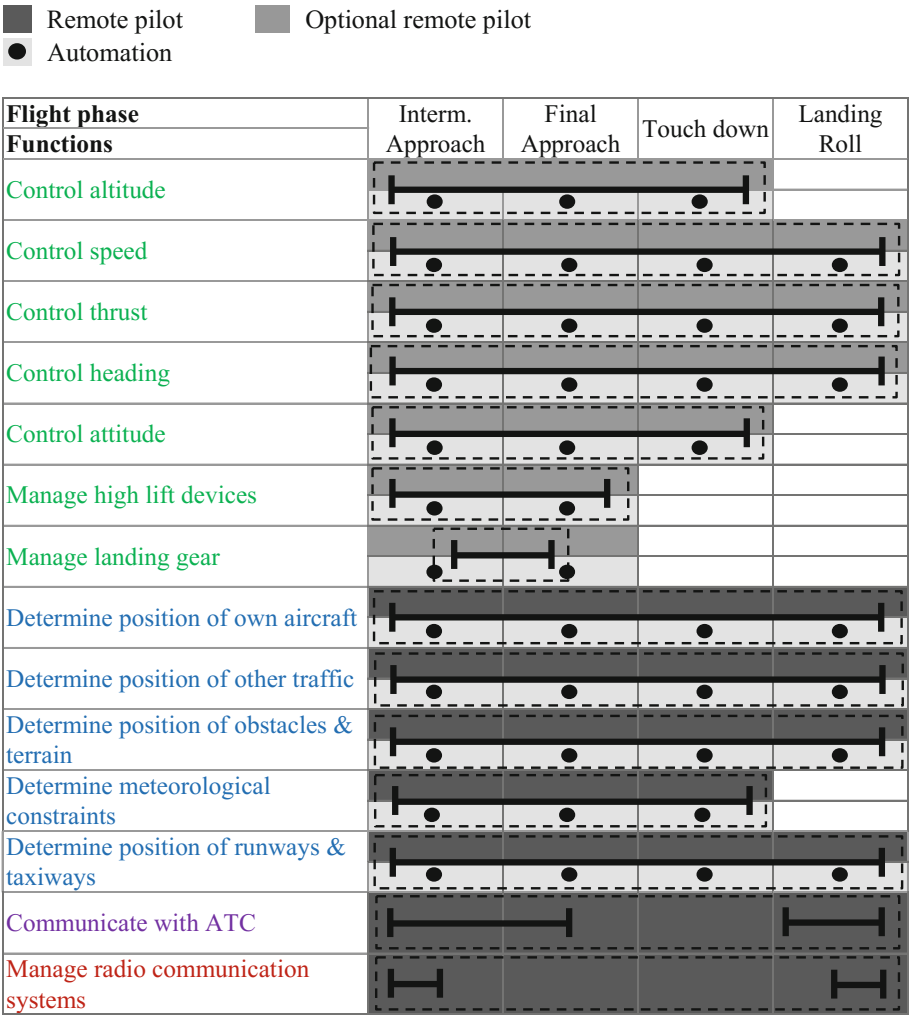


Fig. 5. SOCA-CAT showing a reallocation of functions for future single piloted RPAs during approach and landing phases, based on a supervisory control concept.

4 Discussion

In the presented study, methods from cognitive work analysis were applied to investigate impacts of transitioning from conventional to remote aircraft control on function allocation and information accessibility. The results contribute to deriving information requirements for RPS designed for remote single pilot operations in controlled airspace. The envisioned role of the remote pilot will mainly be to supervise the automation and configure the systems by e.g. entering a new speed according to ATC instructions, while automation executes the functions necessary to achieve the desired system state (i.e. reducing thrust in order to reduce the aircraft's speed to the instructed speed value).

The shift to supervisory control has major implications on the control concept of an aircraft. Whereas in current CPAs automation is optional to use (see Fig. 3), in future RPAs the pilot will optionally intervene (Fig. 5). In order to intervene accurately and quickly, the remote pilot must have awareness of the aircraft's state. As such, the RPS should provide the remote pilot with real time (i.e. with minimum time delay) information on flight parameters, system states and the aircraft's position relative to other aircraft, terrain, obstacles, runways and taxiways. In order to gain awareness of the aircraft's state, sensory information seems to play a role, however its importance is not clear yet. Especially during landing phases, the information that is gathered from the environment seems to play a major role for deciding whether the aircraft is behaving normally or not and whether any (counter) measures need to be initiated. If the pilot does not have access to sensory information anymore, this might heavily affect the quality of the decisions. As such, if the pilot is supervising the aircraft from a remote location with limited access to potentially important sensory information but is supposed to intervene in situations when automation fails, the RPS should provide the remote pilot with means to compensate for the loss of sensory information. Concerning the determination of the own aircraft's position, the RPS could either provide the remote pilot with a window view through cameras installed on board the aircraft (especially during final approach and ground operation) or provide other means of accurately defining the aircraft's position, relative to obstacles, terrain and traffic.

A major implication of shifting from a dual piloted to a single piloted control concept is that the dual control principle is lost. Current operation principles of CPA require continuous cross checking of system states and actions from both pilots. This is naturally not possible anymore for single (remote) pilot operations. Thought should be given on how to deal with this issue. One approach would be to include a supervisory system, supervising the automation and assisting the (remote) pilot with the supervision task. Considering the decision making process presented in Fig. 4, identification of the current system state (done by PF and PM) could then be done by the (remote) pilot and the supervisory system together. The supervisory system could for example inform the (remote) pilot of a deviating parameter, while responsibility of defining a new target state and initiating actions would be left to the (remote) pilot. Another approach would be to leave the entire responsibility to either the (remote) pilot or the supervisory system. Were the supervisory system responsible, it might automatically define a new target state and command the aircraft's automation to execute a certain (counter) measure without interference of the (remote) pilot. The (remote) pilot would merely be

informed of the automation's actions. Research is needed to determine how interaction between a potential supervisory system and the (remote) pilot could optimally proceed.

The developed function reallocation for single piloted RPAs is based on a supervisory control concept. This approach necessitates high degrees of automation involvement. Functions could also be reallocated differently with less automation involvement. However, giving the remote pilot more responsibility of controlling for example *aviate* functions in a sensory deprived environment would not change the information requirements that resulted from the analysis. The RPS should still compensate for the loss of sensory information. Further, if in the future the pilot-aircraft ratio should decrease even more, leaving the remote pilot in charge of several aircraft simultaneously, automation will most likely play a crucial role when realizing RPS for multiple RPA control. As such, monitoring system states will always be a major part of the work of a remote pilot.

Studies should be conducted that explore the importance of sensory information for decision-making. Future research should also focus on the development of system designs that consider effects of the physical distance between aircraft and RPS (such as time delays) and that compensate for the sensory deprivation of the remote pilot. In order to counteract negative impacts of monitoring activity, novel concepts and techniques should be explored for guiding the remote pilot's attention, especially during critical situations.

References

1. Federal Aviation Administration: Integration of civil unmanned aircraft systems (UAS) in the national airspace system (NAS) roadmap, Washington (2013)
2. International Civil Aviation Organization: Manual on remotely piloted aircraft systems (RPAS) (Doc 10019 AN/507), Montreal (2015)
3. Williams, K.W.: Documentation of sensory information in the operation of unmanned aircraft systems (Report No. DOT/FAA/AM-08/23). Federal Aviation Administration, Washington, DC (2008)
4. Vicente, K.J.: Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. Lawrence Erlbaum Associates, Mahwah, New Jersey (1999)
5. Naikar, N., Moylan, A., Pearce, B.: Analysing activity in complex systems with cognitive work analysis: concepts, guidelines and case study for control task analysis. Theor. Issues Ergon. Sci. **7**(4), 371–394 (2006)
6. Stanton, N.A., Harris, D., Starr, A.: The future flight deck: modelling dual, single and distributed crewing options. Appl. Ergon. **53**, 331–342 (2016)
7. Rasmussen, J., Pejtersen, A.M., Goodstein, L.P.: Cognitive Systems Engineering. Wiley, New York (1994)

Workload and the En Route Controller – An Overblown Issue?

Richard W. Rohde(✉)

Fort Hill Group, 660 Pennsylvania Avenue SE, Suite 204, Washington, DC 20003, USA
Rory.Rohde@FortHillGroup.com

Abstract. This is the first in a series of papers on the mental process of the En Route Air Traffic Controller. These papers will explore Situational Awareness, Mental Models, Workload, and a variety of related issues in an attempt to both improve the research community's understanding of the En Route Controller and to enable more productive and applicable future research activities. This paper will describe the learning process and experiences for a typical Developmental trainee.

Keywords: Air traffic control · Mental Models · Human factors · Situation awareness · Training

1 Introduction

Workload is one of the most studied and debated topics in the domain of En Route Air Traffic Control. Countless papers introduce themselves with nearly identical paragraphs about the forecast of large increases in the amount of air traffic in the coming decades, followed by the authors' ideas of technology or other methods of coping with it, usually supported by the results of a Human in the Loop (HITL) simulation they have designed.

The implication is that ever-increasing traffic will result in taskloads that will overwhelm controllers without improved technology, resulting in sector overload, reduced efficiency, increased likelihood of Loss-of-Separation events, or the non-spoken worst case scenario of a mid-air incident.

Employing an extensive literature review combined with the author's in-depth Air Traffic Control knowledge and experience, this paper will examine those implications and the overall topic of workload in the En Route environment.

Please keep in mind that what follows applies only to the En Route environment. Terminal/TRACON operations share many similarities with En Route, but not enough to include them. Tower operations are completely different and a subject for someone with more expertise in that area.

2 Definitions

As several different papers have pointed out (Murphy 2004, etc.), there isn't complete agreement on what is meant by the term 'workload'. Coming up with a conclusive

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_11

definition is beyond the scope of this paper, but for the purposes of this paper a few working descriptions will be defined.

Taskload – Taskload is the major component of workload, and is defined as the demands being put on the controller to safely and efficiently manage traffic situations while complying with all directives such as LOAs, SOPs, and ATC rules (the 7110.65 in the US). Histon and Hansman [1] call this **Situation Complexity.** Kopardekar and Magyarits [2] refers to this as **Dynamic Density**.

Taskload is more than just the situation at the sector. Air Traffic Control is a cooperative endeavor, especially at the En Route level. "Controllers are team players who must coordinate their actions and plans with pilots and many other controllers." [3]

Workload – Is defined as the individual controller's perception of the complexity of the Taskload. This will vary due to the controller's overall skill level and mental and physical state at that moment. Histon and Hansman refer to this as **Cognitive Complexity.**

Situational Awareness (SA) – Situational Awareness (SA) is the 3-dimensional picture of the sector and traffic that the controllers have in their heads. The controller uses radar and other information sources to update Situational Awareness, and then uses this mental picture to make decisions. Controllers refer to SA as 'The Flick' or 'The Picture'.

3 Why Does Workload Matter?

As noted earlier, the primary concern about excessive workload is that it could lead to overload, resulting in a loss of SA and potentially leading to a serious LOS event.[1]

However, several studies such as Hilburn's *Cognitive Complexity in Air Traffic Control* [4] show no significant correlation has yet been shown between workload and errors.

Redding et al. [5] posited that "…errors tend to occur most frequently during periods of relatively light workload…committed by good controllers who…fail to monitor an aircraft plainly on their scope. These results suggest that lack of vigilance in performing the key cognitive task Maintaining Situational Awareness is a frequent source of error." [It must be noted here that there is serious concern that many of the proposed initiatives to decrease controller workload will result in an overreliance on automation increasing instances of lost vigilance].

A second concern is the loss of efficiency. After safety, the main mission of most ANSPs is to increase efficiency through programs such as free-flight, TBFM, Optimized Profile Descents, etc. As Hilburn put it, "excessive complexity drives taskload, which indirectly drives workload, which raises the risk of overload, which ultimately sets an upper limit on sector capacity." [4] As the sector gets busier, the controller will be able to offer fewer additional "services" such as shortcuts, discretionary descents, and advisories.

[1] The fact that as of this writing there has never been a major accident in the En Route environment in the US attributable to controller error somewhat belies this concern, though one could argue that this a "Black Swan" statistic.

But as Howell et al. [6] showed in their research, once Center traffic levels reach 30%, the average inefficiency remains mostly constant, with only a slight increase after 70%. Further, Myers et al. [7] showed that En Route delays are an insignificant issue with their 2005 study which found that the average flight was delayed an average of less than 2 min above 12,000 ft. This figure includes unavoidable weather delays.

At this point one might wonder if things have changed since these studies were done over 10 years ago.

They have. Domestic Air Carrier Traffic Operations have actually decreased in the US by 20% in the period from 2005 – 2016[2]. Eurocontrol reports a slight decline in overall air traffic operations for the same period[3]. Traffic will inevitably increase again, especially as UAS's start to become IFR qualified, but the projections of increasing traffic and congestion have been significantly delayed at the very least.

For the sake of argument, let's assume that traffic will grow again sometime in the near future at such a rate that efficiency will be significantly affected and safety somewhat compromised. What is the best way to measure workload and test initiatives designed to mitigate it?

4 Current Methods for Mitigating Workload

In the cockpit, workload is a very serious issue, especially during takeoff and landing. Taskload is highest at the beginning and ending of the flight which, according to separate studies by Boeing [8] and Airbus [9], is also when almost 90% of all accidents occur. The departure and arrival phases of flight is where fuel is most likely to be a concern - too much to land again quickly after takeoff, not enough on arrival to explore other options. This is also the phase of flight where there is little room to recover from an altitude mistake. And if the flight crew does become overloaded and consequently makes an error that leads to an accident, they will personally suffer the physical effects.

The En Route controller, on the other hand, has many resources to call upon when taskload begins to exceed capacity. And if an LOS occurs, it may be as minor as two aircraft passing within 4.9 miles of each other. In most cases, the most serious repercussion would be a temporary decertification followed by a short period of remedial OJT. But there are several barriers in place to prevent excess taskload from arising.

The first line of defense for sector overload is the Traffic Management Unit (TMU). Well before the sector reaches at saturation point, Traffic Management Coordinators (TMCs) monitor projected traffic levels throughout the center using tools such as Traffic Situation Displays (TSDs). When a sector's traffic levels are projected to rise near capacity, the TMC will alert the Front Line Manager (FLM) for the affected sector. Together they will work out a strategy to mitigate the sector overload, most likely involving restricting the rate of traffic coming into the sector or rerouting aircraft around the sector, laterally or vertically.

[2] https://www.transtats.bts.gov/Data_Elements.aspx?Data=2.
[3] http://www.eurocontrol.int/articles/statistics.

If the TMC doesn't notice the pending saturation alert in time, the second line of defense is the FLM. The FLM also has access to a TSD, and has the additional advantage of being physically located near the affected radar teams. FLMs are more aware of basic traffic patterns for sectors in the area and know what times of day different sectors generally become busy. They are also aware of the capabilities of the controller(s) staffing the affected sector. The FLM can mitigate sector overload by alerting TMU, using proactive tactical maneuvers, increasing sector staffing, and monitoring the sector themselves either directly or remotely.

Sector staffing is a shared responsibility. Unlike a flight crew, the controller can get help[4]. The controller often is first one to recognize that weather, such as turbulence or storm cells, is disrupting the traffic flow. Both the controller and the FLM can see the URET screen or strip bay and notice if it is becoming full of aircraft. A good FLM will also be attuned to clues that an individual controller is getting busy such as louder or faster transmissions, swinging a leg, moving around in her seat, etc.

At a one-person sector a data (planning) controller can be added to assist the radar controller. If still more help is needed a third person in the form of a coordinator (L-side) can also be inserted to add an extra set of eyes and handle interphone calls. If it is a combined sector, it can be split. As a last resort, traffic coming into the sector can be shut off until the workload has returned to a safe and manageable level.

"Air traffic control is so intrinsically the product of a team that part of the work of each controller can often be done by his or her colleagues if the controller is inexperienced, overloaded, or for any reason less efficient than normal." – Hopkin [10]

[The one shift where both strategic and tactical assistance are not readily available is the midnight shift. Typically, TMU is not staffed overnight and there is one supervisor responsible for all the controllers on duty. Two certified controllers are on duty for each area, but once traffic slows on or the other is usually on an extended break. While there were a great number of factors that led to the mid-air over Uberlingen, these were a couple of the major ones cited in the official report [11].]

At a tactical level the controller has many tools to manage building workload, mostly in the form of "task shedding." "The competent controller generally knows the rate at which he or she can complete tasks, and this knowledge is actively managed by the ATCO to avoid overload." – Loft [12]

As taskload rises, there are several strategies controllers can use for mitigation. The controller will prioritize all potential tasks, deferring the lower priority ones if possible or "shedding" them if necessary. [1]

Both Corver et al. [13] and Knorr and Walter [14] identify "Trajectory Uncertainty" as a major contributor to controller workload. This uncertainty can be lessened by "reducing degrees of freedom" [1] through the use of "Positive Control", which usually involves vertical separation. Aircraft will be leveled off or issued crossing restrictions that ensure separation.

[4] There are rare exceptions, the most famous being the crash-landing of UAL232 in Sioux City where an experienced fourth pilot was able to assist the three-man crew and most certainly prevented the results of the accident from being much worse.

Other task shedding strategies can be employed. They can eliminate optional advisories, revert to minimal phraseology, and employ strategies to reduce the need for coordination. In rare circumstances they may even use unconventional tactics such as asking the flight crew to coordinate a speed or heading when checking on the next frequency.

Sperandio [15] observed that "controllers used a variety of adaptive strategies, such as processing fewer aircraft variables and reducing verbal c high taskload."

"This is called *workload management*. If possible, they revert to routine actions, standard procedures and 'simple' solutions that need less attention and that gain time, for instance, by a lower load of radiotelephony. Depending on the evolving situation (routine – non-routine), they switch between low and high workload (cf. vigilance)." – Oprins [16]

This is similar to driving a car on a busy highway and encountering weather and/or unpredictable traffic. The driver will put away their cell phone and sit up and scan constantly. They can also adjust their speed, turn off audio entertainment that might be distracting, and even cease conversing with other passengers.

5 Measuring Workload

Most studies that include measurements of controller workload use subjective methods. Controllers are asked self-rate how busy they were during a simulation, most often using NASA-TLX. The self-reporting, post-event nature of the NASA-TLX may limit the applicability of findings to real world scenarios.

Different controllers will evaluate taskloads differently. Even controllers of similar abilities may define "very demanding mentally" differently – what Histon and Hansman refer to as **Perceived Complexity**. Because the rating process usually takes place after the scenario to avoid interfering with SA, the subjects may not accurately remember how "busy" they felt, nor can it provide a dynamic assessment of workload. On top of that controllers, like all professionals, are proud of their skills and less likely to admit to being "overloaded". Additionally, most studies don't use current controllers but due to financial and/or logistical reasons must rely on students or retired controllers and supervisors. Neither of these groups will have the skills to be representative of current controllers, potentially further compromising the validity of certain types of studies.

Thus, the only real use for a subjective measurement like this is to compare an individual's subjective rating in the same study under different conditions (with or without a Decision Support Tool, etc.).

Before Boeing or Airbus delivers a new model passenger jet to the major airlines, it undergoes many tests. One of the more impressive of these tests is the wing stress tests, which you can see on-line through a simple search[5]. The test is exactly what it says – increasing pressure is exerted on the wing until it snaps. This is an objective measurement of "wing overload".

[5] One of many examples is here: [https://www.youtube.com/watch?v=rak2HldVp9M].

If we truly want to a more accurate measurement on Controller workload, why don't we use a similar procedure in our HITLs? Of the innumerable studies read for this paper where attempts are made to measure "workload", none have attempted to calibrate the workload scale. How can we expect whomever is rating controller workload to give an accurate assessment if they can only see the "zero end" of the scale?

The process would be very simple. Design the first scenario of whatever simulation is being run to be open-ended with an ever-increasing workload, like a game of *Tetris*. Explain to the controller beforehand the exact goal of the scenario; for them to experience "Level 7" (or whatever the highest rating on the scale being used is) so that in the subsequent trials there will be a "yardstick" to more accurately measure workload. Then run the scenario until workload gets maxed out and overload occurs. To lessen subjectivity further, "overload" could be defined as when an LOS or some other benchmark occurs.

Secondly, as Manning [17] notes, a simple solution for making subjective workload assessments more accurate and standardized is to have another person providing the assessment during the scenario. Another controller would be ideal because both for obvious reasons and because they are experienced at knowing how busy other controllers are, a necessary skill for providing OJT, which is a requirement of all controllers. This would result in more objective dynamic workload measurements throughout the scenario.

Another option is to obtain objective measurements through techniques such as secondary performance-based measurements. However, for secondary performance-based measurements to be valid, they must duplicate the primary task being measured. If visual performance is being measured, such as scanning, then the secondary task must involve additional scanning. If cognitive workload is being measured, the secondary tasks must be designed so that it is always cognitive. But this presents its own difficulties - asking the participant to answer 11 times 12 could be cognitive for one person and non-cognitive for someone else (System 1 vs. System 2 in the lexicon of Daniel Kahneman) [18].

Secondly, there must be a reason for the participants to complete the secondary task while also performing the simulation to the best of their abilities. If the subject is not motivated to complete both tasks, then the performance-based measurement will not be effective.

The Situation Presence Assessment Method (SPAM) shows some promise as noted by Zhang [19], but he also goes on to note that "no such study has been conducted (on SPAM) using high-fidelity ATC simulations with subjects who have ATC experience".

Here again it benefits all concerned if the difficulty level of the scenario could be "maxed out" to provide a defined level where overload occurs. However, to be able to truly create a scenario with a taskload that can be ratcheted up to guarantee maxing out the controller's workload abilities, the HITLs will need to be designed to be much more realistic.

6 HITL Scenario Design

The topic of HITL scenario design could fill an entire paper.

The following is a brief example of the lack of realism in most simulations:

Scenario Controller: "American Ninety-Five, descend and maintain one-one thousand."

Remote Pilot responding instantly: "American Ninety-Five roger, descending to one-one thousand."

Within a few seconds the simulated aircraft begins a relatively rapid descent at a uniform rate to 11,000 ft.

Total time elapsed, less than 15 s.

Compare that with:

Controller: "American Ninety-Five, descend and maintain one-one thousand."

"(Garbled due to blocked transmission) roger, descending to one-one thousand."

"All aircraft standby. That last transmission was for American Ninety-Five. American Ninety-Five, did you copy your descent to one-one thousand?"

"American Ninety-Five, roger, descending out of 2-3-0 for eleven thousand."

The controller must now make another broadcast to make sure another aircraft didn't answer the original clearance.

Total time elapsed, at least 30 s.

All this time he must put his "to-do list" on hold and now that 30 s have passed priorities may have changed. As he is about to key up to issue another clearance he receives a call on the interphone with a request from another controller. Meanwhile, American 95 is just starting the descent with a mode C readout of 229…

There are many ways in which HITLs do not accurately simulate actual air traffic conditions, too many to cover here. But lack of trajectory uncertainty and blocked, unanswered, or incorrect readbacks are two major ones that directly affect simulated workload.

Testing proposed new technologies under high workload situations is especially important because that is when the technology is needed. Wiener and Curry [20] pointed out that the effect of automation often is to reduce workload when the workload is low and to increase it when the workload is high. Nadine Sarter, an industrial engineer at the University of Michigan, and one of the pre-eminent researchers in the field, made the same point in a different way: "Look, as automation level goes up, the help provided goes up, workload is lowered, and all the expected benefits are achieved. But then if the automation in some way fails, there is a significant price to pay. We need to think about whether there is a level where you get considerable benefits from the automation but if something goes wrong the pilot can still handle it." [21]

7 Conclusion

While the growth of Air Traffic Operations in the US and Europe has stalled over the last 10 years, it is highly likely that this trend will be reversed and traffic will grow over the next decade. Even without this growth, if NextGen initiatives are successful in

increasing acceptance rates at high capacity airports, at the very least the En Route sectors near that airport will experience a corresponding increase in traffic.

To maintain the efficiency and safety of the En Route ATC system in the US, effective ways will need to be found to manage this increased taskload. To do that, those researching potential strategies need to understand exactly what workload is for the En Route Controller, how it works, and how to effectively test and measure it. Hopefully this paper has provided a step in that direction.

References

1. Histon, J., Hansman, J.: Mitigating Complexity in Air Traffic Control: The Role of Structure-Based Abstractions. MIT ICAT, Cambridge (2008)
2. Kopardekar, P., Magyarits, S.: Measurement and prediction of dynamic density. In: Paper Presented at the 5th USA/Europe ATM 2003 R&D Seminar, June 2003, Budapest, Hungary
3. D'Arcy, J., Della Rocco, P.: Air Traffic Control Specialist Decision Making and Strategic Planning—A Field Survey. FAA-DOT, March 2001
4. Hilburn, B.: Cognitive complexity in air traffic control: a literature review. In: EUROCONTROL. Brussels, Belgium (2004)
5. Redding, R., Ryder, J., Seamseter, T., Purcell, J., Cannon, J.: Cognitive Task Analysis of En Route Air Traffic Control: Model Extension and Validation. Human Technology, Inc., McLean (1992)
6. Howell, D., Bennett, M., Bonn, J.: Estimating the en route efficiency benefits pool. In: 5th USA/Europe Air Traffic Management R&D Seminar. Budapest, Hungary
7. Myers, T., Brennan, M., Klopfenstein, M.: A new approach to analyzing airborne delay. J. Guidance Control Dyn. **31**, 1793–1801 (2008)
8. Boeing: Statistical Summary of Commercial Jet Airplane Accidents, Worldwide Operations 1959–2015. Boeing, Seattle (2015)
9. Airbus: A Statistical Analysis of Commercial Aviation Accidents 1958–2015 (2015)
10. Hopkin, D.: The measurement of the air traffic controller. Hum. Factors, J. Hum. Factors Ergon. Soc. **22**, 547–560 (1980)
11. Johnson, C.W.: Final Report: RFQ/137/04 Review of the BFU's Überlingen Accident Report (2004). http://www.dcs.gla.ac.uk/~johnson/Eurocontrol/Ueberlingen/Ueberlingen_Final_Report.PDF
12. Loft, S., Sanderson, P., Neal, A., Mooij, M.: Modeling and predicting mental workload in en route air traffic control: critical review and broader implications. J. Hum. Factors Ergon. Soc. **49**(3), 376–399 (2007)
13. Corver, S.C., Unger, D., Grote, G.: Predicting air traffic controller workload: trajectory uncertainty as the moderator of the indirect effect of traffic density on controller workload through traffic conflict. Hum. Factors (2016). doi:10.1177/0018720816639418
14. Knorr, D., Walter, L.: Trajectory uncertainty and the impact on sector complexity and workload. In: SESAR Innovation Days, Nov–Dec 2011
15. Sperandio, J.C.: Variation of operator's strategies and regulating effects on workload. Ergonomics **14**, 571–577 (1971)
16. Oprins, E., Zwaaf, D., Eriksson, F., van de Merwe, K., Roe, R.: Impact of future time-based operations of situation awareness of air traffic controllers. In: ATM 2009
17. Manning, C.: Measuring Air Traffic Control Performance in a High-Fidelity Simulation. CAMI, FAA, January (2000)
18. Kahneman, D.: Thinking Fast and Slow. Farrar, Straus, and Giroux, New York (2011)

19. Zhang, C.: The Effect of Situation Presence Assessment Method on Air Traffic Control Students' Workload and Performance in High-Fidelity Simulations. Arizona State University (2016)
20. Wiener, E.L., Curry, R.E.: Flight-deck automation: promises and problems. Ergonomics **23**, 995–1011 (1980)
21. Langweisch, W.: The human factor In: Vanity Fair, October (2014)

Data-Driven Pilot Behavior Modeling Applied to an Aircraft Offset Landing Task

Felipe M.S. Turetta[1(✉)], Helon Vicente Hultmann Ayala[2], Luís G. Trabasso[3,5], Leandro S. Coelho[2], and Jens Alfredson[4,5]

[1] EMBRAER - Systems Modeling and Simulation, Av. Jose Candido Da Silveira 2000, Belo Horizonte, Brazil
felipe.turetta@embraer.com.br
[2] PUCPR – Industrial and Systems Engineering Graduate Program, R. Imaculada Conceição, 1155, Curitiba, Brazil
helonayala@gmail.com, leandro.coelho@pucpr.br
[3] Department of Mechanical Engineering, Instituto Tecnológico de Aeronáutica, DCTA, São José dos Campos, Brazil
gonzaga@ita.br
[4] Saab – Aeronautics, Bröderna Ugglas Gata, Linköping, Sweden
Jens.Alfredson@Saabgroup.com
[5] Department of Computer and Information Science, IDA, Linköping University, Linköping, Sweden
{luis.gonzaga.trabasso,Jens.Alfredson}@liu.se

Abstract. This paper shows studies for the development of a mathematical model that adequately represents a pilot behavior in the specific task of offset landing, using data-driven modeling techniques. Flight test data was used for the identification procedure. Considerations on the pilot's cognitive process and mathematical modeling possibilities were discussed to select the most appropriate inputs and outputs for the model. This data was used to identify the model using artificial neural network techniques. The models obtained were validated against the identification data and different data not used in the training process to evaluate the quality of the models. Conclusions include the difficulties of showing the generalization capabilities of those non-linear models and further studies.

Keywords: Human factors · Human pilot modeling · Neural networks

1 Introduction

The use of modeling and simulation techniques is rising quickly in every field of engineering as computational power grows and the need for to detect design issues early in the development process rises to shorten lead times, lower costs and increase product maturity. However, for tasks requiring a human operator in the loop, the usefulness of simulation may be reduced. That is because it may be difficult to model human behavior and include it in the simulations. Even if human in the loop simulations are used, this

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_12

prevents the analyst to run many simulations (such as a Monte Carlo) due to the impossibility of a human (or even a team of humans) to perform hundreds or even thousands of tests points in a feasible time.

For this reason, mathematical representations of the human operator are useful in a variety of scenarios. As stated by Turetta et al. [1], "science is very far from being able to represent all the complex human behavior with mathematical equations, but in some specific cases, it is possible to isolate a specific action that is required for an engineering test."

The first mathematical models of the human pilot were developed by McRuer [2] with linear and quasi-linear models. Those models kept evolving to optimal control models in the 1970s by Wierenga [3] and Kleinman [4], and then to nonlinear models such as Hess's proposal of 'pulsive control behaviour' [5]. Lone and Cooke [6] did an extensive review of those models.

Regarding the use of data driven techniques for modeling the human aircraft pilot, focus has been given on the landing task as it is recognizably the higher workload phase of flight (not considering military tasks). Martens [7] used neural networks to model and evaluate human behavior during landing in a light windshear condition. Mori, Suzuki and others [8, 9, 10] have modeled the pilot approach using stochastic switched linear models for the approach portion, and neural networks to model flare maneuver, which they stated to be highly nonlinear. The experiments performed in this paper are aligned with these observations.

The remainder of this paper is organized as follows. Section 2 contains the behavioral model development rationale. Section 3 presents the results and discussion using the identified models. Finally, the conclusions and further work are the contents of Sect. 4.

2 Model Development

2.1 Offset Landing Task

The offset landing task is indicated by Flight Test Guide for Certification of Transport Category Airplanes, FAA AC 25-7C [12], of the U.S. Department of Transportation, as one of the tasks to be performed during flight test to evaluate the likelihood of airplane pilot couplings, or pilot induced oscillations. This is a phenomenon that occurs when the phase between the pilot inputs and the airplanes dynamic response are in such a way that airplane oscillatory movement may occur causing discomfort or, at higher levels, even loss of aircraft's control.

The task consists in performing the approach to landing with a predetermined localizer offset, and at a given altitude, to correct the offset and land the aircraft. It is an operationally feasible situation that may occur if the airport localizer antenna is poorly regulated and the approach is performed under low visibility operations. In this case, the pilot performs the entire approach with an offset (as s/he is flying using instruments rather than visually) and at the time s/he acquires visual contact with the runway, the correction of the offset is required at a lower altitude than would be desired.

2.2 Input and Output Models Considerations

Some variables are fixed for this mathematical model, because they can adversely modify the aircraft dynamic response and increase considerably the complexity of the model. The landing gear is always down and locked and the flaps positions are the same in all test points. CG (center of gravity) position and aircraft weight that can also modify the aircraft dynamic response so it is desirable that they should be fixed or be considered as model inputs. Since a limited amount of test data was available, initially CG and weight variations were not considered, and this resulted in unsatisfactory models. Filtering the test points with only similar weights and CG positions, the model quality improved significantly. Finally, the test points were done with a constant thrust setting, and that no rudder input is used. This means that the outputs of the model are the control column lateral (that controls the ailerons, shown as y1 in the figures) and longitudinal (that controls the elevator, shown as y2 in the Fig. 1 displacements.

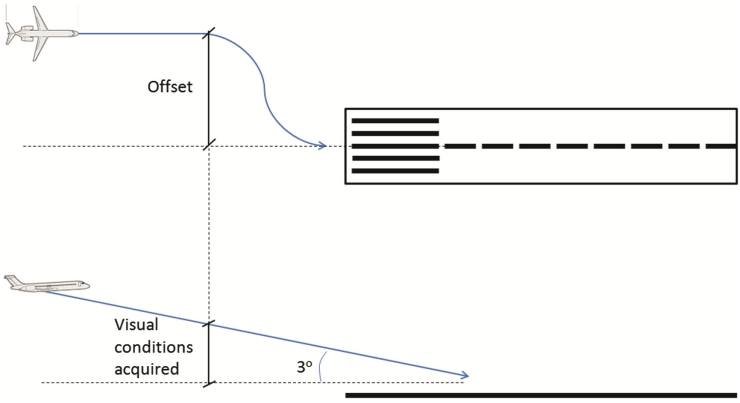


Fig. 1. Offset landing task.

The definition of the model inputs is a more complex task, as multiple factor may influence the pilot behavior in this situation. Mori and Suzuki's [8, 10] analyzed the visual cues the pilot used during approach, and make those cues the inputs of the model. In the end, those cues are a function of airplane position in relation to the runway, airplane speeds and the dimensions of the runway, cockpit seat and windows. Since this work is a data driven approach, it is possible to use directly the airplane position in the algorithm, as it can be transformed into the visual cues by geometric transformations. Therefore, the first inputs for the model were defined as the position of the aircraft in relation to the runway threshold (x, y and z).

Additionally, it is desired that the model is capable of responding to wind gusts, so the wind components in the three directions was also included as model inputs (Xwind, Ywind and Zwind).

Aircraft speed was also considered as an input because it influences the aircraft dynamic response and because the pilot knows that s/he cannot command high angles at low speeds.

The Euler angles were also considered as model inputs because they are highly considered by the pilot during maneuvering, and it is expected that they would reduce the control inputs gain at high angles to avoid stall (pitch, roll and yaw).

At last, it was considered that the load factors in each direction could also be considered as inputs, for a similar reason than the Euler angles, as the pilot would tend to reduce the control inputs gain at high load factors to avoid damage to the airframe (Nx, Ny and Nz). Later, it was observed that those inputs were not necessary as the load factors are not large during this specific operation.

The considerations above defined the model as having the following inputs and outputs:

Inputs:

- Set 1 (coordinates): X, Y and Z coordinates in relation to runway threshold
- Set 2 (winds): Xwind, Ywind and Zwind
- Set 3 (speed): Aircraft speed
- Set 4 (angles): Pitch, yaw and roll angles
- Set 5 (load factors): Nx, Ny and Nz.

Outputs:

- Control column lateral displacement
- Control column longitudinal displacement

The inputs are structured in sets because models with different subsets of those inputs are evaluated. The outputs are the same for all models in this paper.

3 Results and Discussion

3.1 Model Identification Techniques

Initial attempts to use linear mathematical models showed that the dynamic behavior was not captured. Switching to NARX (Nonlinear AutoRegressive with eXogenous input) models and Artificial Neural Networks improved significantly the results at the initial trials, but the models were still having low correlations and R2 coefficients. It was thus defined that artificial neural networks were seemingly more appropriate to fit this kind of problem.

Still, it was necessary to improve the model quality and too many factors were at hand to consider. An optimization approach using all variables was considered, but since it would dramatically increase the complexity of the whole procedure, a sequential strategy has been adopted, namely: analysis of the maneuver, re-evaluation of the inputs with a single landing, evaluation model order and complexity with a single landing and finally evaluation of multiple landings.

3.2 Analysis of the Maneuver

Due to initial difficulty of obtaining a good model, one question that arises was about the complexity of the maneuver and the possibility of modeling it entirely with a single model. As pointed by Mori and Suzuki [8, 10], the flare maneuver is highly nonlinear and represents a completely different behavior than the initial portion of the approach, and this made them divide the initial approach and the flare maneuver into two different models. Even for the initial part of the approach, they used three different linear ARX (AutoRegressive with eXogenous input) models that were switched by a Markov chain depending on the deviation from the intended glide path. The task under study herein is even more complex because it contains the offset correction. Initially, the desire was to model the complete maneuver, but the tryouts indicated that modeling even a normal landing with no flare was a complex task (regarding the neural network generalization capability, as it is shown below). Therefore, it was decided that the evaluation of the flare would be left for further studies, and focus would be given on the offset correction only. All data below 20 m of height above ground level was removed from the data sets to remove the flare behavior.

3.3 Reevaluation of the Inputs - Single Landing

Although all the inputs described in Sect. 2.2 influence the pilot inputs, it may be possible to model the behavior with fewer inputs to simplify the model and save computational cost. It is shown that for a single landing identification, a neural network with 10 neurons and fifth order of delays could deliver an almost perfect correlation and R2 with the actual data, indicating an over fitted model. So, to evaluate the contribution of the variables, models with a fixed order of 3 and complexity of 5 neurons (single layer and sigmoid activation functions) were generated for the same landing, using a different set of inputs. This was made to allow the observation of the variation in the model quality with the inputs sets variation. The results are seen in Figs. 2, 3, 4, 5 and Table 1.

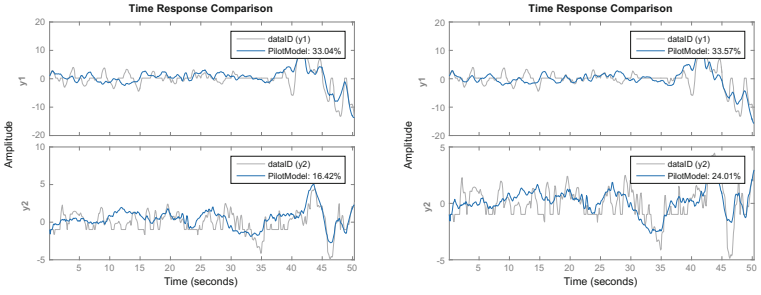


Fig. 2. Left: all inputs sets - Right: no winds set

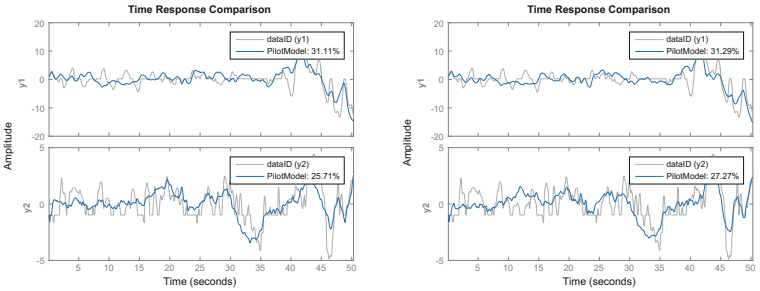


Fig. 3. Left: no angles set - Right: no speeds set

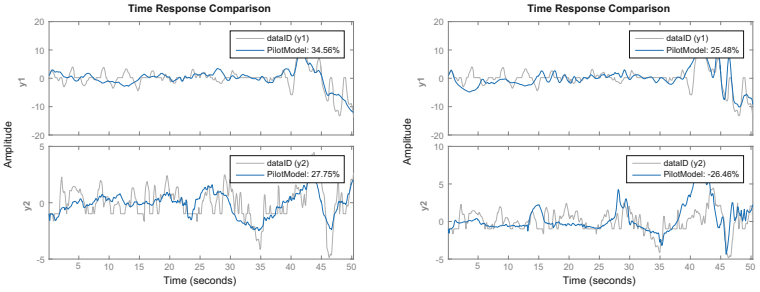


Fig. 4. Left: no load factors set - Right: coordinates only

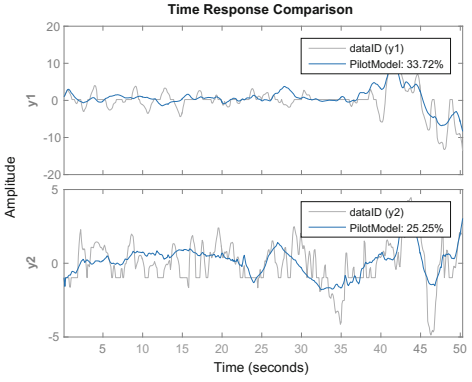


Fig. 5. Coordinates and angles only

Initial analysis suggested that it could be possible to use only the coordinates, but that model lost the capability of fitting y2 (longitudinal stick input). The reason for this result is traced to the pilots' need to know - at least - the angle of attack in order to calculate his maximum longitudinal stick input. It could be argued that adding only the pitch angle to the net would resolve this issue, but since the maneuver contains the offset correction that calls for large bank and roll angles, it is wise to include all the three angles as net inputs (as the angles are dynamically coupled). Table 1 shows that a model using

the coordinates and angles only showed results as good as the one with all variable sets. From this point on in this paper, all models are using those sets of inputs.

Table 1. Quality of models varying with inputs set.

Inputs	Correlation:y1	R2:y1	Correlation:y2	R2:y2
All inputs set	0.750	0.552	0.663	0.302
No winds	0.761	0.559	0.668	0.423
No angles	0.730	0.525	0.676	0.448
No speeds	0.733	0.528	0.695	0.471
No load factors	0.758	0.571	0.695	0.478
Coordinates only	0.703	0.445	0.416	0
Coordinates/angles	0.773	0.561	0.675	0.441

3.4 Model Order and Complexity - Single Landing

When identifying a model for a single landing and comparing the model to its data, it was observed that a perfect match could be obtained for any specific landing. Figure 6 shows that the metrics improve with higher orders and a greater number of neurons, with the median approaching 100% with ninth order or 9 neurons. It also shows that there is a significant spread between different landings, meaning that for a specific landing, a good model may be obtained with a simpler model. This is an interesting result and prompts the question of how it is possible to quantify the human variability. In other words: what is the variance of human behavior in that specific task? This metric, if developed, could be applied to help validate a human model, meaning that maybe a model does not have a particularly good fit to the test data, but it can still be considered as a "human like" behavior, due to this variance.

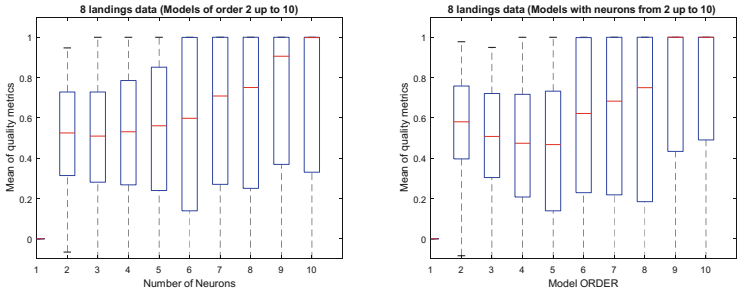


Fig. 6. Evolution of the mean of the model quality metrics (R2 and Correlations) with model order and number of neurons.

Another interesting result, can be seen in Fig. 7, notice that the fourth order model has a better fit than the sixth order one, the eight order model has an even worse fit, and finally the tenth order model has an over fitted response. However, notice that in the eighth order model, the fit is almost perfect in the beginning of the maneuver, diverging

at the middle. This indicates that there are two different behaviors contained in the data, and this model is able to fit the first behavior (which is the normal approach sequence). By the middle of the data, the second behavior appears and the model is no longer good (which is the offset correction). This shows that the strategy adopted by Mori and Suzuki in 2010 of having different models for different behaviors in the same maneuver is interesting because it allows using simpler models, rather than capturing different behaviors in a more complex model.

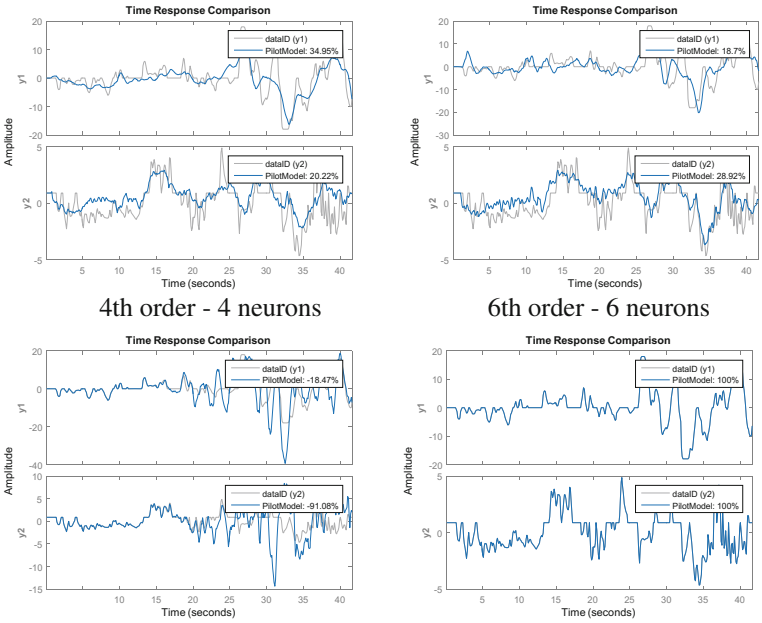


Fig. 7. Evolution of fit of a single landing with order and neurons number increase.

3.5 Evaluation of Multiple Landings

As seen above, when comparing the obtained models with the data sets used in the network training, it is possible to obtain good models with a considerably low complexity such as fourth order and four neurons (correlations around 70% and R2 around 45%). However, to be able to use the model in, for example, Monte Carlo simulations with large variations in the variables such as wind conditions offset size etcetera, the model should have considerable generalization capabilities. That is, the ability to reproduce human-like behavior in scenarios that did not appear in the training data set. This appeared to be a challenging problem, as even very complex models could not show good correlations when exposed to landing data that was not considered in the training set. Additionally, just increasing the model complexity started to prove an unfeasible option as the computational cost grows very quickly. As a measure of this cost, using an Intel(R) Core(TM) i7-4810MQ CPU @ 2.80 GHz with 16 GB RAM (Random Access

Memory), the learning process to obtain a single model from data merged from 7 landings with fourth order and four neurons took 10.22 s. The model from the same data but with tenth order and 10 neurons took 137.40 s.

In order to observe the generalization, the combination of 7 landings (out of 8) to be compared with the remaining one was made, with neurons and order varying from 2 to 15 (resulting in 1568 different models). This training process for all those models took 71 h.

Figure 8 shows that the results quality decrease significantly when compared to the data of a single landing (Fig. 6). This result indicates the need of a strategy change to obtain a model with good generalization capabilities. In addition, Fig. 8 shows that apparently there is no observable improvement in the quality of the model even when the order/quantity of neurons increases up to 15.

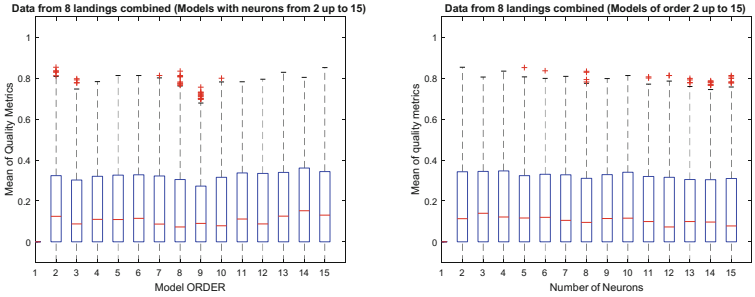


Fig. 8. Evolution of the mean of the model quality metrics (R2 and Correlations) with model order and number of neurons for models identified with 7 landings compared to an eighth one.

Changing the reference for model caused observable changes in the quality of the model. Figure 9 shows an example for the lateral stick input correlation only (but other metrics show a similar behavior). Notice that when landing 1 is used as a reference, the quality is always below 40%, while when the reference is landing 8, the quality is around 70%. Taking landing 2 as a reference, it is possible to see a relation between the increase in the order and the increase in model quality. This shows the complexity of the problem and that further analysis is required. Maybe some landings are outliers, or maybe they contain some characteristic that are not present in the other landings, causing the generalization to decrease when they are not used in the learning process. If that is the case simply increasing the number of landing available for training may solve the issue, but at this point, such data was not available.

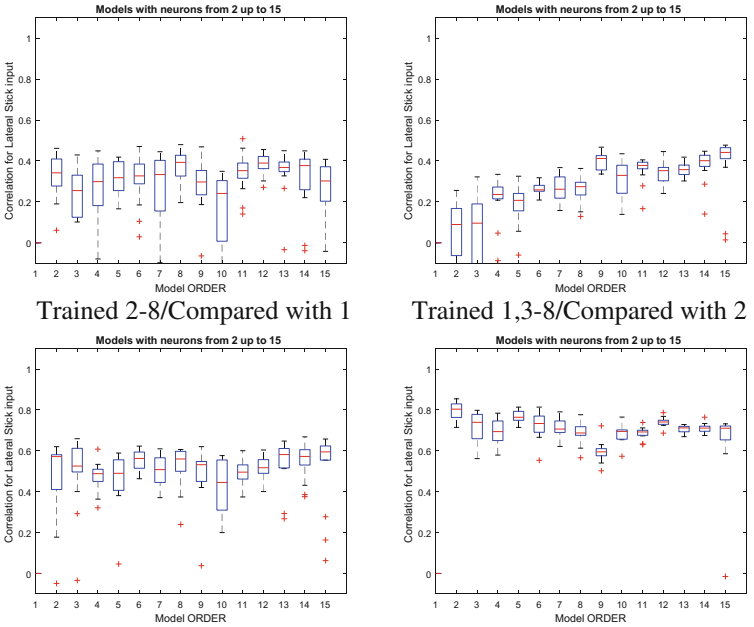


Fig. 9. Comparison of quality of the model (Correlation only) when changing the reference landing for comparison

4 Conclusions

In conclusion, it is possible to model a single landing using a neural network structure, but producing a network capable of reproducing different landings apparently requires further studies. Increasing the number of landings available for training is one obvious possibility but it has operational difficulties related to obtaining and filtering the data.

Other further studies include using an optimization algorithm considering all variables to discover the contributions of each variable and to try to obtain a model with better generalization capabilities. Another possibility is to design a specific experiment to make the identification, as this study was conducted with reused data from tests that were designed for other purpose. This would allow better signal/noise relations in some variables that could uncover interesting results. A possibility would be to start the modeling procedure with a simple normal landing with no offset, and then include growing offsets in the maneuver in such a way that is possible to obtain better understanding of how the offset modifies the pilot behavior.

Other future research is studying the variance of human behavior as suggested in item 3.4 would be a good addition to help validate human models.

Acknowledgements. This study is partially supported by the National Council of Scientific and Technologic Development of Brazil (CNPq) under Grant 150238/2017-7.

References

1. Turetta, F.M., Ayala, H.V.H., Trabasso, L.G., Alfredson, J.: Data-driven pilot behavior modeling applied to a VMCG determination flight test task. In: Proceedings of the Aerospace Technology Congress, Stockholm (2016)
2. McRuer, D.T., Jex, H.R.: Review of quasi-linear pilot models. IEEE Trans. Hum. Factors Electron. **HFE-8**(3), 231–249 (1967)
3. Wierenga, R.: An evaluation of a pilot model based on Kalman filtering and optimal control. IEEE Trans. Man Mach. Syst. **10**, 108–117 (1969)
4. Kleinman, D., Baron, S., Levison, W.: An optimal control model of human response Part I: theory and validation. Automatica **6**, 357–369 (1970)
5. Hess, R.: A rationale for human operator pulsive control behavior. J. Guid. Control **2**, 221–227 (1979)
6. Lone, M., Cooke, A.: Review of pilot models used in aircraft flight dynamics. Aerosp. Sci. Technol. **34**, 55–74 (2014)
7. Martens, D.: Neural networks as a tool for the assessment of human pilot behavior in wind shear. Aerosp. Sci. Technol. **3**(1), 39–48 (1999)
8. Mori, R., Suzuki, S.: Neural network modeling of lateral pilot landing control. AIAA J. Aircr **46**(5), 1721–1726 (2009)
9. Mori, R., Suzuki, S.: Modeling of pilot landing approach control using stochastic switched linear regression model. AIAA J. Aircr. **47**(5), 1554–1558 (2010)
10. Suzuki, S., Sakamoto, Y., Takahara, H.: Analysis of human-pilot control inputs using neural network. AIAA J. Aircr. **43**(3), 793–798 (2006)
11. Entzinger, J.O., Suzuki, S.: Modeling of the visual approach to landing using neural networks and fuzzy supervisory control. Aerosp. Sci. Technol. **14**, 118–125 (2010)
12. Federal Aviation Administration. Flight test guide for certification of transport category airplanes. AC No: 25-7C (2012)

Motion and Time Study on Space Maintenance Mission

Hanzhao Qiu[1,2], Weining Fang[1(✉)], Li Wang[3], Bo Wang[3], and Beiyuan Guo[1]

[1] State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University, No. 3 Shang Yuan Cun, Haidian District, Beijing 100044, China
{15116370,wnfang,byguo}@bjtu.edu.com
[2] School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, No. 3 Shang Yuan Cun, Haidian District, Beijing 100044, China
[3] National Key Laboratory of Human Factors Engineering, China Astronant Research and Training Center, Beijing 100094, China

Abstract. Whether the arrangement of the space maintenance mission is reasonable directly affects the work efficiency of the space mission and the work intensity of the astronauts. Aiming at the problem of lack of historical data, small sample size and non-repetitive task, based on the ground experimental data, the motion and time study on space maintenance mission is discussed. Motion study is processed base on motions and their objects. Forty one motion codes which are suitable for space maintenance mission are formulated. And these motions are classified. Based on time series, moment estimation method and Wilcoxon signed rank test respectively, three methods are proposed to determine observation time of motions and its accuracy. Taking a ground experiment as an example, the method can calculate all observation time of motions. And this can be the foundation for the calculation and forecast of the time of a space maintenance mission.

Keywords: Industrial engineering · Space maintenance mission · Motion and time study · Time series · The method of moments · Wilcoxon signed rank test

1 Introduction

The practice of the Mir space station and the International Space Station has proved that there will be a lot of failures in the formation and operation of the space station [1], while maintenance mission relies mainly on astronauts [2]. From the experience of the ISS, workload of space maintenance mission is high. In particular, maintenance mission out of the cabin occupies a large number of astronauts' extravehicular activity time [3].

However, there are few studies on space maintenance mission. Time of astronauts in orbit is precious. Whether the arrangement of the space maintenance mission is reasonable directly affects the work efficiency of the space mission and the work intensity of the astronauts. In order to solve the problem that the time of space maintenance mission is tight and workload is high, and it is difficult to arrange the time reasonably, this paper carries on motion and time study on space maintenance mission, and provides the foundation for the rational development of the space maintenance mission plan.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_13

2 Present Study of Motion and Time Study

Many researches have performed to forecast or set standard time through motion and time study. Utilizing the activity-bades costing concept and time study, Yuanyuan et al. [4] use historical cost data to estimate the time and cost of the new product. Christmansson et al. [5] combine the SAM method in motion study with the weight, force and workspace to form ErgoSAM. It can be used to predict the operating time and workload. And experiments show that the method can predict the physiological workload. Lauring et al. [6] evaluate the workload of the operational task by applying a workload evaluation to the motion in MTM or WF. Nakayama et al. [7, 8] put forward the work performance entropy and establish the trend equation which can be used to evaluate the performance and to calculate the standard time.

The above methods develop standard time which has applied to the actual production. The common characteristics of these methods are that they are all based on a large number of historical data and through a certain method to calculate the standard time of new products. However, space maintenance mission neither has historical data, nor can do a large number of experiments to obtain a lot of samples to develop standard time. Therefore, it is necessary to use other methods to establish the standard time which can be used to predict the time of space maintenance mission. This paper proposes three methods to determine observation time of motions for the establishment of standard time.

3 Motion Study on Space Maintenance Mission

Motion study is the foundation of time study. Because motion study's focus is on operation process improvements, it is necessary to refine the mission to operation or motion level which can thereby simplify the entire mission. And the effect of the motion study is ultimately assessed by time study.

The commonly used methods of motion study are shown in the Table 1.

Table 1. Comparison of general motion study methods

Name	Definition format	Advantage	Disadvantage
Therblig	Therbligs	Simple, easy to use	Lowly accurate
Work factor system (WF)	Therbligs + work factors	Moderately detailed and accurate	A little difficult to master and use
Method time measurement (MTM)	Therbligs + operating condition	Detailed and accurate	Hard to master and use
Modular arrangement of predetermined time standard (MOD)	Therbligs using limb as unit	Simple, easy to master and moderately accurate	Limited use range

As showed in Table 1, the Therblig's method is too simple to analyze the precise motion. And motion study in WF takes factors such as body parts, moving distance, resistance and human control into account. Besides considering moving distance, MTM considers the form and situation of movement and the weight change in the movement. While MOD omits the measurement of the moving distance, and gets the time of the therblig according to the parts of the body that used. In motion study, all therbligs are related to the body parts, while the work factors and operating condition are closely related to the properties of the manipulated objects. No matter which method is used, therbligs are the foundation and key. Therefore the analysis of the basic therbligs is required.

Assembly line operations of operators are repetitive motions in enterprise. Unlike the operational tasks in the enterprise, maintenance missions in the space station are pretty diversified. According to the experience of the International Space Station, in-cabin maintenance equipment includes general-purpose computer, pump, fan, storage battery, ventilation duct, sensor, panel, adsorbent bed, biofilter, water separator and so on [9]. Differences in objects make operation of maintenance missions different. However, therbligs in missions are the same. Through the analysis of five typical space maintenance missions, it is found that each mission contains the same therbligs.

It can be seen from four therblig definitions in Table 1 that only the therbligs themselves cannot divide motions well. There are usually some variable actions in operations whose existence will affect the accuracy of the time study later. Considering the space station has small space with uniform components, this paper combines four therblig definitions. The motion of space maintenance in this paper is defined as therblig combining with operating object, while the body parts that are used and the classification of the therbligs are also introduced. Some objects which have almost the same size or whose size do not affect the operation time can be classified as a class. Codes of therblig are shown in Table 2.

The Table 2 gives some of the actions involved in space maintenance missions on the ground and their corresponding codes and definitions. These actions are divided into three categories, namely the first category, the second category and the third category. The first category is effective actions which are the necessary action to carry out space maintenance missions. The second category is auxiliary actions. These actions are sometimes necessary actions, but their existence will delay the implementation of the first category, thereby reducing the efficiency. The third category is invalid actions which do not carry out any maintenance operations.

The purpose of motion study is to develop a reasonable maintenance process and to make a basis for time study. Therefore, it is necessary to analyze the causes of the second and third category, and eliminate the third category while reduce the second category as much as possible.

Table 2. Codes of therblig

Category	Therblig	Code	Definition	Operation objects	Body parts used
The first category	Orientation	DW	The act of orienting using eyes		Eyes
	Covering	G	The act of covering for a hole	1. Protecting plate	Big arm
				2. Toolkit	wrist
	Push	TU	The act of applying force in order to move something a little	1. Ratchet wrench	Two forearms
The second category	Distinction	BB	Distinguishing and finding desired object from similar objects using eyes		Eyes
	Search	X	The act of looking thoroughly in order to find something		Head, mainly eyes
	Arrangement	ZL	The act of making things as an orderly grouping condidered as a unit	1. Detector	Two forearms
The third category	The inevitable delay	BY	The operation is paused or delayed due to uncontrollable factors		
	The evitable delay	KY	The operation is paused or delayed due to improper operation and other subjective reasons.		
	Rest	XX	Stop working caused by fatigue		

4 Time Study on Space Maintenance Mission

4.1 Steps of Time Study on Maintenance Mission

Based on the motion study, time study reduces the invalid motions as far as possible, and develops standard time. Standard time can be used to assess whether the operator's operation has reached the standard level, and can be used to predict the required operating time before proceeding and to schedule the operation. Therefore, the schedule is based on the accuracy of the standard time which directly affects the possibility of the scheduled plan under normal circumstances.

The current space maintenance mission does not exist historical data that can be utilized to develop standard time. So indirect methods can not be used to develop standard time. At the same time, the short-term characteristic of the space maintenance mission makes the working sampling method not applicable. So only continuous observation method can be used.

The basic steps for time study on space maintenance mission are as follows.

STEP 1: Data collection
STEP 2: Task decomposition(action division)
STEP 3: Observation and record
STEP 4: Outliers elimination
STEP 5: Calculation of observation time
STEP 6: Determination of normal time
STEP 7: Determination of standard time

(1) Data collection

Motion study is the foundation of time study. Its focus is on the improvement of working methods. So it is necessary to refine the mission to the operation or action level first, and then simplify the mission. The effect of motion study is ultimately determined by time study.

(2) Task decomposition

The content of task decomposition is actually motion study. And the result of motion study is taken as the observation object of time study in the next step. The results are shown in Table 2.

(3) Observation and record

Due to the particularity of the space maintenance mission, it is impossible to carry out on-the-spot observation and time the mission. By using a head mounted camera, the action can be recorded at the operator's angle without affecting the operation. And it is convenient for timing later.

(4) Outliers elimination

After getting the operation video, operation time of every motion is measured. When using continuous observation method to analyze the video, it is needed to record the name of motion and its corresponding time. This paper uses Observer XT 10 to assist in timekeeping and statistics of motion time. Through the mark of the start and stop time of the motion, Observer XT 10 records and statistics each duration of each motion.

After the duration of each motion is obtained, the data need to be processed. Due to the possible external factors, the fluctuation of the motion time is beyond the interval of normal fluctuation, and these outliers need to be removed to ensure the accuracy of the observation time. Triple standard deviation method is used here to eliminate outliers.

In the triple standard deviation method, according to the principle of normal distribution, 99.7% of the normal motion time should be in the interval of the mean plus or minus triple standard deviation. And the excess is the abnormal value. After the outliers are removed, the normal interval needs to be recalculated, and then the other outliers

are removed. Take "loosen screw" as an example, in the case of the ground test six space maintenance missions, a total of 16 observations are gotten as follows. 1.277811, 1.277811, 1.66671, 0.888912, 1.444482, 1.166697, 5.944599, 1.555596, 0.277785, 2.22228, 1.833381, 1.388925, 3.33342, 0.944469, 2.388951, 0.500013

The mean and standard deviation of the above observations are as follows.

$$\mu = 1.75699,\ \sigma = 1.339508 \tag{1}$$

According to triple standard deviation method, the fluctuation interval of the normal data can be obtained as $(\mu - 3\sigma,\ \mu + 3\sigma) = (-2.26, 5.78)$. It can be determined that the observed value 5.944599 is abnormal value and should be removed.

After the removal, the normal interval is recalculated as $(-0.82, 3.78)$. All observations are in this interval, so that the remaining data are normal values which are in the normal interval of fluctuations. Observed values of other motions are processed like the above step to ensure high accuracy of standard time.

(5) Calculation of observation time

After removing the outliers, the observation time can be calculated from the qualified data. And the mean of them is usually calculated as the actual operating time. However, the sample of the space maintenance mission is less, while the environment is complex and the interference factors are more than on the ground. The accuracy of the direct use of the mean can not be assessed. Therefore, three methods are proposed to calculate the observation time and to determine its reliability in the next section

(6) Determination of normal time

Through the above steps, it is possible to determine the observation time of each motion of an operator. Through the evaluation work, it is modified to be a normal time for each motion. What time study needs is the time it takes for a skilled operator to operate at a normal operating speed. But the actual state of the operator is not steady and the speed of operation is not necessarily the normal speed, so the observation time needs to be evaluated. In this paper, the speed evaluation method, which is based on the standard operating speed of the qualified operator, is used to judge the actual operating speed of the operator. The evaluation coefficient in the form of percentage is used to adjust the normal time value. The computation formula is as follows.

$$\text{Normal time} = \text{observation time} * \text{evaluation coefficient} \tag{1}$$

(7) Determination of standard time

Standard time consists of two parts, normal time and allowance time. Allowance time refers to the time delay due to various reasons, mainly including the allowance time of operation, physiology, fatigue and other allowance time. The allowance time is related to the operator's personal characteristics, the nature of the work and the environmental factors. However, the relationship between them is so complex that ILO has not adopted the standards related to the determination of allowance time. Because of the lack of research on the fatigue of space maintenance, this paper takes the standard of medium

fatigue allowance rate as standard. And its value is 10%. The following formula determines the standard time.

$$\text{Standard time} = \text{normal time} * (1 + \text{allowance rate}) \tag{2}$$

4.2 Calculation Method of Observation Time

Since the observation time is the basis for the establishment of the standard time, the next step is to put the observation time to the standard time. A small deviation of the observation time will lead to a large deviation of the standard time, which in turn affects the reliability of the standard time. Therefore, the accuracy of observation time must be guaranteed. Using arithmetic mean as the observation time neither reflects the confidence of the estimated value when the sample is sufficient, nor eliminates random error as much as possible and ensures the accuracy of the observation time when the sample is small. Hence, this method needs to be improved.

Based on the time series method, the moment estimation method and the Wilcoxon sign rank test, three different methods of determining the observation time are proposed

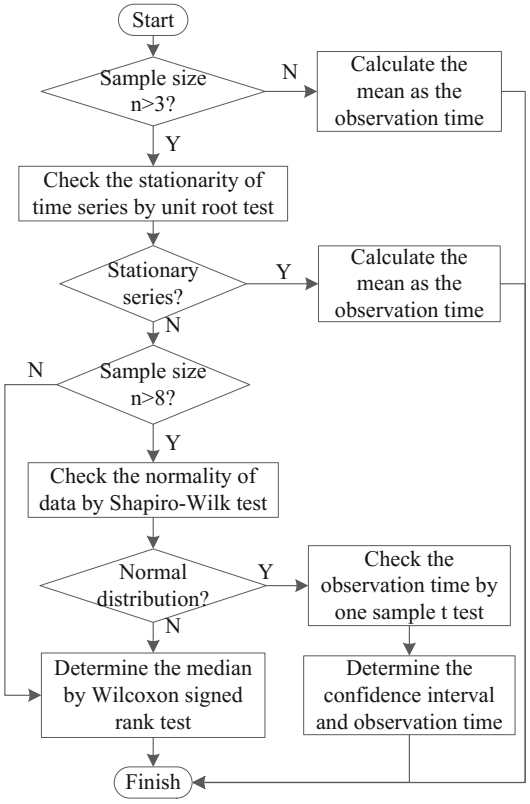


Fig. 1. Flow chart of the determination of observation time

in this paper. The application of the three methods are different. Flow of determining observation time is shown in Fig. 1.

Observational Time Based on Time Series Method. Time series is a concept in econometrics, and can be defined as follows. If a time series process satisfies the following properties: mean and variance are constant and independent of time, covariance is only related to the interval between two periods and has nothing to do with the period itself, then it is a covariance stationary processes [10].

From the above definition, it can be seen that for a stationary time series, the mean, variance and covariance of its future sample points are still equivalent to the corresponding values of the existing samples. Thus, it can be proved that the mean of the existing samples can express the mean of the population, that is, the observation time of the motion, by proving the stationarity of the existing sample time series.

The stationary test of time series mainly uses augmented dickey-fuller test (ADF test) in the unit root test. If there is a unit root in the time series, the sequence is a non-stationary time series. Thus it is possible to check whether the motion observation time series has a unit root to determine whether the sequence is stationary.

For motion observation time series x_t, use ADF test to verify the existence of unit roots. The ADF test is done through three models.

$$x_t = \delta x_{t-1} + \sum_{i=1}^{m} \beta_i \Delta x_{t-i} + \varepsilon_t \tag{3}$$

$$x_t = \alpha + \delta x_{t-1} + \sum_{i=1}^{m} \beta_i \Delta x_{t-i} + \varepsilon_t \tag{4}$$

$$x_t = \alpha + \beta t + \delta x_{t-1} + \sum_{i=1}^{m} \beta_i \Delta x_{t-i} + \varepsilon_t \tag{5}$$

Where $\sum_{i=1}^{m} \beta_i \Delta x_{t-i}$ is lagging, ε_t is the white noise, α is intercept items, and β_t is the trend.

The test hypothesis is

$$H_0{:}\delta = 1 \qquad \text{(Unit root exists, the sequence is not stationary)} \tag{6}$$

$$H_1{:}\delta < 1 \qquad \text{(Unit root does not exist, the sequence is stationary)} \tag{7}$$

The t test is performed under the least square method on the above model in the order of model (5), model (4) and model (3). And then using the ADF distribution threshold table to confirm whether to reject the null hypothesis. When the test rejects the null hypothesis, that is, the sequence is stationary, the next test can be stopped, otherwise the test is continued until the model (3) is tested.

In this paper, R is used to conduct the ADF test for each motion. The test results which satisfy the requirement of stationary under $\alpha = 0.05$ are as follows. The operation object is represented by the serial number, as showed in Table 2.

Observational Time Based on Moment Estimation Method. Through the stationary test of time series, it can be concluded that the mean value of the sample is the mean of subsequent sample points. And the mean value can be used as the observation time to calculate the standard time. However, in practice, the stationary time series is difficult to obtain, and a tiny disturbance will lead to instability of sequence. Therefore, it is not possible to require all sequences to meet the requirements of stability. In this paper, we propose an observation time determination method based on the moment estimation method to determine the observation time and evaluate its confidence (Table 3).

Table 3. ADF test results (motions satisfies stationarity)

Name	Operation objects	Dickey-Fuller	P value
Orientation		−4.9292	<0.01
Covering	2	−2.4149	0.01707
Distinction		−1.9801	0.04748

The moment estimation method is the method of using the sample moments to estimate the corresponding parameters in the population. The simplest moment estimation method uses the first order sample origin to estimate the population expectation and uses the second order sample center moment to estimate the population variance.

For the subdivided motions, there are no factors in theory that affect their time, and thus each motion time is subject to a normal distribution $x \sim N(\mu, \sigma^2)$. In order to determine the value of μ, normality test of the sample is firstly needed. The normality test is performed using the Shapiro-Wilk test [11].

In this paper, Shapiro-Wilk test is performed on each motion time using R. The motion time which can not be rejected as normal distribution under $\alpha = 0.05$ is as follows (Table 4).

Table 4. Shapiro-Wilk test results (motions satisfied normality)

Name	Operation objects	W	P value
Covering	1	0.90517	0.185
Push	1	0.95305	0.5737
Search		0.92915	0.06586

The observation time of the above motions satisfies the normal distribution. Time of each motion is averaged and rounded to retain a decimal, and it is treated as an expectation. A single sample t test under $\alpha = 0.05$ is performed for the observation time. The results of motions that cannot reject the expected value as the global mean are as follows. (Table 5).

Table 5. T test results

Name	Operation objects	t	df	P value	Confidence interval
Covering	1	0.03522	11	0.9725	(1.68,2.73)
Push	1	−0.112	14	0.9123	(1.05,1.90)
Search		−0.0406	26	0.9679	(3.95,6.60)

Observational Time Based on Wilcoxon Signed Rank Test. The accuracy of the observation time determined by the above two methods is high. And they are feasible for most of the data. However, there are a few motions with a few samples that cannot meet the sequence stationarity and normality. At the same time, due to the particularity of space maintenance mission, there is no chance to repeat the test for specific motions. Generally the nonparametric test is used to determine the time of motions whose distribution is unknown. This paper uses the Wilcoxon sign test to estimate the median of the population, so as to determine the observation time of the residual motion.

The Wilcoxon sign test assumes that median of the population M is not equal to a certain number M_0, and then does the test.

$$H_0{:}M = M_0 \tag{8}$$

$$H_1{:}M \neq M_0 \tag{9}$$

In this paper, the Wilcoxon signed rank test is performed using R for the remaining time of motions. And the result of motions which cannot reject the null hypothesis under $\alpha = 0.05$ are as follows (Table 6).

Table 6. Wilcoxon signed rank test results

Name	Operation objects	P value
Arrangement	1	0.4697
The inevitable delay		0.5966
The evitable delay		0.4358
Rest		1

Based on the above three methods, the observation time of motions of space maintenance mission is determined as follows (Table 7).

Table 7. Motions observation time results

Name	Operation objects	P value
Orientation		1.76
Covering	1	2.2
	2	1.91
Push	1	1.5
Distinction	1	4.09
Search		5.3
Arrangement	1	6.6
The inevitable delay		11.3
The evitable delay		11.1
Rest		68.7

5 Discussion and Conclusion

Space maintenance mission is critical to the safety and normal operation of the on-orbit space station. This paper carries out the motion and time study on space maintenance mission. Firstly, space maintenance mission is divided into 41 motions according to its' characteristics. Then, the operation time data of each motion is obtained by measuring the ground experiment of the space maintenance mission. Finally, based on the existing data, this paper proposes three methods based on time series method, moment estimation method and Wilcoxon signed rank test to determine the observation time for different data types. This paper provides the basis for the workflow optimization, workload analysis and schedule establishment of space maintenance mission.

Acknowledgments. This paper is supported by Beijing Municipal Science & Technology Commission (Grant No. D161100000116002) and the foundation of National Key Laboratory of Human Factors Engineering (Grant No. SYFD18061610).

References

1. Teng, X., Chen, Q.: Research on strategies of on-orbit maintenance in foreign space stations and its enlightenments. Space Med. Med. Eng. **25**(6), 475–478 (2012). (in Chinese)
2. Zhu, Y.: Lessons learned from ten-year construction of international space station. Spacecr Eng **19**(1), 50–59 (2010). (in Chinese)
3. Fan, W.: Study on system characteristics of international space station. Spacecr. Eng. **21**(2), 94–100 (2012). (in Chinese)
4. Jiao, J., Tseng, M.M.: A pragmatic approach to product costing based on standard time estimation. J. Syst. Sci. Syst. Eng. **19**(2), 738–755 (2001). (in Chinese)
5. Christmansson, M., Falck, A.C., Amprazis, J., et al.: Modified method time measurements for ergonomic planning of production systems in the manufacturing industry. Int. J. Prod. Res. **38**(17), 4051–4059 (2000)
6. Laurig, W., Kühn, F.M., Schoo, K.C.: An approach to assessing motor workload in assembly tasks by the use of predetermined-motion-time systems. Appl. Ergon. **16**(2), 119–125 (1985)
7. Nakayama, S.I., Sakai, R.: An evaluation of work performance using a work achievement quotient. Int. J. Prod. Res. **23**(5), 879–885 (1985)
8. Nakayama, S.I., Nakayama, K.I., Nakayama, H.: A study on setting standard time using work achievement quotient. J. Int. J. Prod. Res. **40**(15), 3945–3953 (2002)
9. Zhang, W., Xia, Q.: Study on design and verification method of space station maintainability. J. Manned Spacefl. **20**(2), 134–138 (2014). (in Chinese)
10. Baltagi, B.H.: Econometrics. Springer, Berlin (2008)
11. GB/T 4882-2001: Statistical interpretation of data-Normality tests (2001)

Air Traffic Controller Additional Work Load as a Result of Aircraft Dynamic Separations

Julio Roa(✉)

Air Transportation Systems Lab, Civil and Environmental Engineering Department, Virginia Tech, Blacksburg, VA 24061, USA
julioroa@vt.edu

Abstract. For some years, the International Civil Aviation Organization (ICAO), Federal Aviation Administration (FAA), Eurocontrol and Japan Aerospace Exploration Agency (JAXA) have proposed several initiatives as part of the effort to reduce wake mitigation separation that could increase airport capacity. Among different proposed wake mitigation concepts, dynamic wake separation has been identified as a promising method to increase airport capacity in the near-term. Dynamic wake separations take into account the environmental and aircraft dependent parameters that influence wake behavior. The intent of this paper is to study and discuss the additional workload and complexity to air traffic controllers as a result of dynamic separation standards associated with aircraft recategorization (RECAT Phase 3). There is a potential capacity benefit that could result from taking into consideration dynamic parameters; at the same time, there is an additional workload being imposed on air traffic controllers due to the complexity of non-static separations.

Keywords: Wake vortex · Dynamic wake separations · Air traffic controller · Monte carlo approach · Simulation · Sustainability

1 Introduction

In the beginning of the 1970's with the introduction of heavier passenger planes like the Boeing 747 and the Donald Douglas 10, wake hazard potential of these airplanes to smaller airplanes was studied by NASA and Federal Aviation Administration (FAA) as part of NASA studies on aircraft trailing wake vortices. A series of flights combinations and aircraft configuration with NASA's Boeing 727, Boeing 747, Cessna T-37 and Learjet were conducted in order to record data and gain knowledge on wake vortex [1].

Wake turbulence has been shown to be dangerous to aircraft that encounter it as they can exhibit cross flow velocities 360 km/h or higher in their core region depending upon flight conditions and airplane parameters [2].

An aircraft encountering wake vortices experience considerable roll and altitude disturbance. To avoid this danger, aircraft separations were developed in the 1970s, based on the technology available at that time, that allow time for wake to be dissipated and displaced from an aircraft's path (Fig. 1).

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_14

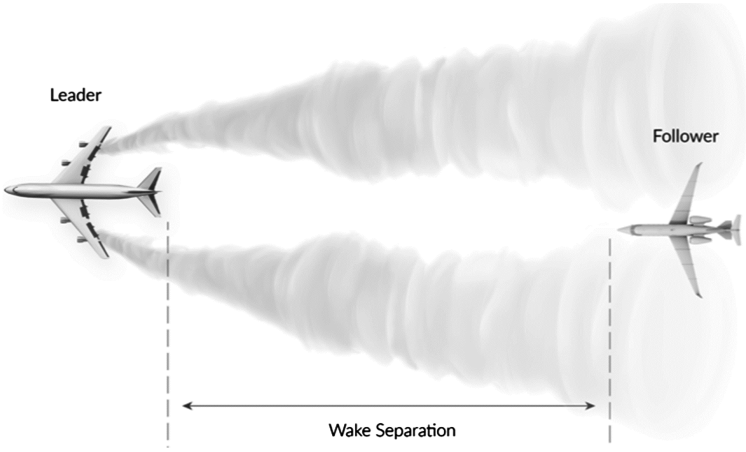


Fig. 1. Representation of wake separation.

To avoid hazardous wake encounters, the FAA and ICAO grouped similar aircraft in terms of size and weight carrying capabilities; and developed wake vortex separation static matrices between leading aircraft groups and follower aircraft groups. Due to the limited knowledge of wake vortex safe separations between aircraft might be overly conservative [3].

2 Achieving Dynamic Wake Separations

As knowledge about wake vortex has increased in a period of over 30 years, aircraft categorizations of wake vortex have been created and modified. Wake vortex separation standards have evolved from five legacy groups to six groups under Recategorization I (RECAT I).

Current wake vortex separations degrade aviation efficiency when traffic congestion limits airport capacity during landing and takeoff [4].

As more aircraft types are introduced into the NAS, the wake separation standards between aircraft need to be revised to ensure runway throughput and airport capacity is maximized without compromising safety. Without any improvement in the separation distance used today, other Next Generation Air Transportation System (NextGen) technologies such as Automatic dependent surveillance – broadcast (ADS–B), will have a less significant role in improving the operational capacity.

The Federal Aviation Administration (FAA) is leading efforts to increase the capacity of the National Airspace System (NAS). The main constraint on airport throughput is the runway capacity.

Currently, runway throughput performance is dependent upon wake turbulence separation or minimum radar separation. Minimum radar separation is driven by communication/surveillance and collision avoidance between a leading and following aircraft. Wake turbulence separation is driven by the wake circulation strength produced by a leading aircraft and the wake circulation strength capacity of the following aircraft.

A better understanding of wake vortex behavior and its interaction with environmental parameters, aircraft-dependent parameters, and aircraft limitations can increase utilization of airport assets while reducing airport environmental footprints and cost of aircraft operations (see Fig. 2).

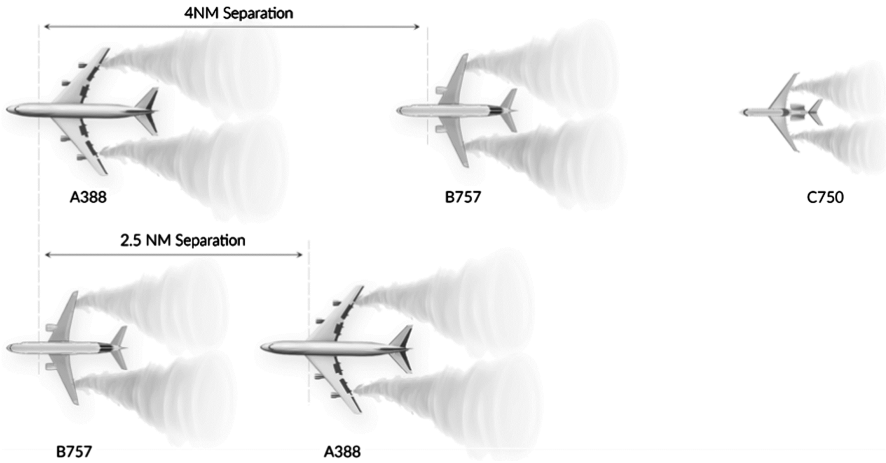


Fig. 2. Static aircraft separation.

3 Air Traffic Controller Workload

Air traffic controllers are tasked with ensuring the safe and efficiency of aircraft from ground operations to approach and landing operations at the destination (see Fig. 3).

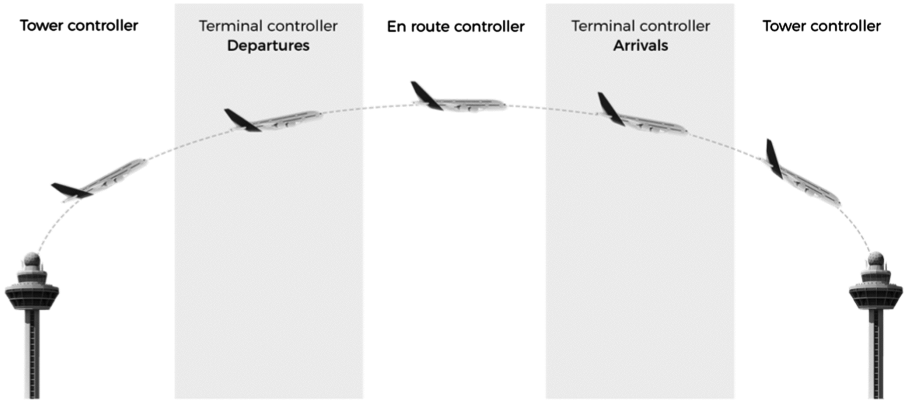


Fig. 3. Air traffic controllers at various stages of flight.

Safety of the national airspace system is ensured, by guaranteeing established minimum separation between aircraft, a separation defined vertically and horizontally

around each aircraft. Currently, the minimum separations on approach are 2.5 nautical miles but these separations have different values depending on the type of operation and the region that is being flown. Table 1 presents recategorization I separation matrix for arrivals.

Table 1. Wake turbulence separation matrix for arrival - Recat I.

Lead/follow	A	B	C	D	E	F
A	2.5	5	6	7	7	8
B	2.5	3	4	5	5	7
C	2.5	2.5	2.5	3.5	3.5	6
D	2.5	2.5	2.5	2.5	2.5	5
E	2.5	2.5	2.5	2.5	2.5	4
F	2.5	2.5	2.5	2.5	2.5	2.5

Air traffic controllers are responsible for managing aircraft operations throughout several phases of flight including ground operations from the gate to the taxiway to the runway takeoff and climb operations to reach a cruising altitude, cross-country flight to the destination, approach and landing operations at the destination, and taxi to gate.

As advances in aircraft surveillance technologies and communication systems such as satellite-based navigation is fully implemented. The capability of monitoring aircraft under a higher frequency than current surveillance systems is achieved. It is clear that the limiting parameter to increase runway throughput is wake turbulence separation.

When aircraft approach a runway, the final segment between the final approach fix and the runway threshold varies as a function of the common approach length. Technologies such as Airport Surface Detection System Model X (ASDE-X) allows controllers to track movement at and around the airport. ASDE-X combines the data of surface radar, multi-lateration sensors, airport surveillance radar, automatic dependent surveillance, and the terminal automation system to determine the position of and identify aircraft on final approach as well as on the airport surface [5].

Environmental factors can exacerbate or attenuate the operating conditions of aircraft in flight. Some of those factors are environmental turbulence, air density, temperature, Brunt-Väisälä frequency, and crosswind speed over the entire space of flight tracks.

The national airspace system is maximized by the optimizing the capacity of the current airspace, the limiting factor in this process is the rate of arrivals at an airport, particularly peak hours at the large hubs or after weather interruption such as winter storms.

Every airport has a capacity in terms of the number of aircraft it can receive per unit of time. The goal of the air traffic control system is to meet that capacity by optimizing the flow and diminishing average waiting times. Currently, this can be accomplished by delivering arrivals at regularly spaced intervals and sequencing departures when the gap inter-arrivals allow it.

Recategorization phase III not only poses the challenge for a larger aircraft separation matrix but also for a future dynamic separation matrix based on current airport and aircraft conditions.

The complexity of pair-wise dynamic wake separation not only requires advanced decision support tools that can help controllers coordinate and direct arriving and departing flights but also the increase in complexity taxes the controller's processing capabilities and may constrain the resolution with which the oncoming airport traffic can be visualized (Fig. 4).

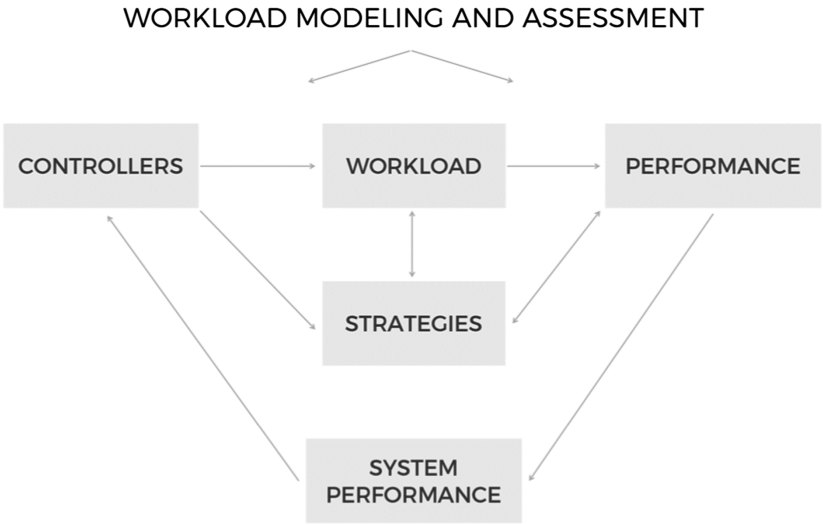


Fig. 4. Air traffic controller interrelationships between workload and performance.

4 Conclusions

Due to the lack of variability and small size of current wake vortex separation matrix, the nature of most air traffic control decisions is routine enabling controllers to select appropriate procedures to apply once the situation is identified. Research suggest that experienced decision makers are able to identify patterns associated with problems rather than comparing options in terms of their predicted outcomes [6].

In the traffic controller environment workload and stress make the working memory more vulnerable to information loss and when the operating procedures are inconsistent situational awareness is more vulnerable. If recategorization III, where to be a dynamic wake separation matrix based on live aircraft and environmental parameters; future automation alternatives with the ability to manage traffic separation as well as predict conflicts and optimize solution when conflict appears need to be considered. The interaction and dependency between the air traffic controller and the automation tools need to be studied and fall back plan needs to be put in place to account for toll failure or disruption of the system.

The forecasted increase in air transportation operations could overwhelm the capacity of the airport system, especially at main hubs. In order to take measures before this happens separation between aircraft could be safely reduced if live dynamic

parameters that have proven to influence wake vortex behavior and decay are taken into consideration. The impact of dynamic separations on air traffic control workload and complexity needs to be studied and alternate solutions proposed to mitigate risk and maximize capacity gains.

Acknowledgements. The contents of this material reflect the views of the author only. Neither the Federal Aviation Administration nor the United States Department of Transportation nor National Aeronautics and Space Administration nor Virginia Tech makes any warranty or guarantee, or promise, expressed or implied, concerning the content or accuracy of the views expressed herein.

References

1. Dan, V., et al.: Characterizing the hazard of a wake vortex encounter. NASA Langley, Hampton (1998)
2. Breitsamter, C.: Wake Vortex Characteristics of Transport Aircraft. Institute of Aerodynamics, Technische Universität München, München (2010)
3. Naoki, M., et al.: Airport terminal traffic simulation applying reduced wake vortex separation. In: 10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference. American Institute of Aeronautics and Astronautics (2010)
4. Holzäpfel, F., Kladetzke, J.: Assessment of Wake-Vortex encounter probabilities for crosswind departure scenarios. J. Aircr. **48**(3), 812–822 (2011)
5. FAA Airport Surface Detection Equipment, Model X (ASDE-X). Air Traffic Technology (2014)
6. Wickens, C.D., et al.: Flight to the Future: Human Factors in Air Traffic Control. National Academies Press, Washington (1997)

Habitability Issues in Long Duration Space Missions Far from Earth

Giorgio Musso[1(✉)], Simona Ferraris[1], Franco Fenoglio[1], Antonio Zafarana[2], Adriana Salatino[2], and Raffaella Ricci[2]

[1] Thales Alenia Space Italia, strada Antica di Collegno 253, 10146 Turin, Italy
giorgio.musso@thalesaleniaspace.com
[2] Department of Psychology, University of Torino, Turin, Italy

Abstract. Living in space today means to stay at an altitude of about 400 km above the Earth surface, on the orbital International Space Station (ISS). In the '70s the man reached the Moon but a manned space exploration mission, beyond low Earth orbit and cislunar space, might significantly increase adverse psychophysical effects on human wellbeing. Nowadays, a manned mission to deeper space, such as for example to Mars, is one of the greatest psychological challenge that has never been faced by the humankind. Due to the enormous distance between Earth and Mars, astronauts sent to Mars will be the first human beings who will lose a direct visual link with their Home Planet. Human responses to this and other extreme conditions that might be encountered during long duration missions into deep space are still unknown. In addition, the acute and long term effects of altered gravitational input on the central nervous system and their impact on sensorimotor and cognitive functions need to be clarified to assure maximum performance capabilities during spaceflight and planetary explorations. Our current knowledge on psychological and cognitive effects of orbital spaceflights or analogue environments is not sufficient to reliably assess the specific risks of human mission into outer space. New psychological challenges of mission to Mars will be analyzed with respect to three different areas: individual response and small crew interactions in isolated, confined, and extreme environments (ICE); human adaptation and performance in different gravity environments; concept and methods of psychological countermeasures. The needs of crew members to effectively and safely live and work in space are now referred to missions orbiting around the Earth and have been managed through specific human factors requirements applicable to the ISS. Future manned exploration missions need to reinforce these requirements to design an environment suitable for a safe stay during manned space missions far from Earth. The recommendations of astronauts who have experienced long term stays in space are collected and analyzed to be translated into requirements to be implemented in future space habitats. The analysis of what we have now and what is thought to be relevant to ensure crew wellbeing and performance during long term stays in space is a critical step to assure the success of deep space human missions.

Keywords: Space · Astronaut · Space habitats · Long duration space mission · Psychology

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_15

1 Introduction

Science fiction books told us about travels to other galaxies with velocities larger than velocity of light, meetings with friendly aliens, star wars with ugly space monsters, etc. The beginning of the Space age (late fifties) opened the door to a magic box, full of scientific discoveries, made mostly by robotic satellites and spacecrafts. However, early Space trips clearly demonstrated that Space environment is extremely hostile to human beings. Man is not made for living in Space. Even if Space medicine, during the years since Gagarin flight, made an outstanding progress in supporting human presence at orbital stations, the radiation hazards and hypo magnetism problem are still opened issues without visible paths to their solution.

Also current research on human psychological and sociological effects of life in Space is based on on-orbit near-Earth experiences and therefore may have limited generalizability to long distance and long duration Space expeditions, such as a mission to Mars or to a near-Earth asteroid (which is currently considered by NASA).

In the case of Mars (or, more generically, in case of long duration spaceflights in deeper Space far from Earth), new stressors will be introduced due to the great distance involved in journeying to the Red Planet. For example, the crew members will be relatively autonomous from terrestrial mission control and will need to plan their work and deal with problems on their own. They are expected to experience significant isolation as the Earth becomes an insignificant dot in the heavens, the so-called Earth-out-of-view phenomenon, the effects of which are still largely unknown. Again, on the surface of Mars (as an example), the round trip communication delay time, depending on the relative position of Mars and Earth, ranges from 6,5 to 44 min approximately, increasing the sense of isolation [1].

The psychological effects of factors that are proper of autonomous long duration missions, such as Earth-out-of-view phenomenon, difficulties of communication, awareness of the impossibility to quickly return to Earth, confined habitat, isolation, etc. need to be analyzed and addressed with proper countermeasures.

Behavioral issues associated with isolation and confinement are currently studied by researchers on ISS or using terrestrial analogues. These studies collected extended evidence on emotional and psychological states and create a rank-ordering of behavioral observations addressing suggestions to help designing equipment, procedures, and activities to maintain mental wellbeing in future missions.

In a NASA experiment (HI-SEAS, Hawaii Space Exploration Analog and Simulation) [2], people living in isolation for a year to prepare for missions to Mars said “We were always in the same place, always with the same people”. These kind of experiments are aimed to plan the methods and means of control and monitoring of the conditions of habitat and crew during long-term crew stays in confined and harsh environments.

The International Space Station (ISS), that now provides on-orbit research and technology development activities and, as an engineering test bed for flight systems, operations and technologies, is critical to future long duration missions.

The former approach adopted for maintenance of the ISS was based on a robust logistics system able to return large sub-assemblies for repair to ground. This choice was made early in the ISS design when the US Shuttle was the main cargo transportation

system. It is now clear that after the retirement of the US Shuttle, the available download capability will be only a fraction of what is today. In addition, exploration missions outside Low Earth Orbit (LEO) will have very limited capability to return hardware to ground for repair and, in some cases, no possibility at all. This means that future interplanetary missions need to be self-reliant with respect to troubleshooting and repair and the ISS can be used now as a test bed to investigate and develop everything is needed for preparing future long duration missions far from Earth, including methods for in situ maintenance and repair.

The current Logistics scenario for the ISS is heavily constraining in terms of logistics, since it is based on "spare" concept: it foresees as baseline the ORU's (Orbital Replaceable Units) to be replaced on orbit using spares, but in case of failure it can be shipped back to ground either for repair or to be discarded.

In contrast, the development of interplanetary missions cannot rely on a maintenance scenario based on a spare policy. The interplanetary missions need to be self-reliant with respect to troubleshooting and repair, and the ISS can be used as a test bed to investigate and develop methods for in situ maintenance and repair.

Man is part of the system that must be made reliable to guarantee the success of the mission. Past and current manned missions/station (as the ISS) on LEO have relied on vehicles with rapid return capability for crew rescue. Either the spacecraft itself has this capability (e.g. Mercury/Gemini compared to Apollo) or a dedicated return vehicle (e.g. ISS, use of Soyuz). If return from LEO takes hours, return from the Moon takes days and return from Mars takes months. Therefore future long distance/duration missions will not have the capability to perform immediate crew return. The solutions implemented for in situ maintenance and repair become applicable to the man, since in this case it has to be considered a "fallible component".

Man is a critical system component that must be reliable. A "failure" of the "man" (intended as a component of the overall system) can compromise the success of the mission. Therefore, during long duration missions the crew need to undergo a strict program of preventive maintenance (physical, psychological, …).

Currently, the longest human spaceflight is that of Valeri Polyakov, who spent 437 days on orbit, in Space. All the longest spaceflights (see Fig. 1) were performed in orbit around the Earth (LEO).

CREW	COUNTRY	YEARS	STATION	NUMBER OF CONSECUTIVE DAYS IN SPACE
Valery Polyakov	Russia	1994-1995	MIR	437
Sergei Avdeyev	Russia	1998-1999	MIR	379
Vladimir Titov Musa Manarov	Soviet Union	1987-1988	MIR	365
Mikhail Koriyenko Scott Kelly	Russia United States	2015-2016	ISS	340
Yuri Romanenko	Soviet Union	1987	MIR	326
Sergei Krikelev	Soviet Union Russia	1991-1992	MIR	311

Fig. 1. The top longest space flights in history [3].

"*What are the most important habitability issues to be considered for long duration spaceflights?*". This question has been asked to Clayton Anderson, NASA astronaut who spent 152 day aboard the ISS.

In his answer he mentioned three different habitability elements:

- Window
- Private room
- Large room

Again, the statement "Windows are an essential part of vehicles and habitats for short and long-duration missions to optimize task performance and crew well-being" has been evaluated with the maximum score by another astronaut with experience of long term stay on orbit (6 months).

All these aspects represent the key factors to be taken into account to drive the design of the next habitats for the future manned Space exploration missions.

The challenge of Space habitat design is to create a living and working environment suitable for being recognised by the astronauts as "home"(the astronauts home in orbit), limiting the effects of the stressors as much as possible. It has to be comfortable, functional and safe enough to make astronauts feel well and thus perform effectively.

2 Psychological Highlights

2.1 Psychological and Cognitive Issues

A manned mission to deep Space, beyond LEO and cislunar Space, such as for example to Mars, is one of the greatest psychological challenge that humans have ever faced. Several stressors might condition the astronauts' mental state and their impact is expected to augment in human long-duration missions.

The main psychological issues are related to life in isolated, confined, and extreme environments (ICE). The most relevant of them comprise sleep problems, alteration in the sense of time, asthenia, career motivation, homesickness, anxiety, depression, and emotional problems [4, 5]. It is ascertained that astronauts are stressed and this represents an important variable because stress plays a crucial role in human behaviour, cognitive functions and emotions. During long duration spaceflights, chronic stress may result in decreased energy, intellectual impairment, decreased productivity, increased hostility, anxiety, sleep disorders, miscommunication, and impulsive behaviour [6]. The astronaut's mood has several psychological and behavioral effects. For example, a positive mood helps to minimize errors, make faster decisions, promote prosocial attitude and improve the state of physical well-being [7]. On the other hand, mood disorders negatively affect the quality of interpersonal interactions and interferes with normal operating conditions. For example, symptoms of asthenia include difficulty in concentrating, fatigue, and sleep disturbances [8]. A series of possible mental disorders might also potentially occur during long-duration mission - as classified according to the criteria of the DSM V (The Diagnostic and Statistical Manual of Mental Disorders) and/or the ICD 10 (International Classification of Diseases)- such as delirium (an alteration of consciousness that causes problems in concentration, focusing and having a coherent

stream of thought), adjustment disorder (a severe and negative emotional response to a tragedy or significant change in one's situation, leading to considerable discomfort and significantly interfering with daily functioning of the person) and neurasthenia (a progressive negative psychological response to the isolation and rigors of a long-duration mission) [4, 9]. To prevent the emergence of severe psychological conditions, currently on the ISS, a psychiatrist or clinical psychologist has a weekly Private Psychological Conference (PPC) with the crew members. The information gathered during the PPC are used to periodically update the Integrated Medical Model (IMM), a statistically-based tool for forecasting possible risks for the crew health. It includes three behavioural categories: behaviour emergency, depression and anxiety. For NASA, behaviour emergency are behavioural or psychiatric conditions which can seriously jeopardize the mission, for example, a brief psychotic episode due to a tragic event like the death of a loved one or a disaster. Although until now, no emergencies have occurred during USA Spaceflight, their probability of occurrence increases with the increase of the journey length.

Importantly, in long-duration missions additional psychosocial issues are related to interpersonal relationships and group dynamics, including the need for privacy, the impact of crew diversity and leadership styles on small group dynamics. As reported by astronauts, participating in studies conducted on the ISS, adaptive and maladaptive features of ground-crew interactions, processes of crew cohesion, tension and conflict, language, cultural differences, gender biases represent additional psychosocial variables [10, 11]. As the duration of the spaceflight increases and the crew becomes more heterogeneous, psychological and interpersonal factors acquire a greater relevance. Extreme conditions such as prolonged confinement, isolation, and longer communication time delays will require the crew a higher level of interpersonal compatibility and capability to work autonomously, adapting to unforeseen challenges [12]. Thus, the team composition and the configuration of member attributes, become crucial factors to maximize the success of the mission. Good crew members' cooperation, coordination and communication are fundamental elements to increase team performance and teamwork [13]. Team performance and individual well-being are positively correlated with a good leadership that requires the knowledge of several leadership models [14]. The characteristics that may mitigate or exacerbate situations and influence team performance are homogeneity of personalities, shared interests, shared values and norms and other demographic factors such as common language, gender, expertise, age, ethnicity, and nationality [15]. Team functioning and the relationship with control personnel is often compromised by cultural differences and communication delay. Although not much attention is paid to the conflict between Mission Control Center (MCC) and spaceflight crew, the evidence is showing that this conflict might greatly impact the mood of astronauts [13].

Another important aspect to take into account in long-duration missions is the exposure to altered gravitational input, that greatly impact sensorimotor function. Microgravity is perhaps the most salient variable that affects the behaviour of individuals. Research has shown that microgravity in Space settings can affect a wide range of neurobehavioral outcomes. Such outcomes include cognition and mental imagery [16], neurovestibular function [17], posture and movement [18], and visual stability [19]. While vestibular and sensory outcomes in microgravity settings are fairly well

understood, there is a relative lack of knowledge regarding how microgravity affects cognitive and affective functions. The impact of long-term microgravity exposure is even less understood; thus, microgravity constitutes a major risk that needs to be addressed in planning long-duration missions.

2.2 Habitability Issues

Beside the psychological issues described above, there are a series of additional stressors due to the environment and its habitability. Habitability refers to the features of the spacecraft which will take crew members to their destination and return them to Earth upon completion of their mission. The main habitability issues in long-duration mission far from Earth are currently identified as related to the lack of external view, the size of the habitable volume, and the lack of privacy.

Windows provide direct, non-electronic, through-the-hull viewing and are essential to mission safety and success, as well as to maintaining crew psychological and physical health and safety. They are essential for piloting and robotic operations, and permit safe viewing through hatches, stellar navigation, vehicle anomaly detection and inspection, and environmental and scientific observations. Importantly and in relation to psychological wellbeing, windows provide SA (Situational Awareness) of the external environment and support crew photography (a primary on-and-off duty activity of on-board crews). With respect to cameras or display systems, windows do not have the 'failure mode', that may not be operable during emergencies when most needed. A possible countermeasure that has been proposed is the inclusion of a window with a portal (~0.5 m^3 [1.7 ft^3]) in the common area and another window with viewing area to support mission activities when external views are required.

Another important factor of life in Space is related to the size of the habitable volume. The physical Space on ISS or in any Space vehicle is limited and social density is another stressor (NCR, 1998). Indeed, longer duration missions requires expansion in the physical volume of the habitable modules to accommodate mission tasks and personal needs. The total habitable volume for single crewmember and for the whole mission crew should be increased with duration, particularly if the mission is not able to be logistically resupplied. Due to the nature of the mission, the volume of the spacecraft will be necessarily small, although it has been shown that the volume of a habitat can impact a number of outcomes. For example, as discussed below, isolation and confinement in small volumes can lead to stress and other related outcomes as a result of sensory deprivation [e.g., 1]. Indeed, confinement, isolation, and stress that accompany a Space mission tend to increase with duration. This creates a psychological need for additional volume. A larger volume for the crew quarters than in the past, may provide privacy and restoration that will be needed in the long-duration exploration mission owing to the increased period of isolation and confinement. Quarters are clustered together to provide alternative social space and can be personalized (e.g. hanging pictures, varying positioning of bedding). Now in the ISS, the volume for each individual crew quarter is of 2.1 m^3 (74 ft^3). NASA is planning to increase this volume to 5.4 m^3 (190.70 ft^3).

Another important factor influencing the crew members' well-being is related to privacy. Also this aspect assumes a greater relevance as mission duration increases. Lack

of privacy and the constant presence of other people have been reported to impair individual well-being and they are among the most adverse psychosocial stressors of long-duration mission. [20]. Given the increased need for privacy and the occurrence of territorial behavior under prolonged isolation and confinement, the provision of sufficient personal space and private quarters represents one of the most important psychological countermeasure. Personal space needs to be increased with mission duration, although this requirement may be limited by technical constraints.

Finally, additional stressors in the physical environment that is unique to Space include: a growing accumulation of garbage, limited facilities for sanitation, the need for constant vigilance. Currently, the noise and vibration of ISS are acoustic stressors, as well as the low level of illumination, that represents a photic stressor.

Most of the above listed factors are of some psychological importance, and represent important habitability issues to be addressed for the success of future of Space missions. Since a well-designed living and working environment can promote the crew's performance and well-being, the habitat design might also be regarded as psychological countermeasure during a Space flight.

2.3 Questionnaire

On the basis of previous reports and the published literature we built an ad hoc questionnaire aimed at collecting the opinion of astronauts on specific psychological and habitability issues during long-duration missions on the International Space Station (ISS). In the present paper we present preliminary evidence from data collected on one astronaut who experienced two long term stays on the ISS, with the intent to further develop this tool and administer it to an extended sample. The questionnaire on life in Space and its future development might provide useful insights into habitability requirements to be possibly applied in future artificial Space environments (i.e., Space stations). The semi-structured questionnaire, administered via email, was composed of 3 sections concerning: (1) *psychological well-being and crew members interactions* (17 items), (2) *human adaptation and performance in altered gravitational environment* (13 items), (3) *psychological countermeasures currently employed on the ISS, and possible interventions that might be useful to implement in future deep space missions* (7 items).

Section 1. Overall, the interviewee evaluates his experience with the other crew members on the ISS as excellent (item 1), reporting a great cohesion between crew members (item 2) and judging as very important having, once a day, some spare time (for example, at meals) to spend with them (item 3 and 4). The astronaut evaluates the common area on the ISS as very important (item 5) and ranks as follow the first four attributes (out of seven) that might improve its comfort (item 6): (1) a big window, (2) several small windows, (3) a dining table, (4) a bigger size of the room. In relation to psychological well-being, having one day a week to dedicate to other activities than the usual ones is judged of greatest importance (item 7).

Overall, his experience with the Mission Control Centre (MCC) was near to be excellent (scored 9 on a 10 points scale, item 8) and, the most relevant attributes of MCC members, selected among others (emphaty, listening skills, problem solving skills), were *encouraging attitude* and *supportive communication* (item 9). In accordance with

previous reports, the astronaut confirms that the crew's motivation might decrease over time (item 10) and suggests as possible countermeasures, the possibility of having more frequent major operational events (e.g. EVAs) and video conferences (item 11). A greater autonomy of the crew members in planning their work schedule and/or priorities of nonessential tasks is judged as having only average importance (score of 5 on a 10 points scale). When asked to select among items expressing his impression about the work on the ISS he selects, among others (Tedious, Excessive, Stressful, Other), *Adequately distributed between members* and *Time demanding* (item 13). On item 14, the interviewee reports he spent most of his free time on the internet. Interestingly, and in line with previous reports, the possibility to see the Earth during a long-duration mission is deemed as being of crucial importance (score = 10, item 15). On the contrary, the presence of plants on the station environment is judged as less important (score = 5, as Food source and/or *sensory input*, item 16). Finally, when asked to express his opinion on possible countermeasures that might improve the interactions between the crew members and the psychological well-being on the ISS he underlines the relevance of selecting *'the right kind of person to serve as long duration astronaut and then ensure that the right crew complement is assigned'*.

Section 2. Interestingly, when asked how long it takes to adapt to move in microgravity (item 1), the astronaut reports that during the first flight it took a couple of days to be comfortable, 2 weeks to be more coordinated and perhaps 4-6 weeks to become graceful. While during his second flight, he largely adapted on the first day, and it took only few weeks to become graceful. The astronaut's response crucially indicates that motor skills learned during the first flight and relative neuroplastic changes were maintained over time (the time interval between the two flights was of 75 months). On items inquiring about any difficulty in anticipating the trajectory of moving objects (item 2), tactile sensitivity (item 3), Visuo-Spatial Orientation (item 4), and Body Representation (item 5), he does not report any change. The only sensory modalities in which he detected subtle alterations are smell and taste (score of 1 on a 10 points scale, item 7). In relation to Time Perception (item 7), the passing of time is defined as quick for the weeks and slow for the months and the lack of the 24-hour light-dark cycles does not seem to affect time perception (score = 1). On the other hand, a main factor that in the astronaut's opinion mainly affects time perception was the *constant work pace* (item 8). The current colors, labelling and lighting system are judged as being sufficient countermeasures to orient the crew members in absence of the gravitational input (item 9), while other factors that need to be improved (item 10) are the quality of air (score = 7) and the excessive noise (score = 6). The stowage (item 11) and the inventory system (item 12) need to be improved, and the astronaut's suggestion is by using the Radio Frequency Identification (RFID).

Section 3. Thirty minutes of Private Psychological Conference (PPCs), once every two weeks (item 1 to 3) is deemed to have a significant role on the crew's psychological well-being (score = 8). However, during a mission into deep Space, where PPC utility might be challenged by delayed communications, the alternative countermeasures are *increased social support between crew members* (score = 8) and *group cohesion* (score = 8) rather than *crew quarter privacy* (score = 5) or other activities (item 4). In terms of psychological well-being the current habitability of the crew quarters is judged

as just above the average (score = 6) and the 4 main solutions (out of 8 possible items) for their improvement are reported to be the *privacy for personal communication*, *individual recreation (i.e., availability of compact entertainment devices), individual environmental control (e.g., adjustable lighting, temperature), space for stowage of personal items* (item 6). Finally, the out-the-window viewing capability in vehicles and habitats for short and long-duration missions (item 7) results to be an essential factor to optimize task performance and crew well-being (score = 10).

Taken together, these data confirm and further suggest the necessity to find solution to enhance external view, the space habitability size and privacy in the habitat in order to improve the chance of success of long-duration mission in deeper Space.

References

1. Schuster, Alfons (ed.): Intelligent Computing Everywhere. Springer, Berlin (2007)
2. HI-SEAS (Hawai'i Space Exploration Analog and Simulation. https://hi-seas.org/
3. List of spaceflight records. https://en.wikipedia.org/wiki/List_of_spaceflight_records
4. Kanas, N., Sandal, G., Boyd, J.E., Gushind, V.I., Manzeye, D., Northf, R., Leong, G.R., Suedfeld, P., Bishop, S., Fiedlerj, E.R., Inouek, N., Johannesl, B., Kealeym, D.J., Kraftn, N., Matsuzakio, I., Musson, D., Palinkasq, L.A., Salnitskiyd, V.P., Sipes, W., Stusterr, J., Wang, J.: Psychology and culture during long-duration space missions. Acta Astronaut. **64**(2009), 659–677 (2009)
5. Mallis, M.M., Deroshia, C.W.: Circadian rhythms, sleep, and performance in space. Aviat. Space Environ. Med. **76**(6, Suppl.), B94–B107 (2005)
6. Strewe, C., Feuerecker, M., Nichiporuk, I., Kaufmann, I., Hauer, D., Morukov, B., Schelling, G., Choukèr, A.: Effects of parabolic flight and spaceflight on the endocannabinoid system in humans. Revi. Neurosci. **23**(5–6), 673–680 (2012)
7. Amerio, P.: Fondamenti di psicologia sociale. Il Mulino, Bologna (2007)
8. Kanas, N., Manzey, D.: Space Psychology and Psychiatry. Microcosm Press, El Segundo (2003)
9. Slack, K.J., Schneiderman, J.S., Leveton, L.B., Whitmire, A.M., Picano J.J.: Risk of Adverse Cognitive or Behavioral Conditions and Psychiatric Disorders. Behavioral Health and Performance, Human Research Program, Space Medicine Division, NASA Johnson Space Center; Houston, Texas (2016)
10. Kanas, N., Salnitskiy, V., Grund, E.M., Weiss, D.S., Gushin, V., Bostrom, A., Kozerenko, O., Sled, A., Marmar, C.R.: Psychosocial issues in space: results from Shuttle/Mir. Gravit. Space Biol. Bull. **14**(2), 35–45 (2001)
11. Palinkas, A.L.: Psychosocial issues in long term space flight: overview. Gravit. Space Biol. Bull. **14**(2), 25–33 (2001)
12. Sandal, G.R.: Psychosocial issues in space: future challenge. Gravit. Space Biol. Bull. **14**(2), 47–54 (2001)
13. Kanas, N., Salnitskiy, V., Grund, E.M., Weiss, D.S., Gushin, V., Bostrom, A., Kozerenko, O., Sled, A., Marmar, C.R.: Crewmember and mission control personnel interactions during international space station missions. Aviat. Space Environ. Med. **78**(6), 601–607 (2007)
14. Bell, S.T., Brown, S.G., Abben, D.R., Outland, N.B.: Team composition issues for future space exploration: a review and directions for future research. Aerosp. Med. Hum. **86**(6), 548–556 (2015)

15. Landon, L.B., Vessey, W.B., Barrett, J.D., Schmidt, L.L., Keeton, K., Slack, K.J., Leveton, L.B., Shea, C.: Risk of Performance and Behavioral Health Decrements Due to Inadequate Cooperation, Coordination, Communication, and Psychosocial Adaptation within a Team. Behavioral Health and Performance, Human Research Program, Space Medicine Division, NASA Johnson Space Center; Houston, Texas (2016)
16. Grabherr, L., Mast, F.W.: Effects of microgravity on cognition: the case of mental imagery. J. Vestib. Res. **20**, 53–60 (2010)
17. Wood, S.J., Reschke, M.F., Sarmiento, L.A., Clement, G.: Tilt and translation motion perception during off-vertical axis rotation. Exp. Brain Res. **182**, 365–377 (2007)
18. Massion, J., Amblard, B., Assaiante, C., Mouchnino, L., Vernazza, S.: Body orientation and control of coordinated movements in microgravity. Brain Res. Rev. **28**, 83–91 (1998)
19. Koga, K.: Gravity cue has implicit effects on human behavior. Aviat. Space Environ. Med. **71**(9S), A78–86 (2000)
20. Connors, M.M., Harrison, A.A., Atkins, F.R.: Living Aloft: Human Requirements for Extended Spaceflight. NASA SP-483. NASA, Washington DC

Maritime

Identifying Gaps, Opportunities and User Needs for Future E-navigation Technology and Information Exchange

Nicole A. Costa(✉), Monica Lundh, and Scott N. MacKinnon

Division of Maritime Studies, Unit of Maritime Human Factors, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Hörselgången 4, 412 96 Gothenburg, Sweden
{nicole.costa, monica.lundh, scottm}@chalmers.se

Abstract. Since IMO's recent approval of the E-navigation strategy, projects like EfficienSea2 have focused on improving information exchange, safety and efficiency of navigation. To identify current gaps and potential development opportunities, this study investigated the current maritime network, typical work practices and technologies through qualitative research with the Swedish Vessel Traffic Services, coastal pilots and navigation instructors. Maritime operations happen within a resilient sociotechnical system where ship- and shore-based operators work in cooperation to achieve common goals. However, the low interconnection of information technologies was evident, and basic local navigational information was dispersed, sometimes conflicting or inaccessible, increasing uncertainty and complicating decision-making for ship- and shore-based stakeholders. New e-navigation solutions may not only alter work processes but also the organisation of the maritime network.

Keywords: Digitalisation · Connectivity · Safety · Efficiency · Sociotechnical systems · Communication · Control

1 Introduction

The international shipping industry is important to many national economies and global development, carrying out approximately 90% of all world trade [1, 2]. Thus, the safety and efficiency of sea transport have a key impact on sustainable development. The "safety, security and environmental performance of international shipping" are regulated by the International Maritime Organization [3], who has recently approved an E-navigation implementation strategy to be completed by 2019 [4]. The objective is to increase safety of navigation through improved standardised and harmonised communication technologies between ships and shore. In support of this goal, the EfficienSea2 project (http://efficiensea2.org/), within which this study was conducted, aims to identify, develop, test, standardise and implement innovative and smart e-navigation solutions that will improve connectivity and information obtainability between ships and shore, reducing communication limitations and presumably the risk of accidents [5–7]. This study aimed to explore current gaps and overlaps in information exchange between ships and shore, capture the needs of navigational assistance services

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_16

(Vessel Traffic Service (VTS) operators, pilots)[1] and ship bridge operators, and recognise potential e-navigation development opportunities. For this, we focused on understanding current work practices and technologies, and how the maritime network is built today to maintain safe and efficient passage in high traffic density and confined waterways, where most vessel movements are seen and accidents occur [11]. This study presents *research for/into design* [13], offering insight into future e-navigation development projects.

Communication, information integration [10], and trust [11, 14] have been described as key factors in navigational assistance for maintaining safe and efficient operations at sea. Communication establishes a relationship between two or more parties in the exchange of information and feedback [15, 16]. This generates knowledge, helps predict behavioural patterns, and is essential to maintaining good teamwork, safety and efficiency at work.

The general systems theory of Cybernetics, which from the Greek means "steersman" or "rudder" [16–18] refers to a dynamic and complex system being steered, controlled and regulated [17] towards the achievement of system goals [16], allowing the system's behaviour and response to be predicted [18]. Controlling the system depends on communication [18, 19], information processing, planning/adaptation and decision, and is measured through the extent to which the system is achieving its intended goal(s) (*effectiveness*), with a minimal use of resources (*efficiency*), and contributing to the goal(s) of the higher-level system in which it is contained (*efficacy*) [18]. The maritime domain is commonly considered such a dynamic, complex and safety-critical sociotechnical system [9, 10, 20, 21], where a network of interdependent social actors and technological tools interact, oriented towards a common goal [22–24] of running their operations efficiently and safely [10]. Embracing the sociotechnical approach that treats the technical and the social as a whole can cater for a better design of future work systems [25].

2 Methods

This study focused on the Swedish maritime sector. All respondents had studied nautical sciences and/or had significant onboard experience. Four visits took place in the Gothenburg VTS operated by the Swedish Maritime Authority (SMA), beginning with a briefing given by a VTS instructor/operator followed by a combination of naturalistic direct observations with open-ended, semi-structured interviews [26, 27] with a sample of six VTS operators (VTSOs) and one pilot planner. The objective was to learn about their everyday operations in this complex and dynamic work milieu [27].

[1] The VTS is a shore-based organisation that monitors and helps coordinate traffic, providing updated local information to vessels navigating the respective VTS area around ports [8]. VTSs can provide one of the three services: Information Service (INS), Traffic Organisation Service (TOS), or Navigational Advice and Assistance Service (NAS) [9, 10], being INS the most common one [11], present in Gothenburg and Malmö. Pilotage is a service provided to the ship for a fee and is compulsory in many ports. Pilots are mariners with detailed local knowledge who board a ship to help the captain navigate, berth or unberth safely and efficiently in constricted waters [12].

The focus began broadly with how the operators made decisions about safe distances and safe depths, which after initial analysis developed into understanding the technological mechanisms and stakeholders behind it, and gaps in information exchange. Each visit had the duration of 3 to 5 h, beginning at different hours of the morning to capture different traffic patterns, changes of shifts and hand-over procedures. An open-ended, semi-structured interview [26] was done with two coastal pilots from SMA, active in the Malmö area, to learn how they contribute to maintaining safe passage. To discuss further bathymetric and weather data updates to vessels about the Malmö area, an interview with an SMA project partner was held. A semi-structured telephone interview with a Malmö VTSO occurred to specifically discuss reporting formats. Informal interviews [26] and two interviews with walkthroughs were held with three full-mission bridge simulator instructors from the department of Mechanics and Maritime Sciences at Chalmers University of Technology to discuss elements of voyage planning.

The analytical process began with data collection and progressed continuously, having the initial data helped determine what to follow up on next (*theoretical sampling*) [28, 29]. The sessions were audio-recorded, transcribed and highlighted where relevant with *memos*. *Open coding* and data grouping proceeded iteratively and incrementally, based on Grounded Theory [28].

3 Results and Discussion

The results suggest that anti-collision and anti-grounding strategies were important for the participants both in voyage planning and execution for the maintenance of maritime safety. Achieving this common goal in local waters requires a complex network of maritime stakeholders and information systems to share data, coordinate decisions, and monitor/control operations. In the VTS, safety was described as being dependent on (a) the communication of updated information to the vessels, ensuring that this information is well understood to provide the overview of the traffic; and (b) the monitoring of traffic to make sure it follows the "normal patterns", accounting for all parameters that can affect safety. For the pilots, safety was defined as making a voyage plan with safety margins, such as distance to the shallow waters, "*distance to the dangers*". The need for thinking ahead and having a preventive mind-set also factors into safety: "*We always have a plan B*". For navigators, "*If you look at it from a more holistic point of view, I would say that voyage planning is really just risk assessment, actually. (…) to handle the hazards that you have along the route and mitigate these risks. (…); Everything we do and see will be referred to that voyage plan*". However, systems were poorly interconnected and basic navigational information dispersed, sometimes conflicting or inaccessible, increasing uncertainty and complicating decision-making. Moreover, there were disparities in the way ships and shore perceived reality. While the VTS had access to an overview of the traffic and information, each vessel had richer individual knowledge of its surroundings, being focused on itself and its operations (also see van Westrenen and Praetorius 2012 [30]): "*As a VTS you have to be a little bit more split in your view*"; "*It's very different to make decisions from here, depending on vessel size, wind situation (…). That vessel here and there; he*

knows that himself best". Being in the VTS did not make up a full view of reality since, on the screen, vessels did not have fully realistic proportions and distances might seem smaller than in reality. It became thus harder for the VTSOs to predict ship movements: "*The pilot onboard, he can see ahead. You cannot see ahead here*". For this reason, the VTSOs had to rely on the pilots and bridge officers to do as they saw fit and follow the rules: "*Hopefully they're following the rules. Then we don't have to do nothing*". Breaching the expected behaviour or communication patterns alerted the VTSOs to certain vessels as potential "trouble-makers". In this sense, the VTSOs repeated the words "normal" and "pattern" to describe what they expected to see and hear, and what triggered them to intervene and contact the vessels or keep a closer look at them. These terms were used to express expected navigation behaviour and communication patterns (also observed by Brödje 2012), suggesting informal prescriptive norms (e.g., "*Normally the vessel turn starts here on the way*"; "*Normally they go here*"). There were also common procedures that were not based on rules or regulations but that have become recommendations, and behaviour that deviated from the norm was unpredictable and control was lost both from the perspective of the system not being equipped to alert on this as from the VTSOs not knowing when and how to intervene in the best way. How information was shared by the VTS to vessels also followed a pattern and this had an effect on efficiency: "*When it comes to the information we give out, it could be that it's consistent, that it follows the same pattern so it's recognizable; that we don't talk too much – just give the essential information*". Consistent behaviour might become accepted even when it broke some norm at one point (e.g., the pilots on a cruise line always taking a shortcut, "*Yes, a little bit tight. They could use the fairway better, they really could. But that's really common*"; "*I know they [pilots] always go there [over the marked shallow areas] (…) But if it had been another type of vessel, I would have reacted*". So, considering the lack of a full common picture among stakeholders, an opportunity exists for future route sharing to help operators predict the moves of others and plan their own. There is also an opportunity for keeping historical records of normal navigational patterns, and for the treatment of big data.

Other historical records that could benefit the operators were water currents. For the pilots, there were now buoys outside the port of Malmö measuring currents. These were useful for pilots to calculate whether a certain manoeuvre was possible, estimate entering speed and drifting angles. However, these buoys were not working during night time, so the pilots had to guess based on experience, colleagues, history of currents. A VTSO asked one of the interviewers if she knew how to get wind history. This reflects it would be useful if the system could keep a historical database of the wind as well. Incident history, at another point, also seemed difficult to go through, i.e., there was no good search function to find a particular incident: "*I have to think where to find it*". As it was now, they tried to keep incident reports and save forecasts in the email. In summary, the capability to store, analyse and transform historical data about normal navigational patters, distances to vessels or to shallow waters, typical voyage plans, average water currents in certain locations, and common incidents, has proven relevant. Behaviour that breaks norms but that becomes common for certain vessels and therefore accepted by the VTS would be a part of the normal navigational patterns and this could identify potential alternatives for other vessels' voyage plans. If this

knowledge can be integrated among the stakeholders, it should also help predict movements in close-quarter waters.

As for the information dispersity and access limitations, these gaps were identified for different functions, as described in Subsects. 3.1 to 3.9. The integration of systems and data across the international maritime domain can be used to support communication and cooperation among the global operator network in voyage planning and execution. This includes an extended capacity to connect software and hardware more seamlessly (for automatic, regular and accurate updates of the systems with sensor data, e.g., draught).

3.1 Voyage Plans

Voyage planning may need to be done within only a couple of hours before the ship leaves berth, depending on how much time is available from the moment the ship is given a new destination. It is commonly done digitally whilst consulting the printed or electronic onboard publications, which are not in most cases incorporated into the planning software (e.g., the reporting points and what to report to VTS). One can easily get the electronic charts for all locations, but more detailed local information may not be readily accessible and adds to the workload.

Onboard, pilots will look at both voyage plans – the ship's and the pilot's – and see if they match. Considering that pilot fixed voyage plans are downloadable over the internet, the routes should be expected to coincide, but this is often not the case. Limited access to information or having to skim through endless sources of information has a negative impact on vessel voyage planning. The pilots mentioned seeing voyage plans with the wrong draughts or other errors. Pilots will recommend route adjustments to the captain and to continue collaborating onboard. However, some captains might still insist not to proceed any farther in the voyage with the pilot, and that is when the pilot will warn the VTS of the inappropriate voyage plans of that particular vessel, so that the VTS can take more precaution. For this reason, the pilots wished there were stricter regulations in the Malmö area to take on a pilot. This issue reflects shortcomings in incorporating local information in voyage planning and emphasizes the potential usefulness of the VTS accessing ships' voyage plans. Moreover, some shipping companies require additional paperwork and approval when adjusting voyage plans, which is yet one more reason to improve the incorporation of local information in strategic voyage planning rather than tactical (on the go) modifications. Indeed, voyage plans, more than destination, were reported relevant for the VTS to know the ships' intentions and intended routes more in advance and facilitate certain predictions and decisions from the VTS side of which vessels will be likely to affect vessel A: "*Sometimes it's quite difficult from here to see the situation*".

3.2 Hidden Closest Point of Approach Rules

Pilots have generally better knowledge of what is happening in their location even compared to captains with pilot exemption (*"Guys with exemption can actually get this*

information as well but then they have it in the website. The pilots have it in the phone. And I can understand that that website is not as up to date (…)") on currents, maximum draughts and depths, the usual traffic flow, and hidden Closest Point of Approach (CPA) rules (e.g., "Like ferries between Helsingborg and Helsingör in the Sound, they have unwritten rules about the regulations and about CPA limits (…)"). Every time the VTSOs were asked about CPA rules for VTS areas, they insisted there were none. Although this is true concerning the regulations, the pilots mentioned "hidden" CPA rules of 0.3 nautical miles (556 meters; around 3 cables) among the ferries in the Malmö area. These norms were not known to the respective VTS nor to foreign ships who navigated through the area infrequently, representing another gap in information exchange and ultimately in knowing how to navigate the local waters.

3.3 Automatic Identification System Information

Automatic Identification System (AIS) information was not always correct or there at all (e.g., draught), and it was limited. For this reason, the radar was generally more trusted than the AIS (also observed by Brödje 2012). Ships could not see each other's destination or berth nor access an overview of information about the whole VTS area traffic, whereas the VTS could see the ships' destinations but not the berth either. The VTSOs had thus to be provided information by the port controller or access other systems such as the port system to see which berth a ship was assigned. VTSOs would also have to check SMA's website to see specific wind speed limits for berthing and unberthing. The VTS itself lacked information that would allow it to expand its assistance to ships, information which was possessed by the ships such as voyage plans, by the port controllers such as berths where ships were scheduled, or by the pilots such as hidden CPA rules.

To get correct departure and arrival times, the VTSO worked with pilot dispatch and port, *"And that's the tricky part, because it's so many people involved (…): it's tugs, it's linesmen, it's crane operators, it's agents, it's paperwork (…). We try to work together; get the latest information all the time"*. All this missing information could potentially be updated through AIS data to all operators.

3.4 Draught

Draught was the most valued parameter concerning anti-grounding measures, yet it proved to easily become a possible source of ambivalence between stakeholders. It was the Pilot Organisation board who assessed and established the maximum draught allowances for each berth and channel within a VTS area. The draught of a vessel determined which channel in the VTS area the ship should take and to which berth it should head. It also determined, along with ship size, if it required a pilot and at which pilotage level. VTSOs, pilot planners and pilots usually first received a vessel's draught from the port agent. These were generally incorrect for any given vessel, so they confirmed it with the AIS information and/or contacted the ship. The port agents themselves, though, relied on the ship captain to fill out information forms and provide

them with the draught, which was often incorrect or forgotten, forcing the port agents to guess. There was no standardised format to exchange this information, and in the end this diminished trust among shore-based operators and limited various stakeholders: "*(…) agents you can never trust*". Moreover, the AIS required the bridge officers to insert the information manually. Draught information given by the AIS transponder and in the pilot planning system were sometimes conflicting for the same vessel. Although the AIS is something the VTSOs relied upon on a daily basis, they preferred to confirm with the bridge officers directly. The port control also had a draught value. The bridge officers might not know the draught themselves if they have not recently measured it. Having automatic AIS transponders that measure it regularly cost the shipping companies money, and it was not mandatory. As captured in an interview with an SMA member, under-keel clearance values might change drastically in certain areas/channels over time, but were not regularly measured. In this case, the ships were basing their decisions on obsolete values and actually running a risk of grounding. Thus, more frequent measurements should be customary. In the end, it is relevant to consider how many times one vessel will get the same question about the draught from different service providers.

For the draught, other parameters were deemed important (e.g., the water level and the currents). The charts did not account for these parameters, so operators had to refer to weather services such as the Swedish Meteorological and Hydrological Institute (SMHI)'s Bizmet for weather forecast and SMHI's ViVa for real-time measurements of water level, currents, visibility, wind direction and speed (e.g., pilots had this information available online on their pilot charts and the VTS had screens dedicated to showing this information). Information like water level could be checked by the bridge officers, but the pilots reported that they might have "*lack of knowledge (…). They don't know that they will need it and maybe not aware that the water level is low. Maybe they do not expect that they should know*". Along the same lines, the pilots reported that vessels did not always realise that they could not pass in a certain passage with a given draught or skip taking a pilot: "*They are creating a lot of issues for the VTS*".

3.5 Shallow Water Contours

Another parameter for anti-grounding was the vessels' shallow water contours and how these were set up seemed to differ between operators, and depended upon ship size, draughts, familiarity with the Electronic Chart Display and Information System (ECDIS) and workload. Whereas pilots had blue shallow water contours on their portable pilot unit/pilot pad (often referred to as "no-go areas") triggering alarms at a 10-m depth, vessels would mark it at 30 m instead, optimised for open waters. However, when they got to the confined areas, they might not modify their ECDIS and this would result in most of the navigable waters being displayed as non-navigable. Allegedly "*quite many seafarers*" either had no awareness of this ECDIS setting – that one could have a route with different shallow water contours and alarms throughout, or the workload to set the route this way became significantly increased. Planning in advance for different alarm limits for different areas was possible, "*but the workload on*

the planning navigator is quite much higher". At the VTS, an alarm was set for when a ship crossed the VTS' shallow water contours (marked in purple and called "the multipurpose areas" or sometimes also referred to as "no-go areas"), long established as a result of past incidents in the area. The alarm was set for different timings depending on the draughts of the vessels. However, because the ECDIS was flexible and used by different vessels with different draughts with therefore different needs with regards to shallow water contours ("*built for everyone*"), the pilots suggested that it should be connected to the draught of the vessel, measured in real-time to be able to trust that the shallow waters were adjusted accordingly. The AIS transponder could be connected to the onboard equipment and systems to capture updated measurements of the ship's draught. Although more costly, this is an opportunity to avoid the need for manual input or dealing with uncertainty, especially in areas with moving seafloors.

3.6 Pilot Scheduling

The pilot planners and the VTSOs shared the pilot planning system. Once the pilot planner updated the pilots' and tugs' schedule, the VTSO got the updated information on their computer. Ships with pilot exemption did not show up on this system, so the VTS would know of them once they reported inbound the VTS area. The vessels were, however, scheduled in the port and therefore the expected arrivals of all ships, taking pilot or not, were actually available on the website of the Port of Gothenburg, and this was information that could be known to the VTSOs in order to have a more complete picture of the traffic.

3.7 Temporary Warnings to Mariners

The VTSOs were informed of temporary operations, such as work being done at the quay, diving operations, construction work, as examples, from the entities preforming the work. These entities would contact the port controller who would authorise (or not) these activities, and once this was authorised, the entities would then inform the VTS that it was occurring. The VTSO would insert navigational warnings into yet another system and report them to the Maritime Safety Information (MSI), mark temporary operations on their ECDIS, and inform passing ships about this through VHF so that they could use slow speed or take another route as appropriate, or this temporary information would also be transmitted through the NAVTEX system [NAVTEX is also a (paper) medium for weather forecasts and maritime emergencies]. This sort of activity needed to be inserted in different systems and communicated with the port controller and the VTSO separately even though in the case of the Gothenburg VTS the port controller and the VTSO sat in the same room.

3.8 VTS Reporting

It was observed occasionally how vessels would report to the VTS inconsistently, miss/skip certain information or ask the VTS which VHF channel to use, what and

when to report to the VTS as examples. Reporting details were clearly indicated through several sources, such as the onboard publications and the SMA website, yet not always done as required ("*Almost 99% of all knows that (…). Sometimes [they might miss it], but then they ask*"). This reduced the VTS' trust in the vessel and alerted them to keep a closer look: "*What else could he be missing?*". This issue can signify poor voyage planning, overreliance on the VTS, or misjudging the reporting requirements. It was also observed that finding relevant information in the onboard publications could be time-consuming and contradictive (e.g., understanding where there is mandatory pilotage). It was also cumbersome to access yet another information source probably better known to the locals than to foreign ships that infrequently travel to a particular location.

One issue identified repeatedly was the interference of smaller vessels and pleasure crafts, especially a problem during Spring and Summer: "*Quite a lot of incident reports are related to close-ups with small vessels*". The VTSOs talked about how these vessels did not usually have AIS or VHF and therefore were not identifiable or contactable. They did not have to report to the VTS and were not included in the VTS system alarms. This was considered a positive thing or else the workload of the VTSOs would increase immensely. On the other hand, they might interfere with other traffic that the VTS did keep contact with, and it became a problem that they could not be contacted: "*(…) if they interfere with the traffic, so this is very good with the AIS and he listens to channel 13, so he's available*"; "*We have to believe that all vessels will have to move, because I can't give way to 20 vessels ahead of me. It's like mission impossible*".

3.9 Local Weather

Local weather could be relayed to pilot exempted vessels by the VTS, and to vessels that did take pilot(s) by the pilot(s) themselves and by the VTS. Ships could access most weather information but they mostly had to rely on the VTS and pilots for wind information that was more specific to the area and berth where they were headed. It was reported that ships in the Malmö area did not have access to local water level information, so this was the responsibility of the VTS to share. For example, there was a fee to be paid by ships who did not order pilots and tugs at least 5 h in advance. The captains did not allegedly appreciate this rule because this often made them order tugs, but once the service occured the wind might have already calmed down. This was difficult to predict and became an added cost to the shipping company especially if ships did not have access to local weather information.

4 Conclusions

Field studies, observations and interviews were employed in this study to investigate work activity performed within the dynamic and complex maritime system [27, 31]. The sample consisted of experienced mariners in the Gothenburg and Malmö areas, considered two of the most traffic-intense areas in Sweden. Using the VTS as a starting point, information describing the conduct of a large number of vessels and roles played

by the stakeholders could be efficiently captured. Further field studies on the perspective of the individual vessels being aided by the pilots and the VTS are recommended for future research. The wide array of information sources and communication channels, and the unpredictability of events and of the external environment were observed. The general conclusions taken here resulted from empirical evidence that is indicative of the mechanisms of the overall maritime network and offers direction on what to investigate when going back to the field and on the development of future e-navigation solutions.

Shipping is a self-organising network where each ship is its own principal decision-maker and responsible for maintaining own safe operations. Whereas in general terms the domain is a loosely coupled [32] sociotechnical system with distributed control [9, 33], safe navigation and manoeuvring of large vessels within VTS areas is an activity that requires coordination between the bridge officers of the ship, the VTS [9], and often the coastal pilots [10]. Here, decision-making is based on communication between ship and shore operators and on local information integration for preparation and prediction [10]. The stakeholders, who are mostly not co-located [14], are linked and aided by information technology devices which were seen to not be necessarily aligned with the devices of other stakeholders and therefore caused gaps of information, of perception of the other operator's perspective, and of trust. Even for those shore-based workers who were co-located, the lack of information integration triggered information sharing repetitions. The information contradictions, overload or lack thereof produced uncertainty and compromised performance (*efficiency* and *efficacy*) and control [18, 19]. Information proved to be rather compartmentalised, implying additional procedures and communication channels among the local stakeholders to bridge the information circulation gap, whilst the individual ships still relied on local services to capture local information. The VTSOs have used the same system for the past decade, and each stakeholder had a plethora of own systems with different levels of integration of local information. In addition, as physical location of the operators differs from centre to centre and it has impact on information flow, this expresses a need for synchronised systems for information exchange that does not depend on the physical location of the operators. Furthermore, the issue of harmonising the delivery of VTS has been under debate in the International Association of Marine Aids to Navigation and Lighthouse Authorities (IALA) [34], which should help foreign vessels better understand what to expect from VTSOs as well as how to report. The variety of requirements coming from all potential users and the need to coordinate them for technology design may reflect a wicked problem [13, 35]. Yet, a standardised solution for integrated information broadcast across the domain, taking into consideration software and hardware, is an obvious strategy to aid operators globally in voyage planning and execution, and provide a common baseline to geographically distributed stakeholders. This does not imply that spoken communications through VHF or telephone should be minimized or eliminated, as this helps reduce uncertainty [19], gain insight into the skill of the navigator and builds trust [10, 14], as well as helps the VTSOs obtain and update their mental models of the local traffic [11]. However, data access difficulties, ambiguity and information sharing repetitions occupying the VHF (e.g., draught) or adding to the workload could potentially be eliminated or better managed.

Other needs identified point at evaluating certain regulations/recommendations or lack thereof: (a) hiring a pilot had a key impact on the access to local information, but this was not mandatory in all ports, as well as it was not mandatory to take regular measurements of the draught connected to AIS for increased accuracy; (b) there was no capability to identify and contact smaller crafts, which also proved to be important for maintaining safety. Limited knowledge of ECDIS functions was also referred to as an issue, emphasising that although more advanced systems may bring numerous advantages, training and education cannot be overlooked.

With the emergence of intelligent systems having the capacity to render big data and even the popular notion of a future with unmanned ships, it is important for future research to consider issues of introducing new automated technologies, such as operator overreliance or resistance, the amount of information [21] and how it is presented [36]. It is crucial to consider that more information processing may shift decision-making power to machines; to consider social sustainability and how maritime stakeholders and their roles will be affected or the control/power they currently hold will be shifted, transforming the workings of the overall network; how they will be trained and their skills maximised. Another consideration is whether more access to readily usable local information and integration in voyage planning and execution would expand the potential for a centralised capacity of traffic monitoring and assistance [10] or instead render navigators more and more independent from shore-based assistance services, emphasising the control-distributed nature of the domain (see also van Westrenen and Praetorius 2012 [30]). In a transition towards integrated information systems, there is still a commercial need for manufacturers to develop their own solutions in a competitive market. If the maritime domain is to adopt a new global technological infrastructure, standards and regulations must be put in place and the transition must be managed to incite adoption, especially when different parties will have different requirements.

Acknowledgments. The authors would like to thank the European Commission and its project EfficienSea2 [grant nr. 636329] for financial support.

References

1. International Chamber of Shipping: Shipping and World Trade (2015). http://www.ics-shipping.org/shipping-facts/shipping-and-world-trade
2. Stopford, M.: Maritime Economics. Routledge, New York (2009)
3. IMO: Introduction to IMO (2016). http://www.imo.org/en/About/Pages/Default.aspx
4. IMO: E-navigation (2014). http://www.imo.org/en/OurWork/Safety/Navigation/Pages/eNavigation.aspx
5. Billesø, M.B.: BALTICO Meeting 2016: IHO Navigational Warning CG (S-124) and EfficienSea2 Project (PowerPoint Presentation). Danish Maritime Authority (2016). http://www.sjofartsverket.se/pages/105078/9,%20EfficienSea2%20.pptx

6. Juhl, J.S., Sørensen, A., Hosken, E., Halliwell, M., Pedersen, P., Bondo, P., Nielsen, S., Krog-Meyer, P., Stuer-Lauridsen, F.: Deliverable 5.3: Development of a New Common Port Database Concept and Structure. BIMCO (2016). http://efficiensea2.org/wp-content/uploads/2016/05/EfficienSea2_Development-of-a-new-common-port-database-concept-and-structure-deliverable-5.3.pdf
7. Danish Maritime Authority: 8-9 Nov 2016: Conference – Getting connected to the future (2016). http://efficiensea2.org/event/conference-getting-connected-to-the-future/
8. IMO: Vessel Traffic Services (2016). http://www.imo.org/en/OurWork/Safety/Navigation/Pages/VesselTrafficServices.aspx
9. Praetorius, G.: Vessel Traffic Service (VTS): A Maritime Information Service or Traffic Control System? - Understanding Everyday Performance and Resilience in a Socio-technical System Under Change. Thesis for the Degree of Doctor of Philosophy. Chalmers University of Technology, Gothenburg, Sweden (2014)
10. de Vries, L.: Success factors for navigational assistance: a complementary ship-shore perspective. In: Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2014 Annual Conference, pp. 175–186. Human Factors and Ergonomics Society, Lisbon (2015)
11. Brödje, A.: Hello, Is There Anybody Out There - Just Nod If You Can Hear Me: Exploring the Work of Vessel Traffic Service Operators. Thesis for the Degree of Licentiate of Philosophy. Chalmers University of Technology, Gothenburg, Sweden (2012)
12. IMO: Pilotage (2016). http://www.imo.org/en/OurWork/Safety/Navigation/Pages/Pilotage.aspx
13. Lurås, S.: Systemic Design In Complex Contexts: An Enquiry Through Designing a Ship's Bridge. Dissertation for the Degree of Doctor of Philosophy. The Oslo School of Architecture and Design, Norway (2016)
14. Bruno, K., Lützhöft, M.: Virtually being there: human aspects of shore-based ship assistance. WMU J. Marit. Aff. **9**(1), 81–92 (2010)
15. Flin, R., O'Conner, P., Crichton, M.: Safety at the Sharp End: A Guide to Non-technical Skills. Ashgate Publishing Limited, Aldershot (2008)
16. Miller, K.: Organizational Communication: Approaches and Processes. Cengage Learning, Wadsworth (2012)
17. Woods, D.D., Hollnagel, E.: Joint Cognitive Systems: Patterns in Cognitive Systems Engineering. CRC Press, Boca Raton (2006)
18. Skyttner, L.: Basic ideas of general systems theory. In: General Systems Theory: Problems, Perspectives, Practice. pp. 49–109. World Scientific Publishing, Singapore (2005)
19. Johansson, B.J.E., Persson, P.-A.: Reduced uncertainty through human communication in complex environments. Cogn. Technol. Work **11**(2009), 205–214 (2009)
20. Lützhöft, M., Grech, M.R., Porathe, T.: Information environment, fatigue, and culture in the maritime domain. Rev. Hum. Factors Ergon. **7**(1), 280–322 (2011)
21. Grech, M.R., Horberry, T.J., Koester, T.: Human Factors in the Maritime Domain. CRC Press, Boca Raton (2008)
22. Vicente, K.J.: The Human Factor: Revolutionizing the Way People Live with Technology. Routledge, New York (2006)
23. Fox, W.M.: Sociotechnical system principles and guidelines: past and present. J. Appl. Behav. Sci. **31**(1), 91–105 (1995)
24. Hendrick, H.W., Kleiner, B.M.: Macroergonomics: an Introduction to Work System Design. Hum. Factors Ergon. Soc., USA (2001)
25. Kirwan, B.: Soft systems. Hard Lessons. Appl. Ergon. **31**, 663–678 (2000)
26. Patton, M.Q.: Qualitative Research and Evaluation Methods. Sage Publications Inc, Thousand Oaks (2002)

27. Stanton, N.A., Salmon, P.M., Raffery, L.A., Walker, G.H., Baber, C., Jenkins, D.P.: Human Factors Methods: A Practical Guide for Engineering and Design. Ashgate Publishing Limited, Aldershot (2013)
28. Corbin, J.M., Strauss, A.L.: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage Publications Inc., Thousand Oaks (2008)
29. Czarniawska, B.: Social Science Research: From Field to Desk. SAGE, Thousand Oaks (2014)
30. van Westrenen, F., Praetorius, G.: Maritime traffic management: a need for central coordination? Cogn. Technol. Work **16**(1), 59–70 (2012)
31. Maguire, M.: Methods to support human-centred design. Int. J. Hum. Comput. Stud. **55**(4), 587–634 (2001)
32. Orton, J.D., Weick, K.E.: Loosely coupled systems: a reconceptualization. Acad. Manag. Rev. **15**(2), 203–223 (1990)
33. Hollnagel, E., Woods, D.D., Leveson, N.: Resilience Engineering: Concepts and Precepts. Ashgate Publishing Limited, Aldershot (2006)
34. IALA: IALA's Strategic Vision 2014-2026 (2013). http://www.iala-aism.org/content/uploads/2016/06/iala_strategic_vision-1.pdf
35. Buchanan, R.: Wicked problems in design thinking. Des. Issues **8**(2), 5–21 (1992)
36. Oh, J., Park, S., Kwon, O.-S.: Advanced navigation aids system based on augmented reality. Int. J. e-Navig. Marit. Econ. **5**(2016), 21–31 (2016)

Inshore Navigation's Yacht: Study Cases of Sustainable and Ergonomic Yacht Design

Massimo Di Nicolantonio(✉)

Architectural Department, G. D'Annunzio University, Viale Pindaro 42, 65127 Pescara, Italy
m.dinicolantonio@unich.it

Abstract. Coastal and inland navigation is a typical activity of yachting, which allows people to visit beautiful and pristine seascapes. These places have very special characteristics. Often they have very shallow water and no basis for mooring; at the same time, most of the marine and inland coasts are not navigable because of their nature of marine reserves. This requires a suitable typology of vessel for yachting, like small pleasure crafts. This first phase of the research reports a series of critical reflections about the conscious design of small crafts, in a specific environmental context. It highlights how a competitive and strategical design for the revaluation of the tourism offering, simultaneously with the conscious design of sustainable products, can be a starting point for viewing the relationships and interactions between the different subjects that make up a system innovation.

Keywords: Slow mobility · Inshore navigation · Accessibility · Marine tourism · Solar design · Sustainability

1 Introduction

Coastal systems are areas of significant tourist and environmental interest: nautical tourism represents one of the essential resources for the sustainable development of the planning and the management of services related to tourist accommodation business. The management of the use of coastal systems requires different skills that aim to develop sustainable strategies that combine all the activities for the protection of natural resources. The coast's evolution is governed by oceanographic and terrestrial processes and by human settlement and zoning activities. All these actions influence the strip of territory in contact between land and sea, whose extension depends on the type and coastal climate regime. Since the early 70s, in Italy, all institutional authorities moved in favor of the regulatory framework reorganization. Since the 80s, the program to institute marine protected areas was started. The Italian situation is a clear example of how, over the years, the environmental and legislation conditions have changed (Fig. 1).

New modalities of aquatic environments's usage were necessary (for example the establishment of marine protected areas like Torre di Cerrano in Abruzzo, etc.). This awareness led to the will of conceptual reinterpretation of nautical artifacts, because of new social and usability models, opened to a wider audience. In part, the boating industry today aims to diversify and evolve in favor of the principles of social inclusion and

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_17

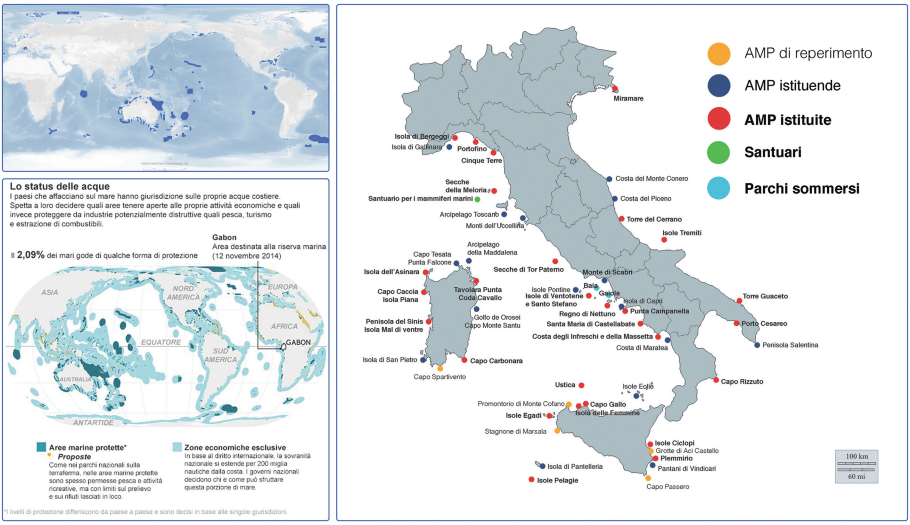


Fig. 1. World marine areas and coastal Italian system

environmental protection. In this scenario, the *slow mobility design* can carry out a search action and the offering of innovative products. Proceeding by stages, investigating the nature of the problem, it was possible to delimit the main aspects and to propose possible solutions using the tools of knowledge. New environmental resource management models can match a design approach aimed to redefine the user-product-environment relationships through the conscious design of nautical artifacts, with the aim of contributing to the sustainable development of coastal and inland nautical tourism.

2 Inshore and Inland Slow Mobility in the Abruzzo Region

The Abruzzo Region's river waterways are predominantly torrential. This unconformable territorial assessment is not slightly usable. A serious action of redevelopment needs of a serious geological intervention. This action of redevelopment especially regards river courses themselves. However, while in this moment the hinterland does not present real opportunities of development, the coastal regional development is much more favorable. The first example is the program named *Bike to Coast*: this program is focused on the redevelopment and renovation of the disabled railway track positioned near the coast. The main purpose is to create a new functional and organized touristic and bicycle track. A second way of action could be to improve the functionalities of the entire touristic coast system by the adoption of an intermodal slow mobility concept named: *Bike and Boat*. The first phase of the research focused on the survey carried out to assess the possibility of tourism development in Italy, according to the implementation of inland and coastal waterways. All the collected data show a widespread desire for system enhancement. There has been many proposals in Abruzzo, as: municipalities, investors, and tourist offices. The data analysis contribute to define a competitive diagram for the

Abruzzo touristic product. The program operates on a large scale, and connects to the European network for the implementation of the touristic slow mobility system. Other european nations have already invested material and human resources and materialized targeted intervention efforts to the development of intermodal mobility. The study for the Abruzzo region focused on the redefinition of the roles and services, for the implementation of touristic inshore and inland slow mobility services, connected to all the other touristic services (Fig. 2).

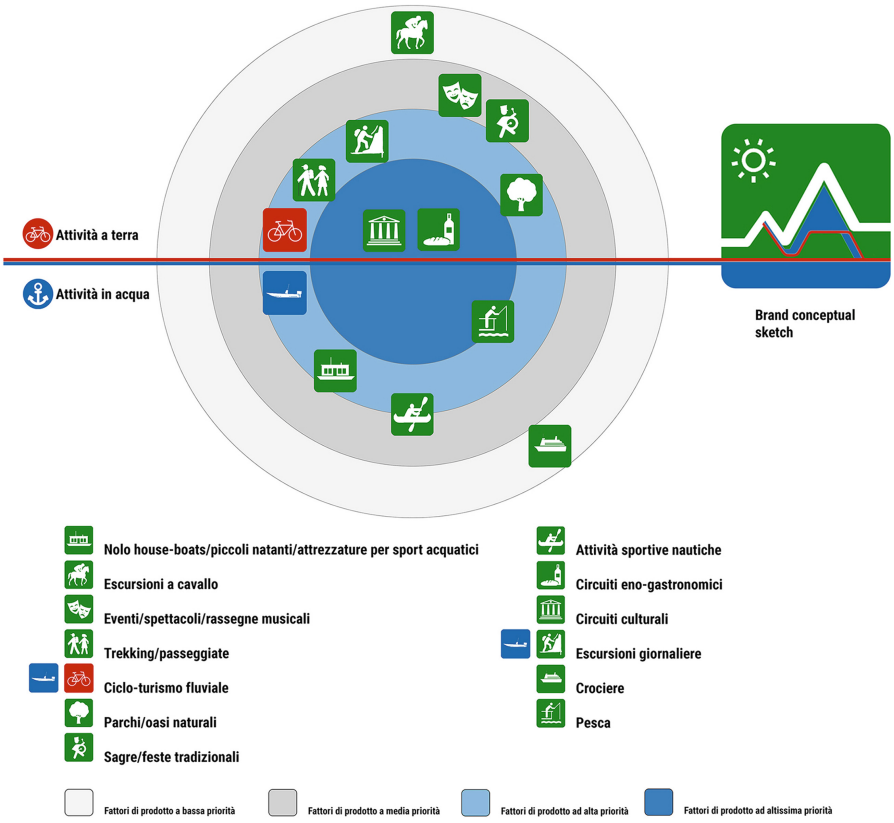


Fig. 2. Competitive diagram for Abruzzo Region

The concept of a service design based on the usability of small pleasure craft is suggested as one of the possible solutions to improve the maritime sector and intermodal transportation. This choice leads itself benefits that help ensure system efficiency, through the development of intermodality, and new criteria for the management of river and maritime territories. Its can be considered an available and feasible for its particularity: cheapness, environmental quality, research of psycho and physical pleasantness. This model requires low impact environmental's solutions, and human based solutions, such as: systems for the good and conscious management of small pleasure crafts, services for river and coastal tourism, services like boat sharing and other options.

3 Design Process

Accordingly to the issues proposed, an analytical investigation process has been activated. A first phase of survey, by the adoption of design research tools, was oriented to the compression of the nautical filed, with specific focus to the small pleasure craft. The exploratory survey was oriented in the world of the nautical sector with specific reference to vessels for the minor pleasure. The contextual analysis took into consideration many aspects: environmental, social, cultural, economic, managerial, technical and technological. New values have been brought out: the social inclusion, the marine eco system protection, the designers sensibility, the diversification of market demand, all these aspects related to the sustainable use of marine environments, innovation, social value of the product, the economic value of the product, ergonomics and living comfort. Every feature has transversely contributed to optimize the whole process, through different stages: reconnaissance, ideational, design, production, marketing. Beginning evaluations were made taking account of the various aspects, such as: accessibility, ergonomics, importance of physical and visual relationship between artifact and environment, relationship between artifact and user, relationship between environment and user. New fruition models needs research and new functional solutions and new formal models. The direction is that of a substantial innovation process. This way requires a survey on new languages. The new models proposed, as an expression of the innovation, must be consistent with the indicated conditions. Innovation means to opposite and overcome a crisis. Innovations means to take advantage and turn negative features into opportunities. Innovation means to propose a strategic model aimed to reduce or eliminate all the conditions that have contributed to spread negativities. The designer's goal is shared by the industry. The first target is to set the guidelines redefinition market's guidelines, to generate a new type of application, which gives more space to the issues relating with ergonomics and energy efficiency, reinterpreted with new linguistic codes. Design, as a tool for innovation, match a fundamental fact: the man, understood as service and product user. The man is at the center of the project [1]; a good designer should always consider how to transform man into potential user, reinterpreting the existing, through a new awareness. In this specific case this belief bases its principles in the search for sustainable solutions. In this context, "The design is intended and lived as a flexible instrument for the interpretation of contemporary social transformations"[1] [2]. Design must be able to express together empiricism, rationality, imagination and control. The project is the projection of an intuition that its able to open new horizons and different perspectives, to trigger the change of the contemporary scenario processes, through figuration of future scenarios, through the Design Orienting Scenario activities. (Figure 3).

[1] Marano A. (2004) Design e ambiente, la valorizzazione del territorio tra storia umana e natura, Poli.Design, Milano.

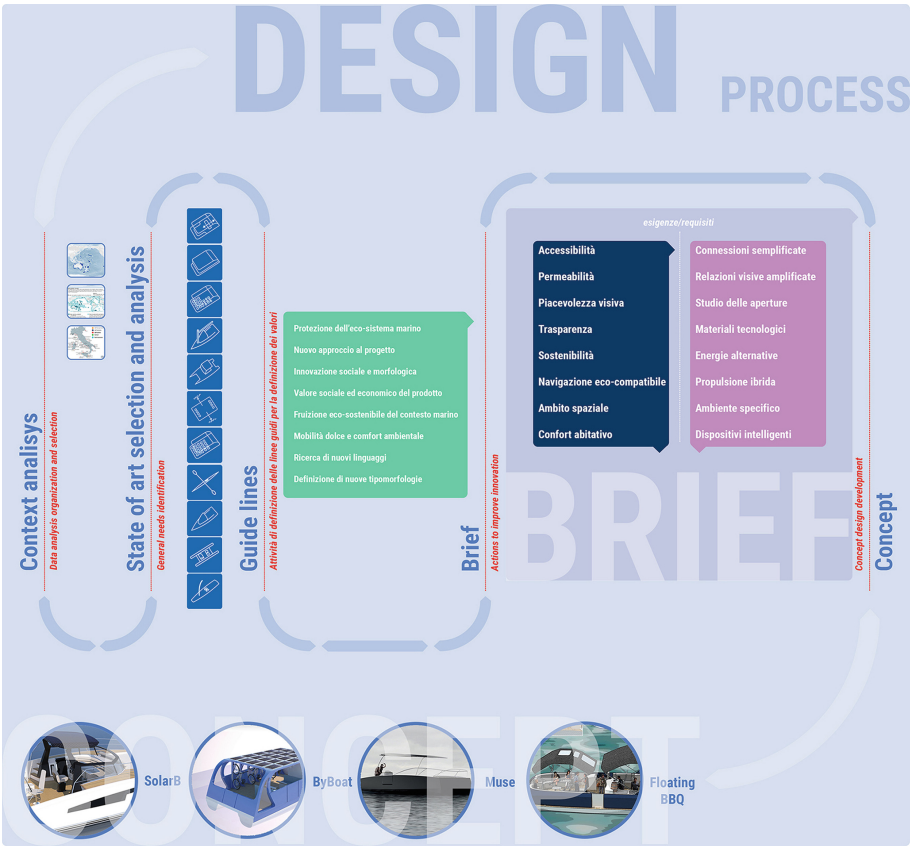


Fig. 3. Design process description

4 Guidelines

The analysis of the context has highlighted a series of key factors that clearly can express the strategies to be taken to investigate the issues relating the small pleasure craft. The conscious usage of entertainment resources results fundamental. The main purpose is to orient all possible targets (users and tourists) by the adoption of decentralized target touristic services. The Pescara Province inland already has an array of recreational character and offers services aimed at the enhancement of river and coastal natural resources. All the essential factors can be considered opportunities, relating to their spatial and morphological characteristics, suitables to suggest new design oriented scenarios. All the characters and factors should be investigate and considered like opportunities to improve, or obstacle to reduce the user experience. For this reason, the *Costa dei Trabocchi* assumes a strategic importance, considering the existing tourist flow. However, the achievement of all touristic points of interest needs a considerable travel time, so the solution proposed should consider all these factors. The seaside cost

its full of recreational activities and opportunities for yachting development. Touristic companies demonstrate a big interest. For this reason all the concept proposed must be able to respond to different modalities of marine contest's usage. Small craft must be able to respond to these request. The touristic offer, relating to the resort activities, turn out to be a potential resource. Small craft seems to be a great opportunity to introduce the idea of floating hotel, because off more than one potential user expressed needing relating with these specific requirements. At the same time the inland an river navigation seems itself an attraction to improve. For this reason, the vessels must respond to the technical and security features to make river navigation possible. Another activity in this area is one that consists of taking customers from the main harbor and simply take them further out to let them enjoy the entertainment sea activities. All this activities require a diversified management and qualified services connected to them. This research presents the first results of the work on the topic of sustainable small crafts for inshore and inland navigation. The market investigation has revealed how the most suitable craft, suited to play this role, can offer a diversified service which is compatible with the different local conditions and different fruition modalities by a specific user. At the same time, the focus revealed how these products can ensure minimum standard of accessibility and management services on-board. The issues of the inshore and inland slow mobility by waterways has been investigated with great attention. Particular attention paid to the external and internal livability of the vessels, exploring ergonomic and technological solutions, like: continuous surfaces, flowing lines, deck organization to host different activities; flexible solutions designed to improve the whole space efficiency. All these aspects have been investigated in compliance with the minimum requirements for accessibility and habitability.

4.1 ByBoat

ByBoat (Bike-Boat) is the first study case presented. This concept is designed as a service boat. Its purpose is to facilitate the intermodal mobility, and to implement the tourism industry along coasts and inland reserves. ByBoat can carry up to a maximum of eight users with the following equipment for cyclists. The entire housing system is organized with several panoramic dinettes. The deck is completely sheltered by a built-in cover. The cover consists in an integrated system of solar cells [3]. This solution helps to increase the energy reserve, which is necessary for the management of on-board services, activities and hybrid-electric motorization. A good level of accessibility on board is guaranteed by the introduction of transformer platforms positioned on the three sides of the boat. From the functional point of view, the access size is over dimensioned because of the particular nature of the means that accompany the users: the bicycle (Fig. 4).

Fig. 4. ByBoat, concept layout and views

4.2 Floating BBQ

The second study case presented is a Floating Barbecue Boat, designed to offer a traveling restaurant service. This small boat is designed for the agryfood business class, who has interest to invest into the restoration service. The purpose is to use the small craft like a promotional vehicle for their food products, and the solution proposed is the management of a mobile catering business. The typology proposed is the multihull, very similar to the pontoon boats used for inshore navigation. This choice allows to organize the entire deck surface, relating to the various planned activities on-board: from service management activities (relating to transportation and restoration), to the reception of the final user. The general arrangement of the deck is very simple and easy to identify: a large living area at the bow away (useful for restoration); a barbeque kitchen housed in the stern; a separate day toilet for public use; the control consolle integrated in the starboard side. Four transformers and retractable platforms provide access from all four sides of the boat. Two retractable solar roofs complete the technological solutions adopted. The roof contribute to shelter by the sun; the fabric curtains integrate photovoltaic films. This option can be useful to accumulate solar energy, together to ensure an optimal management of on-board services, to maintain a good supply of green energy for the entire day. All the solution adopted contribute to improve the efficiency of the small craft in different conditions: inshore navigation or inland navigation (Fig. 5).

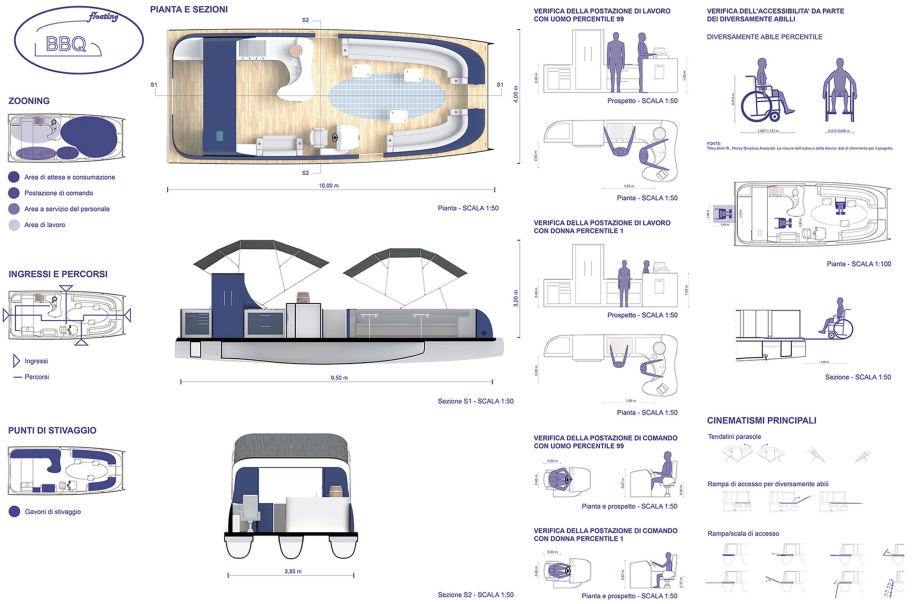


Fig. 5. Floating BBQ general arrangement

4.3 Muse

The third concept proposed, is a multifunctional small craft, designed both for inland navigation and longer cruisers. Muse can be adapted to different conditions. For example: amateur fishing practice, on-board restoration, simple tourist explorations. The main deck is provided with different equipments according to the different nature of the various typologies of entertainment. The multipurpose character can represent an innovative aspect, and contribute to reference the product inside the referred market. However, this small vessel refers to the specific target made of small and medium entrepreneurs who want to invest in tourism and entertainment, promoting services like: boat sharing companies; cultural and tourist attractions sponsored by private associations and touristic offices; hotels organized by offering hiking services or others activites. The product is characterized for its simplicity and for the standardization of useful modules; the modules contribute to perform the general organization of the deck, in relationship to the different activities that may take place on board. Also in this third case, the idea of a fully accessible deck has been studied with particular attention, by the adoption of transforming and retractable platforms of different dimensions. The product can be considered minimal, sustainable (by the adoption of a solar roof), eco-friendly, and easily adaptable with the development of a slow mobility system aimed at enhancing the territorial resources. The product can be used by a wide range of users: families, groups of friends, all user who love yachting in every condition (Fig. 6).

Fig. 6. Muse multipurpose flexible layouts description

4.4 Solar B

Solar B is a pleasurable small opened craft, designed to carry nautical tourism in the different dimensions of daily trips or short cruises. The small craft can host from 2 to 4 users. Particular attention was paid to all the issues related to the livable comfort, sustainable mobility, and security management on-board. The choice of the multihull typological solution allowed to adopt morphological, technological, functional, and distributive solutions which take into account specific requirements linked to the following activities: accessibility (by the insertion of a retractable and telescopic gangway, a staircase for ascending and descending by/to sea, the inclusion of an integrated transformer platform in the main deck to ensure accessibility to extend user classes, stairs and walker rounds to connect stern and bow. Both solutions have been integrated on the left and right transom, under the main deck level); management of navigation and on-board services (low immersion; hybrid propulsion system for inshore navigation; silent and free emission; solar roof and solar sunbathe to guarantee energy's continuous rechargeable cycle as useful solution to the management of on-board services; center cockpit with a panoramic view); High habitability grade for short cruises (a convertible dinette positioned in the main deck; a small day-toilet, a double cabin, a kitchen with the indispensable minimum nautical equipment; all these equipments are integrated in the lower deck by the left and right hulls); heliotherapic activities (organization of a sunbath area in the deck solar bow); amateur fishing and sport fishing (organization of an aft area completely dedicated to fishing activities management); Carbon fiber and fiberglass materials are used to built the hull structures and the deck

house superstructures. The small pleasure craft is constructively simple to implement, because thera are no complex geometry (Fig. 7).

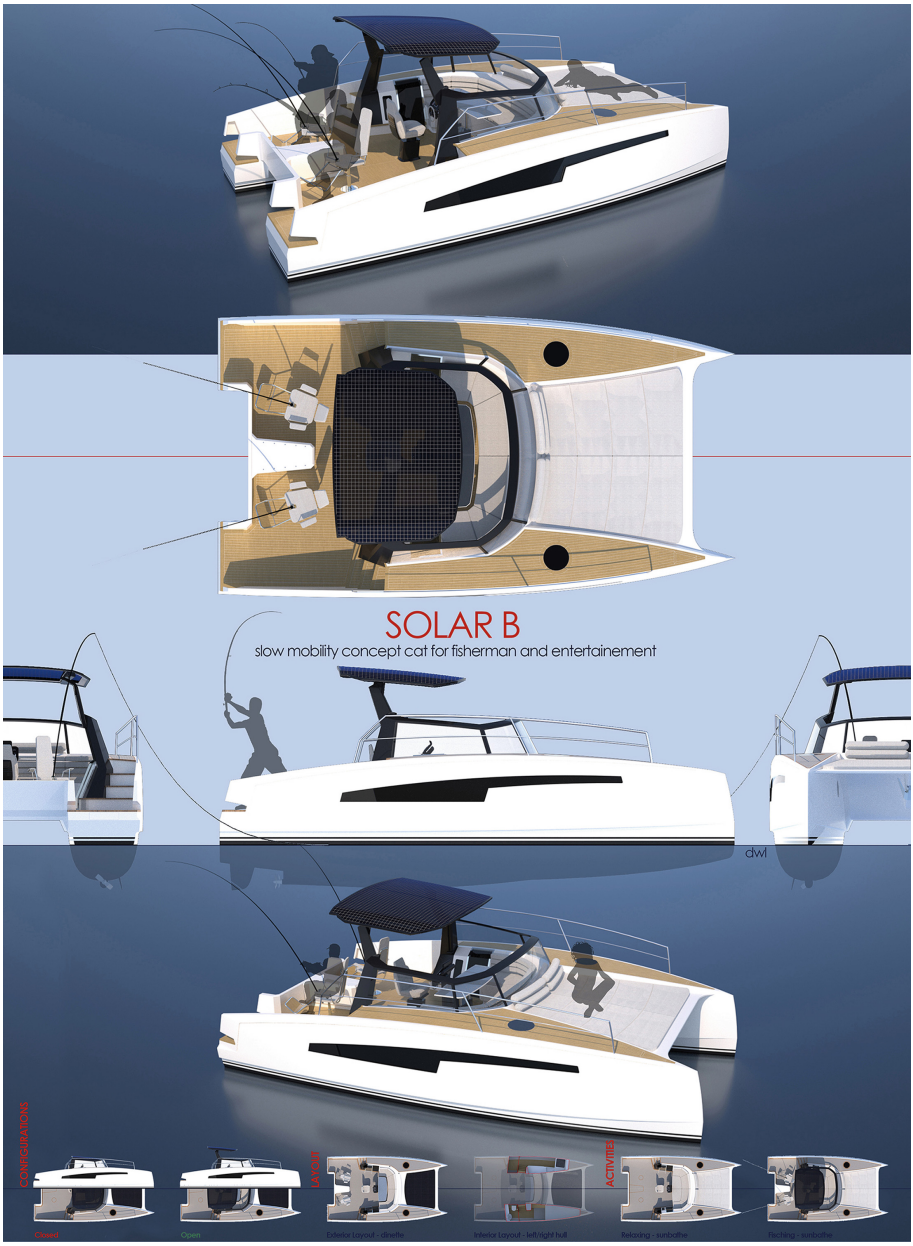


Fig. 7. Solar B, exterior concept description

Each on-board function is clearly defined and readable (access systems, operations management, internal and external connections, stern/bow walkeround connection, different activities organization (sailing, fishing, relaxing). The adoption of the transforming access platform allows the user to get on and off from to the small craf to the dock; at the same time this solution ensures to get on and off by the sea or into the sea. The large outdoor dinette, can turn into a comfortable spare bed (Fig. 8).

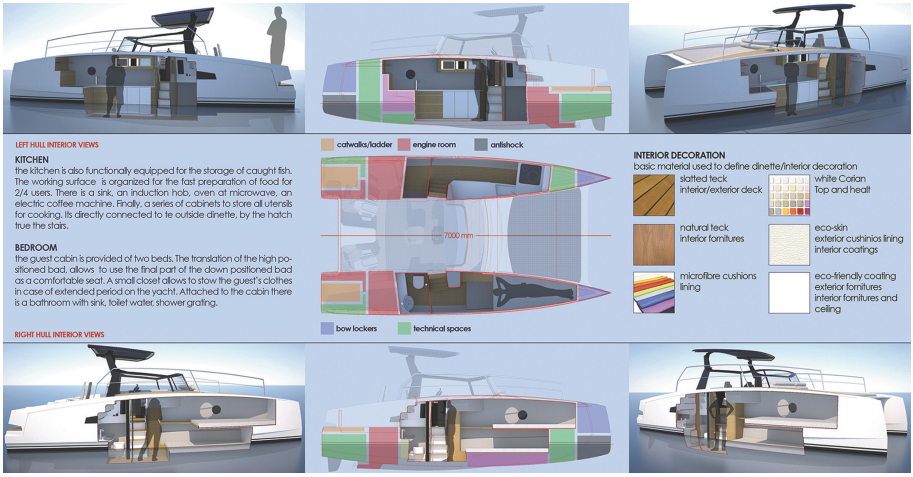


Fig. 8. Solar B, interior layout description

From the architectural point of view, the concept is characterized by the relationship between the strength of the hull lines and the lightness of the small and fully transparent deckhouse. The deck house its opened by the aft, and characterized by the movable solar roof. The solution of the movable roof allows multiple categories of users to use the yacht in any condition. At the same time, the adoption of the closed position of the roof let the users to safety navigate in adverse weather conditions. For the hybrid-electric propulsion and on-board services management, technological and smart solutions have been studied by the adoption of solar cells integrated into the solar roof, electrocroming glazing for the deck house openings to the outside [4] and domotics to ensure an appropriate level of minimum living comfort. These technical solutions have already tested in the marine field, and used on different levels of intervention.

5 Conclusions

The main purpose of the analysis carried out up to this moment, was to highlight the application of a series of design solutions for small boats appointed to the coastal and inland navigation. For example: correct solutions for spatial organization; the choice of particular types of small crafts designed for coastal and internal navigation; small crafts for the transport of users; small crafts for mixed transport of users and things; more in general, small boats designed like interfaces for the development of the tourism service

industry, or in the specific, for the Abruzzo's touristic service; last but not least, the adoption of correct aesthetic and morphological solutions related to the nature of the artifacts. These first suggestions can strongly represent ideas of morphological and functional originality. At the same time these solutions could generate great interest by all maritime filed. The described designed solutions could help to define new possible scenarios and forms of marine environments's usage.

Acknowledgements. This paper refers to the first results obtained by the dossier "River and coastal slow mobility", related to the "Slow Mobility in the Abruzzo Region"research activated by the Architectural Department of d'Annunzio University from Chieti-Pescara, according to Abruzzo Region Governance; the dossier is carring out by the second unit of Industrial design (Proff. A. Vallicelli, E. Baldassarri, G. Di Bucchianico, M. Di Nicolantonio; Ph.D. J. Lagatta; Student: A. Biagi). The first results report about the first step of a research relating the competitive development for the Abruzzo Region; this first part is attributed to Prof. M. Di Nicolantonio. All the depicted marine products refer to: "ByBoat", student M. Iaquinta; "Floating bbq", P. G. Russo; "Muse", students P. Stella, C. Grimaldi, S. Novak; "Solar B, Slow Mobility Concept Cat for Fisherman and Entertainement", author: M. Di Nicolantonio, (project placed the second prize at the 2016 National Competition Premio di Design Nautico). The various paragraphs of the present paper can be considered the consequence of a common discussion and a collective review among author. In Particular, the writing of the various paragraphs can be Attributed to: Prof. Massimo Di Nicolantonio.

References

1. Di Bucchianico, G., Vallicelli, A.: User-centered approach for sailing yacht design. In: Karwowski, W., Soares, M.M., Stanton, N.A. (eds.) Human Factors and Ergonomics in Consumer product Design Uses and Applications. CRC Press, Taylor & Francis Group, Boca Raton (2011)
2. Marano, A.: Design e ambiente. La valorizzazione del territorio tra storia umana e natura. POLI.design, Milano (2004)
3. Rossi, E., Di Bucchianico, G., Di Nicolantonio, M.: 10 Meters daysailer "for all" sustainable technological solutions for easy navigation. In: Vink, P. (ed.) Advances in Social and Organizational Factors. CRC Press (Taylor & Francis Group), Boca Raton (2012)
4. Di Nicolantonio, M., Lagatta, J., Vallicelli, A.: 80 feet sustainable motoryacht: Technological solutions concept of the living spaces on board. In: Ahram, T., Karwowski, W., Schmorrow, D. (eds.) Procedia Manifacturing, "6th International Conference on Applied Human Factors and Ergonomics AHFE 2015 and the Affiliated Conferences 2015. Elsevier (2015)

"A-Shaped" Mast Sailing Yacht: Technological Solution for Easy Navigation

Jessica Lagatta(✉) and Andrea Vallicelli

Architectural Department, G. d'Annunzio University, Viale Pindaro 42, 65127 Pescara, Italy
{jessica.lagatta,andrea.vallicelli}@unich.it

Abstract. Most of the modern sailboats are armed with a sloop sail plan. This type of plan has some aerodynamic and operational drawbacks. The central mast position and its equipment create vortices on the sails (decreasing the efficiency) and, in particular, considerably reduce the habitability of both the deck and interior. In this paper we present a solution to this problems through the concept of a 35 ft yacht armed with a "simplified sail plan". It is an "A-shaped" mast placed in the stern area of the yacht that has been considered in single jib and double jib configuration. This sail plan represents an opportunity of great interest for a *Design for All* approach, because it follows some of it fundamental principles, for example: to enable prehensility, to require minimal efforts and ability to the users, to use of technological solution (also innovative) to meet the requirements and skills of the user, etc.

Keywords: A-shaped mast · Easy navigation · Design for inclusion · Design for all

1 Introduction

Over the last fifty years, the strong expansion in the nautical sector and the subsequent increase in the number of users led – within the sector of yachting – to the need of creating boats for a larger and more diverse type of consumers. At the same time, the growing awareness of social inclusion within the various sports has allowed people with disabilities to have access to activities they could not practise before. These include sailing sports.

Nowadays, designing boats for a broader audience – that is of different age and with different tastes, cultural habits and physical abilities – could represent a competitive solution for the nautical market, from both an economic and social point of view. Thus, the awareness of the existence of a market with non-standard users, partially different from those to whom the design of sailing boats was addressed until a few decades ago, leads to the need to offer new design solutions that actively address this new complexity and the resulting added value [1].

DfA (Design for All), defined by the Stockholm declaration as "design for human diversity, social inclusion and equality" [2], could represent the most appropriate designing approach to meet one of the actual several demands of the market. That is to make increasingly inclusive boats, not segregating ones and able to actively meet the requirements and needs of all kinds of sailors.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_18

However, the development of an inclusive design of a sailboat requires the management of several aspects that are not only related to ergonomics (mainly with respect to the deck plan), but also to mechanics-performance and aerodynamics. As a matter of fact, a sailboat can be defined as a complex dynamic system that interacts with two fluids (air and water) in which the developed forces, which cause its movement, partly depend from the movement itself. Therefore, the design research should mainly focus on the fluid-dynamic issues connected to particular features of the propelling system. As a matter of fact, the latter can play a very significant role in the quality of comfort and safety during navigation.

In short, it can be said that there are three design aspects that are peculiar of the wind-based propelling system and of the naval architecture: (1) the sail plan (2) the shapes of the hull (3) the centreboard plan. This paper takes into particular consideration the first and the third aspects. Some typical design solutions (for devices and systems) for these aspects – already used in other nautical sectors (yachting and competitions) – have been analysed and some suggestions have been made for their application aimed at inclusive design, highlighting the potential benefits.

In particular, part of this paper will illustrate the research (a collaboration between the Department of Mechanics of the Polytechnic University of Milan and the Department of Architecture of Pescara) that allowed to test – in the wind tunnel – the performance of a traditional sloop sail plan in comparison with a sail plan with an "A-shaped" aft mast, with both one and two jibs. The final part of the paper will illustrate a concept-design related to a boat designed according to an inclusive approach, with a "simplified sail plan" and a canting keel.

2 The Problems of the "Traditional" Sloop Sail Plan

Most of the modern small and medium sailboats are armed with a sloop sail plan. The sloop is a type of sail equipment characterised by a single mast, located in the centre of the boat (generally between 35% to 45% of the LWL from the bow) and equipped with a mainsail, ridden on the boom and on the mast itself, and a foresail: jib or genoa ridden on the stay. In addition, there are the sails for running, such as the spinnaker and gennaker, which have no head rope.

However, the sloop sail plan, very widespread because of its simplicity, has some operational and aerodynamic shortcomings such as the following:

- the mast at the centre of the boat significantly reduces the practicability of the deck and the underdeck, affected by the presence of reinforcing structures and pillars;
- the boom and the truck of the mainsail sheet, as well as being in the way on the deck, can represent a hazard for inattentive passengers;
- the several adjustments to be made on the mast and sail require the installation of many expensive equipment such as rails, vangs, traveller, etc.;
- the reduction of the mainsail surface with the boom is always a very difficult operation (also when it can be furled) for a single person, particularly when the weather is not good;

- the central position of the mast and its equipment (spreaders, rigging, etc.) with respect to the sails creates vortexes that affect both the air flow on the mainsail and the air slot that is formed between the mainsail and the genoa, with a negative effect on the overall aerodynamic performance.

Obviously, this type of sail plan also creates several problems to the ergonomics, usability and management of the boat for all kinds of users.

The first step to solve these problems was to take into consideration some design proposals that, over the history, have tried to address these issues (particularly related to aerodynamics) in an innovative way. This led to the creation of a model of "boat with simplified sail plan" (patented by Vallicelli Yacht Design & Co. and others), equipped with an "A-shaped" mast that can use one or two jibs, as described in detail below.

3 "A-Shaped" Mast. Short Description of the State of the Art

As of today, there are many design proposals, which will be referred to below, that have tried to address the above mentioned problems, particularly those related to aerodynamics.

As mentioned above, the most interesting solution – and the one with the best potential for development – is the arrangement of a sail plan with an "A-shaped" mast in the stern of the boat. With this solution, the mainsail is removed and the total sail surface is distributed on two jibs (in the two-sail configuration) or on one jib (in the single-sail configuration), providing some operational and aerodynamic benefits, such as:

- no harmful interference of the mast on the sails;
- no boom and more distance from the main traveller from the deck;
- simple operations for reducing and lowering the sails;
- less equipment on the deck;
- simple structures for fixing the mast to the hull (reinforced frames, pillars, keelson, chain plates, etc.);
- more usable deck surface.

Please note that the position of the mast in the stern of the boat significantly reduces the back stay angle in comparison with a sloop. This causes more compression effort for the "A-shaped" mast. However, this problem is mitigated by the fact that the stressed bars are two and they can also be sized with length-wise extended sections that, if well-designed, can also offer less aerodynamic resistance. Moreover, some more experiments on the engineering field may quickly develop this kind of mast, removing the shortcomings due to the increased compression loads.

In the past, there have been many attempts to reduce the problems of aerodynamic interference on the sails, both for the sloop and for the sail plans with more masts.

The historical origins of the "bipod" or "A-shaped" mast (as we prefer to call it) date back to at least 2500–2000 BC. Portrayals of Egyptian river boats show this kind of mast, which was used because it had a lighter load on the hull. The double shaft of the mast allowed to distribute the compression loads on the sides of those boats – which had a flat bottom and no central keel and were built with lightweight materials like papyrus.

In more recent times, and in the yachting sector, we can find several proposals, more or less original, developed at different degrees of experimental verification. One of the most interesting is the boat Amoco Procyon (1991, USA) by O. Harken and B. Chance, built with an "A-shaped" mast (patented as Bi-POD mast and produced by E. Hall), equipped with a central stay where the mainsail is fixed as if it was a jib. The Bi-POD mast reduces the interference of the mast on the sail entrance side, although it does not removes it completely. Actually, this kind of mast was already featured in some 1930s designs by M. Curry and then in the design (by N. and G. Sironi) that won the Philips competition in 1971. Other 1970s experiments, such as Luna or Vendredi 13 by D. Carter, included schooner-type sail plans only with jibs, no mainsail, where the masts where located aft of the respective jibs. Also the idea, more interesting, of solving the aerodynamic problem using an "A-shaped" mast in the stern of the boat is quite old. There are also some pioneering experiments from the 1950s, such as the catamaran Marara by C.O. Walker, and several patents that are very similar (see: I. Brandin, 1992, SE; K.D. Lehmann 1994, DE; C. Atkinson, P.C. Sewell, K.L. Hamson 1987 CA; R. Wieland 1997, DE; H. Erler 2000, DE).

According to the current available information, all these ideas – conceptually equal –are still hypotheses, never verified with reliable experiments.

Thus, at present, the research performed by the Department of Mechanics of the Polytechnic University of Milan and the Department of Architecture of Pescara (presented in 2008 in Auckland, see bibliographical reference), partly included below, is the only scientific work for the verification of the performance of this kind of sail plan [3].

4 The Experiment in the Wind Tunnel

The experimental work performed, as said, in the wind tunnel of the Polytechnic University of Milan was set as a comparison between the performance of a traditional sail plan (sloop) with that of a "simplified sail plan" ("A-shaped"), of equal sizes, both with one sail and with two sails. In order to perform the comparison, two models (scale 1:10) were created: one for the boat with sloop sail plan and another for the boat with "A-shaped" mast (in the stern). For the first model, the sail plan of a traditional cruiser-racer was adopted: the Comet '51 s (design by Vallicelli Yacht Design & Co., winner of the Italian championship IMS 2007). For the second model, on the other hand, a double configuration was considered: with a single jib and with a double jib (overlapped and not overlapped). The models were made complete with mast, rigging and sails, with the hull cut above the floating line and fixed on the rotating plate of the wind tunnel (Fig. 1).

The data were acquired through the National Instrument Data Acquisition Boards and transcribed in accordance with the Matlab standard. The data acquisition software calculated the forces and moments using the dynamometer calibration matrix.

Moreover, in order to connect the readings of the measurements of the wind force and the shape of the sail, and in order to provide the data in the form of input for CFD calculations, a photogrammetric measurement system was developed in-house for reconstructing the profiles during the performance of the tests.

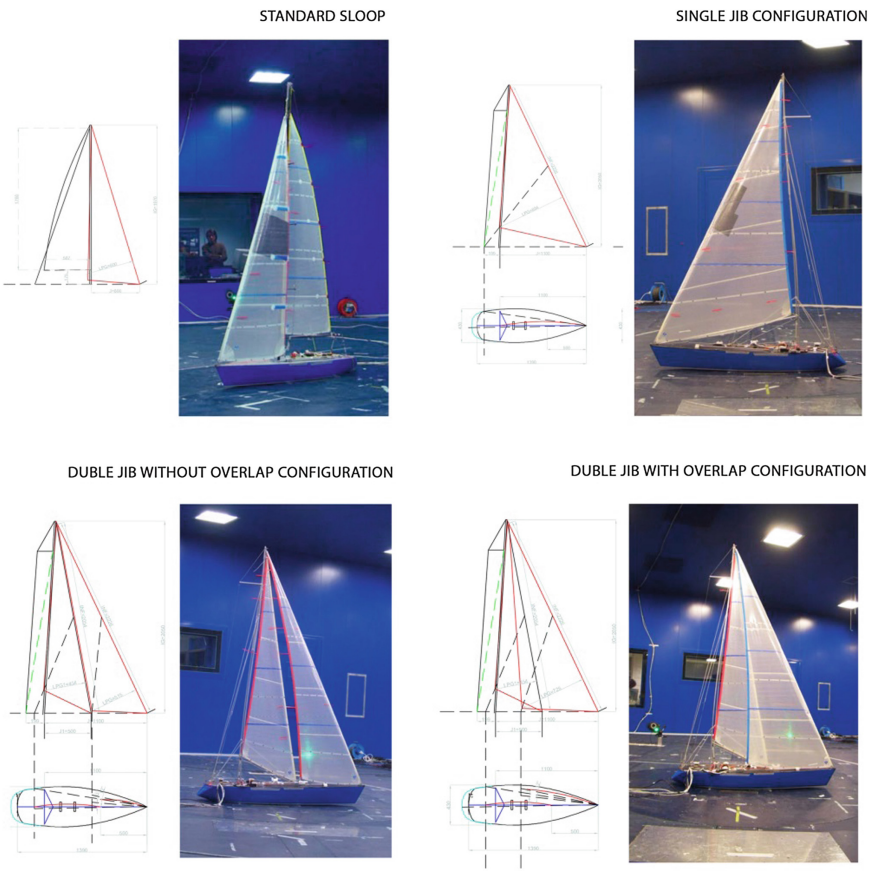


Fig. 1. Test in wind tunnel.

The tests were conducted with the boat in vertical position. Furthermore, in order to have the most complete upwind angle range possible, measurements were made also with apparent wind angles of 22°, 27°, 32° and 42°. In all of the four different solutions analysed (fractioned sloop, single-jib, overlapping double-jib, non-overlapping double-jib), the correct cut of the sails and their ideal adjustment for the best efficiency was calculated.

The data resulting from the tests clearly showed that the double-jib "simplified sail plan" configuration (overlapping or not) ensures a better performance compared to the traditional sloop sail plan. These results can be seen in the graphs related to the VPP made according to the data from the tunnel tests.

5 Case Study (Combination of "A-Shaped" Mast and Canting Keel)

From a design point of view, the use of the "A-shaped" mast and the simplified sail plan can contribute to improving not only the sailing performance but also the ergonomic

aspects of the boat. The shape of the mast, its position and the simplification of the equipment offer designing potential in terms of: usability of the deck (more spacious and free from fittings such as the mainsail truck); habitability of the underdeck (lack of central structural pillars); simplification of the operations and easier access on board.

With this respect, the case study presented below refers to the concept-design of a 35' sailing yacht, where we can find the following effective combination: an evolutionary propelling system such as the "simplified sail plan" (furling and self-taking single-jib) and a particular centreboard plan like the canting keel and, of course, the architecture of the associated hull.

The study aims to show that the combination of technically innovative solutions can become an essential tool for those who want to develop designs in the framework of easy sailing, inclusive design and particularly Design for All that, in the nautical field, aims to direct the design towards new models of safe, easy and accessible boats, while not neglecting the necessary aesthetic appeal for the commercial success.

Therefore, with respect to this specific case study, here is an analysis and a summary of the several benefits offered by the "simplified sail plan" (Fig. 2):

- the sails, without a boom, can be furled from the cockpit;
- the changing of position of the sails, taking and gybing, can be made automatically by the track fixed on the beam that connects the two vertical shafts of the "A-shaped" mast. Thus, the traveller – being above the cockpit – does not represent a hazard for the crew during the operations;
- the benefits in terms of aerodynamic efficiency, compared to a traditional sail plan, allow broader design options that can be aimed to achieve: "simply" a higher speed of the boat; reduced sail surface with the same performance of a traditional system; lighter deck equipment; reduced physical efforts for operations; cost control;
- a simplification of the whole operations system allows to achieve an appropriate ergonomic level for a user-friendly experience;
- facilitating the adoption of an open stern, free from manoeuvring equipment, allows for design solutions aimed to increase the degree of accessibility and to favour the presence of diverse users on board, who can also contribute to the management of the navigation activities if necessary.

All the above highlight the obvious benefits in terms of safety (no risk of collision between users and boom), handling (low physical efforts for manoeuvres), propelling efficiency and, naturally, ergonomics.

As shown in this concept-design (35' sailing yacht), in order to meet the requirement of inclusive design (comfort and accessibility), good use was made of the performance potentials of this particular type of aerodynamic propelling system. Moreover, the comfort requirements have been met by working on the shape of the hull and its accessories, in order to optimize the hydrodynamic efficiency and stability of the boat.

As we all know, a hull can be designed in many different ways, but they all shall be aimed at an effective performance of hydrostatic and hydrodynamic tasks. Its proportions, for example expressed in the length-width relationship, define very different

Fig. 2. Concept '35 mailing yacht, studio Vallicelli & Co.

stability behaviours (with other architectural parameters being equal), that can significantly change the ability to offset the heeling effect produced by the forces generated by the action of the wind on the sails.

Thus, in order to improve the comfort during the navigation in a sailing yacht, one of the design goals consists in finding solutions that reduce the transversal heeling.

With this respect, if we limited to considering the righting parameters, designing a very wide hull could be enough. However, this solution would lead to increasing (with other parameters being equal) the resistance to forward movement. Therefore, we would have a more stable but slower boat. Hence, in order to have a desirable speed, we would need to increase the propelling power (that is the sail surface) and this would obviously result in an increase in the heeling moment. In order to offset the latter, we should redesign a slightly wider hull, triggering an endless cycle.

Obviously this loop (in the search for balance) must find a designing solution. A possible one could be, for example, relating this problem with other aspects of the design. One of these is represented by the centreboard, and particularly by ballasted centreboard.

There are several types of hull appendixes that are responsible for both reducing the leeway of the boat and, when ballasted, contribute to its stability by lowering its overall centre of gravity.

Among those that, over the past years, have been most successful in terms of performance (in the cruising and racing sector), there is the canting keel (Fig. 3).

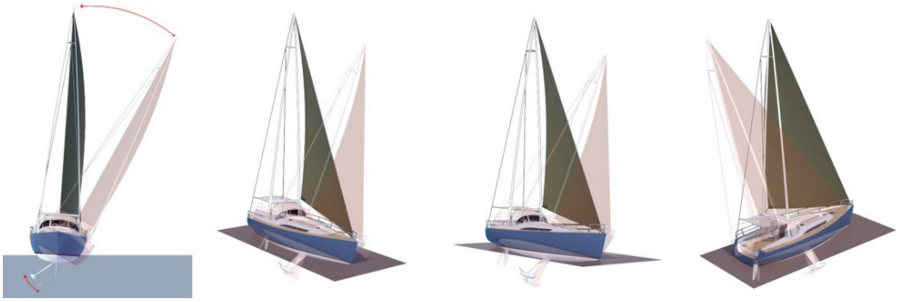


Fig. 3. Concept '35 sailing yacht, use of the canting keel.

This centreboard is fixed to the hull, through a hinge, and equipped with a ballast at its end. Its particular feature consists in that it rotates transversally by means of a mechanical device that is usually controlled by an hydraulic system. Its movement allows to move the ballast centre of gravity to upwind, producing very long lever arms that generate significant righting moments even with moderate weight, thus allowing navigation with minimal or even no heeling.

So, the increase in the transversal stability achieved by an appropriate work on the hull width or moving ballast, offers the possibility of reducing the overall weight of the boat, with clear benefits in terms of ergonomic and nautical performances (reduction of the physical efforts required for manoeuvring).

6 Conclusions

In conclusion, this case study shows that design innovation can be originated by the adoption of solutions resulting from experimental research also performed in different fields. In this specific case, we have seen that the synergic combination of an innovative sail plan (designed for easy sailing) and a canting keel (designed for high-performance racers) can be an interesting solution also for inclusive design (specifically DfA) in the nautical sector.

Acknowledgments. This paper is the result of the combination of two pieces of research. The first one relates to the experiments on the "A-shaped" mast performed thanks to a collaboration between the Department of Mechanics of the Polytechnic University of Milan and the Department of Architecture of Pescara. The second one relates to a research doctorate within the Department of Architecture of Pescara named: "Design for Inclusion in the field of sailing yacht. Design for All guidelines for small sailing boat" (Ph.D. Jessica Lagatta, Tutor Prof. Andrea Vallicelli, Co-tutor Prof. Giuseppe Di Bucchianico). The paragraphs in this text represent the result of a discussion between the authors. In particular, they can be attributed to: Ph.D. Jessica Lagatta (paragraphs 1, 2, 3), Prof. Andrea Vallicelli (abstract, paragraphs 4, 5, 6).

References

1. Lagatta, J., Di Nicolantonio, M., Vallicelli, A.: Design for inclusion. Differences and similarities between DfA and UD in the field of sailing yacht design. In: 6th International Conference on Applied Human Factors and Ergonomics AHFE 2015 and the Affiliated Conferences 2015. Procedia Manifacturing, Elsevier, Amsterdam (2015)
2. EIDD: Dichiarazione di Stoccolma. http://dfaeurope.eu/what-is-dfa/dfa-documents/the-eidd-stockholm-declaration-2004/ (2004)
3. Vallicelli, A., Fossati, F., Muggiasca, S., Viola, I., Martina, F., Nazareth, A.: Wind tunnel and CFD investigation of unconventional rigs. In: High Performance Yacht Design International Conference 3, Auckland (2008)

Assessing Sonar and Target Motion Analysis Stations in a Submarine Control Room Using Cognitive Work Analysis

Daniel Fay(✉), Neville A. Stanton, and Aaron P.J. Roberts

Human Factors Engineering, Transportation Research Group, Faculty of Engineering and the Environment, University of Southampton,
Boldrewood Innovation Campus, Burgess Road, Southampton SO16 7QF, UK
{D.T.Fay,N.Stanton,apr1c13}@soton.ac.uk

Abstract. As part of the Command Team Experimental Test-bed project a submarine control room simulator was constructed. Subject Matter Expert advice was sought throughout to ensure fidelity. However, no dedicated analysis was performed to validate the final result. To resolve this a comparison of Work Domain Analysis outputs was completed to identify levels of fidelity present within the simulator. Outputs were generated from a study at HMS Drakes Talisman, which were compared to previously completed outputs of the simulation facility.

Keywords: Cognitive work analysis · Abstraction Hierarchies · Submarine control room · Sonar · Target motion analysis

1 Introduction

The Command Team Experimental Test-bed (ComTET) project [1, 2] aims to identify and develop new ways of working for submarine command teams. To achieve this a submarine control room simulator was built [2] to conduct statistically high powered experiments exploring design concepts. Subject Matter Expert (SME) input was sought throughout the design, build, and experimentation stages, with SMEs indicating that the simulator was realistic. To formalise this endorsement of fidelity, a comparative examination for select User Interfaces (UIs) (Sonar and Target Motion Analysis (TMA)) using Cognitive Work Analysis (CWA) was performed.

CWA [3, 4] is a framework for modelling complex sociotechnical systems, emphasizing how work should be, as opposed to how it is, conducted [5]. Sociotechnical systems are those formed of technology and actors, human or otherwise. With several trained personnel operating complex technology at any given time, submarine control rooms can be defined as sociotechnical systems. CWA has been successfully used in a variety of domains [6, 7], including submarine control rooms [8, 9]. It has five stages, each for different purposes, although usage is contingent on analysis objectives. For UI analysis three stages are used [6, 10], Work Domain Analysis (WDA), Control Task Analysis (ConTA), and Social Organisation and Cooperation Analysis (SOCA). For this

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_19

comparison WDA was used. WDA documents objects within a system, their affordances, and their purposes [4, 11]. This information is represented by an Abstraction Hierarchy (AH), a diagram that splits a system into five horizontal levels of abstraction, ranging from the systems reasons for existing to the constituent physical objects. For each level text boxes (nodes) describe constituent components. Related nodes in adjacent levels are connected by 'Why-What-How' links, which show for any given node, 'what' it is, 'why' it exists, and 'how' it is implemented. With a completed hierarchy it is possible to assess how a system achieves specific goals, or why different aspects exist. This level of detail permits comparison of simulated systems to their real counterparts, with high levels of similarity indicative of fidelity.

Simulator fidelity is defined as the degree to which a simulator can replicate the actual environment [12], citing [13, 14]. It can be split into several types, each pertaining to a different contributing aspect. Of specific interest for ComTET is the replication of task and physical fidelity. Task fidelity is defined as the degree to which task replication exists [12], citing [15–17]. Physical fidelity is defined as the degree to which the physical environment feels like the real environment [12], citing [18]. These were chosen as they directly impact the results of studies performed in ComTET. If future ways of working, both digital and corporal, are to be accurately assessed it is vital that the same tasks are completed using the same physical objects as a real submarine control room.

Previously completed WDAs for ComTETs' Sonar and TMA [1] meant that only their counterparts in Talisman needed creating. Therefore, a study was planned and realised to elicit the required information from SMEs at HMS Drakes Talisman trainer. The compiled WDAs from this study permitted comparison of these stations to derive a measure of fidelity. Nodes on equivalent levels of each AH were compared to understand similarities and differences. This comparison elucidated the degree of convergence present, with high similarities suggesting high fidelity. Similarly, divergence would indicate low fidelity, and a need to improve.

2 Method

2.1 Interviews

Interviews were held with five Royal Navy Submariners at HMS Drakes' Talisman trainer over three days. All had served at sea for an average of 856 days (SD = 412.25), and had experience operating Sonar or TMA. Some participants occupied supervisory roles, ensuring holistic analysis. Participation was voluntary.

Participants received drawing equipment (pen, pencil, ruler, rubber, and paper) to utilise freely for communicating concepts. Printed copies of AHs for ComTETs' Sonar and TMA stations were available. TMA functionality in ComTET is a Local Operations Plot (LOP). As the LOP is a component of the Submarine Combat Management System (SMCS) in Talisman, SMCS was compared to CoMTETs TMA. Interviews were recorded using an Olympus WS-831 digital voice recorder, with an additional external microphone. Notes were recorded by interviewers using a notepad and pen.

For the interview, participants were asked to describe their job role, describe their console, and explain its usage. They were encouraged to create supplementary material

to aid interviewer understanding and construction of the AH. Each interview was scheduled to last two hours, although this was fluid to allow more time for information elicitation, or less time for duty requirements. The time was provisioned as per Table 1.

Table 1. Interview schedule

Activity	Minutes	Description
Introduction	10	Description research and objectives
Role scoping	30	Interviewees role defined and scoped
Sketching	30	Sketching of console interface
Means-ends analysis	30	Creation of an abstraction hierarchy
Debrief	10	Quick debrief with chance for questions
Slack time	10	Extra time, for any activity above

Before starting, participants were provided with an information pack containing an information sheet, interview schedule, a demographic questionnaire and a consent form. Upon signing to signify informed consent, and filling out the questionnaire, the interview began. Interviewers introduced themselves to the participants, and the aims of the study were reiterated. Detail about ComTET was provided for context. Interviewer notes were taken throughout to assist with AH construction.

Next, participants were asked to describe their role within the command team, and specify objectives that should be met. For each objective, criteria were elicited that would indicate successful, or otherwise, performance. Participants were allowed to answer freely, although were asked to prefix their objectives with ‘maximise’ or ‘minimise’ to assist with defining clear objectives.

After their job role and objectives had been documented, participants were asked to sketch various interfaces from their console. They were asked to add as much detail as they could remember, however they were not expected to create exact screen replicas. They could use any drawing equipment available. During sketching they were encouraged to speak aloud to match drawings to explanations. The next stage, means-ends analysis, was performed either in parallel or subsequently, depending on participant preference. Participants were asked what functionality each individual aspect of their sketched interfaces, and why it does so. More sketches could be drawn if needed, and these were explored in the same manner. For pivotal processes, participants were asked to write the process down using lined paper whilst explaining each point for clarification.

Once participants had finished explaining how their station looks and works, their attention was drawn to the corresponding ComTET AH diagram. It was explained, but they were not required to absorb all shown information. They were asked to identify any differences between their role and its ComTET equivalent. They were asked to disregard the bottom two levels, as the interfaces were different hence generating different Physical Objectives and Object-Related Processes nodes. However, if they had feedback it was noted and incorporated.

Finally, the participants were debriefed. The interviews purpose and what had been covered were reiterated. Participants were asked if they had any questions, and then thanked for their time, concluding the interview.

2.2 Observations

To ensure holistic coverage of stations and to mitigate unintentional omissions from interviews, a series of observations in the trainer were also conducted. Their purpose was to physically observe each console, and to consolidate experimenter knowledge. Personnel volunteered their free time during simulation runs to either provide a running commentary of interface usage, or to demonstrate their usage, answering questions in situ.

Experimenters used a notepad and pen to capture elicited information. For running commentaries, personnel stood behind working operators during running scenarios and vocalised key points of action. Operators were not interrupted as they were training. This procedure enabled experimenters to place understanding of the system into the context of actual usage. For station demonstrations, an operator would talk through specific procedures of importance, navigate through screens so detail sketches as well as notes could be made, and guide experimenters through practical use of the system. There were no fixed timings, nor topics, for observations. Rather they were held ad hoc in response to training demands on the submariners, personnel availability, and experimenter queries.

3 Comparison of Abstraction Hierarchies

Two AHs for Talisman were created, one each for Sonar and TMA. For each station, the nodes on each level were compared to their ComTET counterpart. The higher three levels will be discussed on an individual node basis, however the bottom two will be discussed in a combined holistic manner. This is due to the number of component nodes, and their close relationship.

3.1 Sonar

Functional Purpose. The overall goals of operators in both simulators were largely the same, although Talisman operators observed the additional goal of making effective usage of available Sonar arrays. This is due to physical restrictions placed on Talismans simulated arrays, which have an upper limit on how many different bearings may be listened to simultaneously. In ComTET this restriction does not exist, and is therefore is not taken into consideration for the station.

Values and Priority Measures. There exists significant variability at this level. Only two measures are directly comparable, '*Maximise Contact Detection*' and '*Maximise Signal Clarity*'. The ComTET measure of '*Maximise Known Contact Information*' is split into two for Talisman: '*Maximise Information Passage up the Command Chain*' and '*Maximise Rote Execution of Gun Drills*'. The three remaining ComTET measures, '*Minimise Counter Detection*', '*Maximise Safety of Operations*', and '*Maximise Mission Objective Completeness*', were not included in Talismans measures. This is because senior management take on these roles in Talisman, therefore absolving operators of a need to do so. However, this is not absolute, operators can feed information up the chain

of command if they deem it necessary to directly influence meeting these measures. The remaining measures for Talisman all regard the efficacy of operators when using their station and communicating data. There are no direct comparisons for ComTET, however efficacy is implicitly expected. It is hypothesised that the explicit inclusion of these measures for Talisman can be directly attributed to the values of the Royal Navy in being an effective defence organisation.

Purpose-Related Functions. The general function of Sonar is the Detection, Classification, Localisation, and Tracking (DCLT) of contacts. These are represented in both simulations by '*Detection of Contacts*', '*Classify Contacts*', '*Gain Speed Estimates on Contacts*', and '*Transfer Information About Contacts*'. In ComTET two additional general functions, '*Interpreting/Predicting Contact Actions*', exist. These are undertaken within Talisman, however typically not by Sonar operators, and as such are not included. A General Function specific to Talisman is '*Workflow Construction and Execution*', which relates to measures pertaining to efficiency and accuracy. The emphasis on these measures requires that general functionalities be available to directly support this.

Object-Related Processes/Physical Objects. There were largely common physical objects and associated affordances for both systems. This is due to the prominence of waterfall displays. Generally, all waterfalls consisted of '*Bearing Tape'*, '*Time Tape*', '*Waterfall Area*', '*Traces*', and '*Contact Markers*'. While the representations of these Physical Objects differed, they are largely comparable. Similarly, their Object-Related Functions are comparable. Aspects outside of the waterfalls exhibited aesthetic variability, but afforded similar, if not the same, functionality for both simulations.

3.2 TMA

Functional Purpose. Both ComTET and Talisman had the same overall purposes, to '*Generate a Tactical Picture*' and '*Supporting Higher Command Activities*'.

Values and Priority Measures. Again, there exists significant variability at the measures level, however, comparisons are still present. The combined ComTET measures of '*Maximising Knowledge of Contacts*' and '*Maximise Information Gathered About Contact*' can be equated to Talismans '*Maximise Information Transfer*' and '*Maximise Look Ahead Steps*'. These measures are about collecting as much information as possible for a contact, and ensuring that its future intentions are known. The ComTET measure '*Maximise Solution Integrity*' is present in Talisman as two separate measures, '*Maximise Solution Accuracy*', and '*Maximise Solution Working Time*'. As with Sonar, there were four measures included only in ComTET, as in Talisman they are performed by senior management: '*Maximise Safety of Operations*', '*Minimise Own-Ship Detection*', '*Maximise Mission Objective Completeness*', and '*Minimise Threats from the Enemy*'. Again, Talisman had explicit measures of effectiveness and accuracy. These were '*Maximise Effectiveness*', '*Minimise Errors*', and '*Minimise Communications from the Operations Officer*'.

Purpose-Related Functions. The general purpose of TMA is generating and maintaining the tactical picture. This is represented by the general functions '*Choose Contact*', '*Create Solution*', '*View Solutions*', '*Assess Solutions*', '*Ensure Solution Accuracy*', '*Refine Solutions*', '*Merge Solutions*', '*Split Solutions*', and '*Remove Solutions*' in both simulations. There were no simulation specific general functions.

Object-Related Processes/Physical Objects. Base components were found to exist in both simulations, such as '*Contact Markers*', '*Cuts*', and '*Own-ship'*. However, due to the extended capability of Talismans' SMCS, compared to ComTETs' TMA, similarities existed on a more atomic level than Sonar. This means that while Physical Objects represented similar objects, with similar Object-Related Process, how they were arranged lead to highly divergent interfaces. As expected, the closest similarities were observed comparing LOP implementations, with different areas of Talismans interface diverging to greater extents.

4 Discussion

In both simulations Sonar and TMA exist for the same reasons, but their measurements of success are different. It is thought this disparity is a result of their environment. Talisman is a Royal Navy facility, so its measures of success will be derived from the Navy, whereas ComTET is an experimental laboratory and will use an experiment orientated set of measures. The included measures in ComTET are ones that can be measured experimentally, and can be achieved within the projects scope. The equivalence of Functional Purposes, but with differing Value & Priority Measures, indicates that ComTETs stations suggests that they are appropriately included and evaluated in a contextually appropriate manner.

Usage of different interfaces invariably led to different Physical Objects, and subsequently different Object-Related Processes. Consequently, physical fidelity is reduced. However, an essential degree of fidelity is maintained by common interface aspects, such as Sonar waterfalls or TMA cuts. These interface commonalities are mainly core objects required to carry out Sonar and TMA, suggesting that a reduction of physical fidelity can be attributed to non-essential Physical Objects. Thus, while not exhibiting high physical fidelity overall due to different interfaces, a common core for essential components makes a strong case for medium physical fidelity in ComTET.

Congruent Purpose-Related Functions for both interfaces suggest high task fidelity, despite differing interface designs. These differences create different task completion pathways using the available Physical Objects and their affordances, but do not preclude completion of required tasks. While Object-Related Processes could be interpreted as smaller tasks, for example rotating a dial, they are not factored in determining task fidelity. This is because if one submarine used a dial and another used a slider to control a value for completing a task, it would not change the fundamental nature of the task. For the Royal Navy's varied fleet these differences will be inherently present, but submariners are trained to perform tasks irrespective of them. Therefore, disparate Object-Related Functionality does not detract from the high task fidelity exhibited in ComTET as a result of synergistic Purpose-Related Functionalities.

5 Conclusion

In order to ensure experimental validity regarding Sonar and TMA User Interfaces for ComTET a comparison was required with a submarine control room. The comparison control room chosen was Talisman at HMS Drake. Cognitive Work Analysis was chosen to perform the comparison as the method was created for analysing complex sociotechnical systems. Abstraction Hierarchies had already been completed for ComTET, but not for Talisman. Subject Matter Expert input was sought to construct these diagrams for comparison. A study was performed at Talisman with five currently serving submariners over three days. Participants were interviewed and observed to gain a holistic understanding of their stations User Interface. Once all interviews were completed, Abstraction Hierarchies for Sonar and TMA were created.

Comparison of Abstraction Hierarchies for ComTET and Talisman demonstrated that ComTET exhibits a high degree of task fidelity with a medium degree of physical fidelity for Sonar and TMA. This establishes confidence that findings and recommendations from ComTET would be similar to results from using Talisman, whose generalisability to real world applications is well established as a submarine command team training facility. Predicating real world applicability is vital, ensuring a tangible realisation pathway for outputs of the ComTET project.

Acknowledgements. This work was funded by the Defence Science and Technology Laboratory (Dstl) on behalf of the Ministry of Defence (MOD)

References

1. Fay, D., Stanton, N.A., Roberts, A.: Designing new interfaces for submarines: from cognitive work analysis to ecological interface design. In: Stanton, A.N., Landry, S., Di Bucchianico, G., Vallicelli, A. (eds.) Advances in Human Aspects of Transportation: Proceedings of the AHFE 2016 International Conference on Human Factors in Transportation, July 27-31, 2016, Walt Disney World®, Florida, USA, pp. 413–425. Springer, Florida (2017)
2. Roberts, A., Stanton, N., Fay, D.: The command team experimental test-bed stage 1: design and build of a submarine command room simulator. Procedia Manufact. **3**, 2800–2807 (2015). http://www.sciencedirect.com/science/article/pii/S2351978915007465
3. Rasmussen, J., Pejtersen, A.M., Goodstein, L.P.: Cognitive Systems Engineering. Wiley, Hoboken (1994)
4. Vicente, K.J.: Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. CRC Press, Mahwah (1999)
5. Stanton, N.A., Salmon, P.M., Rafferty, L.A., Walker, G.H., Baber, C., Jenkins, D.P.: Human Factors Methods: A Practical Guide for Engineering and Design. Ashgate, Farnham (2013)
6. Jenkins, D.P., Stanton, N.A., Salmon, P.M., Walker, G.H.: Cognitive Work Analysis: Coping With Complexity. Ashgate, Farnham (2009)
7. McIlroy, R.C., Stanton, N.A.: Ecological interface design two decades on: whatever happened to the SRK taxonomy? IEEE Trans. Hum. Mach. Syst. **45**, 148–150 (2014). http://ieeexplore.ieee.org/document/7042788/

8. Hunter, A., Hazen, M., Randall, T.: VICTORIA class submarine human-in-the-loop experimentation plan. In: Proceedings of the 19th International Command & Control Research & Technology Symposium (ICCRTS), pp. 17. National Technical Reports Library, Alexandria (2014)
9. Stanton, N.A., Bessell, K.: How a submarine returns to periscope depth: analysing complex socio-technical systems using cognitive work analysis. Appl. Ergon. **45**, 110–125 (2014)
10. Burns, C.M., Hajdukiewicz, J.: Ecological Interface Design. CRC Press, Boca Raton (2004)
11. Rasmussen, J.: The role of hierarchical knowledge representation in decisionmaking and system management. IEEE Trans. Syst. Man Cybernet. **15**, 234–243 (1985)
12. Hancock, P.A., Vincenzi, D.A., Wise, J.A., Mouloua, M.: Human Factors in Simulation and Training. CRC Press, Boca Raton (2008)
13. Gross, D., Freeman, R.: Measuring fidelity differentials in HLA simulations. In: The 1997 Fall Simulation Interoperability Workshop, Simulation Interoperability Standards Organization, Orlando, USA (1997)
14. Alessi, S.M.: Fidelity in the design of instructional simulations. J. Comput. Based Instr. **15**(2), 40–47 (1988)
15. Hughes, T., Rolek, E.: Human systems modeling: fidelity and validity: issues of human behavioral representation requirements development. In: Proceedings of the 35th Conference on Winter Simulation: Driving Innovation, Winter Simulation Conference, New Orleans, USA, pp. 976–982 (2003)
16. Zhang, B.: How to consider simulation fidelity and validity for an engineering simulator. Flight Simulation and Technologies. American Institute of Aeronautics and Astronautics, Reston (1993)
17. Roza, M., Voogd, J., van Gool, P.: Fidelity considerations for civil aviation distributed simulations. In: Modeling and Simulation Technologies Conference. American Institute of Aeronautics and Astronautics, Denver (2000)
18. Allen, J.A., Hays, R.T., Buffardi, L.C.: Maintenance training simulator fidelity and individual differences in transfer of training. Hum. Factors **28**, 497–509 (1986)

Visualization in Maritime Navigation: A Critical Review

Victor Conceição[1,2,3(✉)], Maria Beatriz Carmo[4], Joakim Dahlman[2,5], and Ana Ferreira[1]

[1] Instituto Dom Luiz, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisbon, Portugal
acferreira@fc.ul.pt
[2] Chalmers University of Technology, SE-412 96 Göteborg, Sweden
{placido, joakim.dahlman}@chalmers.se
[3] Escola Naval - CINAV, 2810-001 Almada, Portugal
[4] BioISI, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisbon, Portugal
mbcarmo@fc.ul.pt
[5] Naval Postgraduate School, Monterey, CA 93943, USA

Abstract. Innovations in navigation technologies are inducing subtle transformations in the navigator role and navigation processes. Thus, research on maritime navigation, entails the need to further study the designs of all technological artefacts that are or might be used in navigation, and specifically those that involve Human Computer Interaction. To the best of our knowledge, no systematic review about maritime navigation visualization has been carried out. The objective of this review is to fill this gap, by classifying new and innovative ways for the visualization of maritime navigation information. Recent findings on visual attention, offer new strategies and solutions for the representation of navigation information. Operators' visual attention can be largely improved through mitigation of cluttering effects or by guiding their attention. This review of visual attention also highlights the importance of systematic contextual research to understand all the interactions and processes that happen in the maritime navigation domain.

Keywords: Human-systems integration · Maritime navigation · Visual attention · Visualization

1 Introduction

As part of a research assessing navigational performance within the e-navigation concept, a review of the state of the art in information visualization was undertaken. Research over maritime navigation control processes, entails the need to study the developments or designs of all the technological artefacts that are or might be used, and specifically how humans interacts with them, as pointed out by Flach in [1]. Consequently, the main objective of this review is to identify innovative ways for the visualization of navigation information, within the ship domain.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_20

Among the theoretical views presented in the selected literature, this study takes into consideration the overarching elements, such as the navigation team, individuals and technologies, anticipated e-navigation solutions and the user's perspective. Therefore, it provides an important contribution to the existing research on technology-driven directions in the fields of visualization, computer graphics and cartography. Previous studies already addressed elements of this work, such as the comprehensive review on information visualization in the maritime domain, done by Davenport and Risley [2], listing suggestions about human factors, visualization requirements and collaborative visualization systems. They also identified new mapping solutions and display technologies, pointing out some issues such as lack of automation. Even though it reflects the use of similar data and information display, the study was not focused in maritime navigation, being mainly motivated by the requirements of maritime surveillance systems, without referencing specific design solutions. Another research applies spatial cognition in the design of a navigation system, but not in the maritime domain [3]. This work will then transfer the locus of attention to the shipborne systems supporting navigational functions.

Thus, the first part of this study focuses on the classification of pertinent topics drawn from the navigation processes and the context of maritime navigation. This conceptual analysis is followed by an explanation of the methodology and selection systematization criteria of the publications. Subsequently, the discussion embraces a synthesis of the findings related to visual attention and some reflections over their application in the light of the maritime navigation processes. Finally, we present derived conclusions, new research questions and some proposals for further exploration.

2 Conceptual Model of Maritime Navigation

Over the last two decades, the navigators have seen huge changes in the bridge settings, due to the extremely fast pace in the assimilation of new technologies. These changes are inducing silent transformations in the navigator role, which demands a review of the maritime navigation process [4–7]. More recently, automation, broad band communications and internet supported the probable operations of remote and unmanned vessels bringing additional challenges [8]. This trend in maritime navigation technology, highlighted in Fig. 1, has open-up a discussion on how to redefine the way ships are operated and designed. Despite the common acceptance of the advantages brought by digital information and computing technology, the overall complexity of technological support systems and regulatory framework for marine navigation emphasized the needs for new system design. This effect of amplification has been identified as an important driver of Cognitive System Engineering (CSE) development [9].

The role of the navigator is shifting to tasks more related with planning and monitoring, execution and surveillance are being undertaken by automated systems such as autopilot or automatic detection and tracking RADAR (ARPA - Automatic Radar Plotting Aid). Those functions fit into the second stage of human machine dependency (supervisory control) [12]. The third stage – fully automatic control – would correspond to unmanned vessels operations.

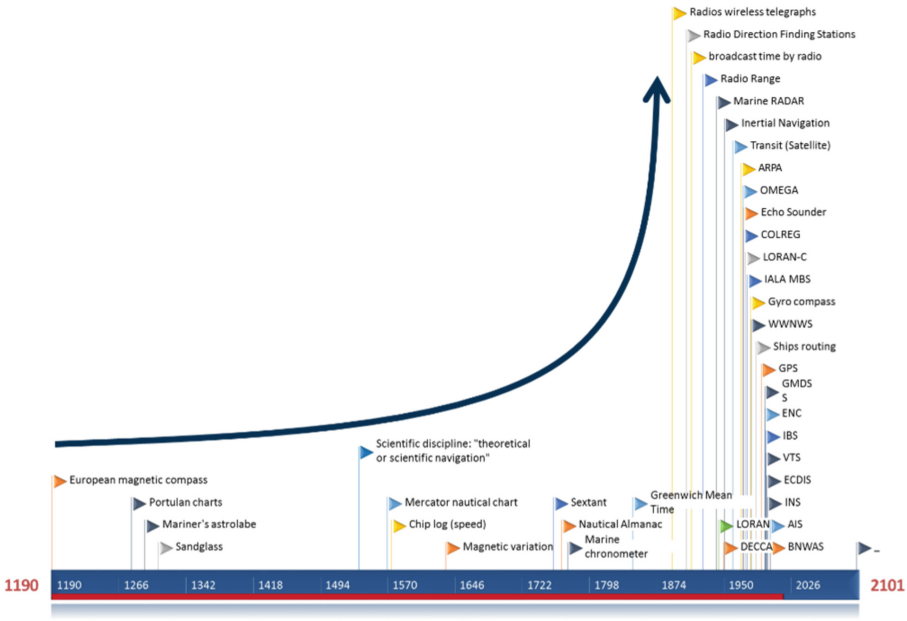


Fig. 1 Time line visualization of the trend in maritime navigation technologies and techniques

This alteration in the navigation tasks are stressed by the huge amount of data and information provided by on-board sensors, databases and shore based maritime services, leading to rising apprehensions about the workload and its effects on situation awareness and decision making [10–13]. Bainbridge [14] identified several automation glitches, such as over reliance, trust and feedback in human-machine collaborations, human drop of motivation and skills when engaged in monitoring tasks. The relevance of automation feedback and representations of its behaviour was highlighted in a ship the grounding study [15]. The same study points out the need for changes in the information systems to present new forms of highlighting changes and events, supporting the anticipation of changes, and facilitating the cognitive work engaged in the search and scanning in the displays.

Maritime navigation turned out to be a very complex and large-scale sociotechnical system comprising human and man-made entities that interact with each other and operate in a rough environment [16, 17]. Improving the maritime navigation performance, requires the understanding of how the process works and its context. Therefore, we need to address this problem as a joint human-technological activity. Some challenges were identified by Klein *et al.* [18] for the success of human-agent team activities, comprising issues like mutual **predictability**, **common ground**, **visible status** and **intentions**, **attention management**, **collaboration** and **negotiation**. The thinking, computation and the decision-making processes are no longer only dependent of the operator himself, but are also socially distributed among the elements of the team, it is happening in the individuals and the cognitive tools of the sociotechnical systems [19]. To develop a collaborative spatial decision-making tool, Antunes et al.

[20] overviewed several decision models to derive six different requirements, the **support of perception**, **retention**, **knowledge externalization**, **divergent/convergent activities**, **recognition** and **task/pattern management**. The collaborative view of this human-technological system leads to the assumption of an ecological perspective, where the navigation function is directed by the joint human-agent system, in opposition to the traditional navigator egocentric design.

In the most common manual of maritime navigation [21, 22], we may identify the key navigation tasks: **setting objectives**, **planning**, **execution**, **monitoring**, and **revising or adapting the plan**. From a different view, Jul and Furnas [23] proposed a navigator model which considers the following nonlinear tasks: **setting a goal**, **selecting a strategy**, **collecting information**, **perception**, **assessing**, **creating a cognitive map** and **moving**. The last four tasks represent the wayfinding/motion loop. Darken and Peterson [24] defined wayfinding as the cognitive element of navigation, involving the **formation of strategies** and **tactic** that will guide the movements. The development and use of the cognitive map are an essential element of the wayfinding. In this view, they consider navigation as the aggregate task of wayfinding and motion. Bjerva and Sigurjónsson [25] proposed another concept of wayfinding and suggested that it comprises three steps: **cognitive mapping**, **wayfinding plan development**, and **physical movement**. Based on observations of pilots' mental workload, Westrenen [26] proposed a navigator model with a three-stage decision model, comprising **tracking**, **short-term planning** and **long-term planning** behaviour. In this view, long-term planning concerns mostly with the voyage planning, prior to sailing, and can be associated with the previous concept of goals setting and strategy selection. Short-term planning, comprises local observation and information collection necessary to make decisions over the control of the vessel. Tracking corresponds to the assessment of the movement and correlation with the initial plan.

In the view of control theory, distributed control emerge from the increased interaction between the agents, which supports self-organization and adaptability, when facing uncertainty or unpredicted constraints [27, 28]. Time and predictability are the major determinants of controlling system [9], thus in situations of low predictability and available time we will find a reactive decision type, where the operator responds without having foreseeing the events.

The conceptual navigation model adopted for this study may be visualized in Fig. 2 (a). It comprises the three-main functions, forming a goal, defining strategies and moving. Several nonlinear tasks are required in each stage and, all of them, have deliverables (goals, plan and control actions over ship). To support the categorization of visualization factors, the process of navigation is contextualized in the work space, Fig. 2(b), and available time, Fig. 2(c), together with expected control modes, Fig. 2(d), cognitive and decision making processes, Fig. 2(e). As an example, while sailing, the navigator tries to minimize reactive decision (i.), as they are not thought, they are usually linked to emergency situations and can only be reinforced thru experience and training. Since it is impossible to eliminate all the uncertainties, it might be worth to better prepare the ship for those events. Thus, a sound planning is the first mitigation measure

(ii.), supported by cognitive process that helps to build a cognitive map of the physical voyage. At the same time, several alternatives are considered and tested. This process will further support the perception and understanding of the real world (iii.), and control is based on planned responses or projected events. Therefore, information systems, should be design to support the identified cognitive and decision-making processes, considering the navigation function, context and available time. For instance, observing Fig. 2, we can see that in the most front end stage, when acting in the conduction of the ship, events are occurring in the bridge and typical control modes are reactive and proactive. These control modes require continuous attention to stimulus and the operator must react in short time. From this case, we may argue that by enhancing attention abilities we strengthen the reaction possibilities. We may also focus on higher stages processes, such as enhancing the estimation and understanding cognitive processes, so the operator will use the proactive control mode, based on predictions that support better decisions.

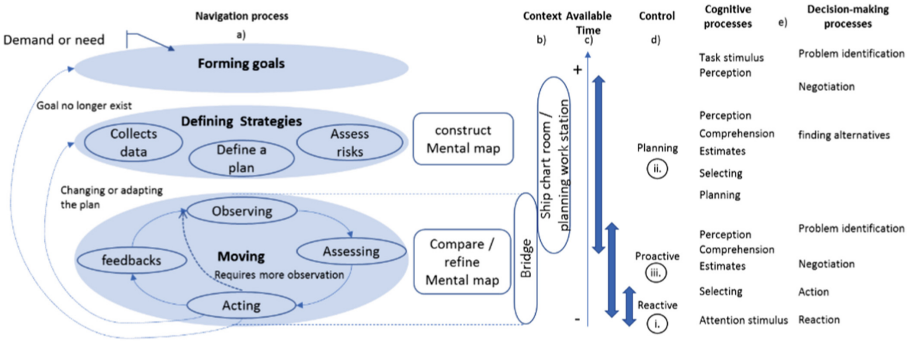


Fig. 2 Identifying the cognitive and decision-making processes within the navigation functions

3 Methodology

The literature research applied in this review was conducted in several databases, namely SCOPUS, Science Direct, Springer Link, EBESCOhost and ResearchGate. The articles were selected based on their relevance to visualization in the navigation and orientation domains. The following search terms were used: visualization, visual attention, visual perception, memory workload, cognitive workload, visual search, 3D visualization. Due to the large amount of results, additional keywords combinations were made with: wayfinding, navigation, orientation, map, decision-making, situation awareness. Studies in eNavigation domain and related with these areas were also retrieved. Finally, a comprehensive search was performed in a selection of conference proceedings over the last 5 years (Symposium on Information Visualisation - InfoVIS, International Conference Information Visualisation - IV).

Inclusion and exclusion criteria were driven by criteria such as: scientific Quality (theoretical basis, with empirical data set), scientific Context (references to related work), significance: originality, quality of presentation (clarity, illustrations); qualifications (credibility of authors and institutions), Studies addressing the relevant topics, published in English and in the recent years. To overcome the omission of significant studies, a selection of previous reviews that address the focus of the review question was made, providing not only a synthesis of the research already undertaken, but also an initial guidance for this follow up review.

Very few studies were conducted in the context of maritime domain, on other hand more were find in the field of air navigation. Apart from the above criteria, no review protocol was registered. The results were combined in the categories listed in Table 1, that were drawn from the navigation, cognitive and decision making processes illustrated in Fig. 2. This paper presents the analysis on the *Visual Attention* category.

Table 1. Categories and associated concepts used to guide de classification.

Cat ID	Category	Associated concepts
CAT1	Visual attention	Uncertainty, guided attention, visual search
CAT2	Visual memory	Information overload
CAT3	Visual perception	Strategies to predict and identify
CAT4	2D/3D	Space-time visualization, multi-attribute
CAT5	Wayfinding	Visualization in support of, strategies
CAT6	Planning	Creation mental map, learning, experience
CAT7	Multiple task	Support of, collaborative decision process

4 Results and Discussions

The results of the 24-selected sources, most of them published in journals (20), are summarized in Table 2. Publication years varies from 1980 to 2016, averaging 2008, and the majority are from the field of psychology and cognition, with some from computer vision and cartography. The dominant paradigm is experimental, together with four meta-analysis and few qualitative studies. Metrics are mostly associated with performance evaluation, measuring reaction times (RT), search times (ST) and target detection (TD). The author (first author) familiarity with maritime navigation helped in the identification and interpretation of the relevant consequences of the findings presented in the literature review. It is acknowledged that not all the theoretical and relevant publications may be presented, however we believe that this selection provides sufficient ground to support the claim of new design requirements in the visualization of navigational information and decision support for navigation control.

The viewers' role in his perception of visual information depends critically on where his attention is focused and what is already in his mind prior to viewing an image [29]. Additionally, human vision rapidly and automatically categorizes visual images into regions and properties (preattentive processing). Treisman's Feature Integration Theory [30], claims that if the target has a unique feature, one can simply access the given

Table 2. Integrated summary of the selected publications

Cit.	Year	Discipline, field	Methodology	Context
[28]	2012	Computer Vision & Pattern recognition	Meta-analysis	Attention and visual perception survey
[29]	1980	Psychology, cognition	Quant., 9 experiments Perfor. eval., ST + RT	Feature-integration theory hypothesis
[30]	1989	Psychology, cognition	Quant., 4 experiments Accuracy of TD & RT	Efficiency of visual selection, T-N similarity and N-N similarity
[31]	1994	Psychology, cognition	Computer simulation, literature review	Model of visual search, Guided search
[32]	2007	Psychology, cognition	Quant., 2 experiments Perfor. evaluation	Boolean map hypothesis
[33]	2001	Psychology, cognition	Quant., 3 experiments Perfor. evaluation	Visual attention, circular spots of various sizes
[34]	2016	Psychology, cognition	Quant., 5 experiments Perfor. evaluation	Visual search for target object in cluttered scenes
[35]	2012	Cognition, behaviour	Qualit., 1 experiments Perfor. Obs. + gaze	Ground traffic control decision support system
[36]	2004	Psychology, cognition	Meta-analysis	Review of guiding attributes for deployment of visual attention
[37]	2007	Computer vision & pattern recognition	Meta-analysis	Visual attention, Taxonomy of Clutter Reduction
[38]	2008	Psychology, cognition	Quant., 1 experiments Perfor. evaluation, RT	Color and location in a visual search
[39]	2015	Cartography	Descriptive - Quant.	Map Viewer Design for Seniors
[40]	2012	Cartography	Quant., modelling (computer science)	Map design, automatic symbolisation
[41]	2016	Cartography	Modelling (computer science)	Map design, distortion perception
[42]	2011	Visualization	Meta-analysis	Color use in visualization, survey
[43]	2010			

(*continued*)

Table 2. (*continued*)

Cit.	Year	Discipline, field	Methodology	Context
		Psychology, cognition	Quant., 3 experiments Perfor. evaluation	Visual search of low prevalence targets
[44]	2013	Psychology, cognition	Quant., 2 experiments Perfor. evaluation	Cognitive load, multiple displays
[45]	2012	Computer vision & pattern recognition	Quant., 3 experiments Perfor. simularion	Visual attention, feature type, layout impact on performance
[46]	2015	Computer Vision & Pattern recognition	Descriptive-Qualit., Metho. Case Study	Decision-making, uncertainty
[47]	2014	Computer vision & pattern recognition	Descriptive - Quant., Questionaire + 1 exper.	Visual attention, graphics analysis
[48]	2012	Cognition, behaviour	Quant., correlational analyses	Visual perception, spatial memory persuasive geocommunication
[49]	2016	Psychology, cognition	Quant., 2 experiments Perfor. evaluation	Visual attention, visual and semantic influences
[50]	2012	Psychology, cognition	Quant., 1 experiment Perfor. simulation	Visual attention assessment in HUD, methodology
[51]	2003	Psychology, cognition	Quant., 1 experiment Perfor. evaluation, RT	Visual attention, mapping spatial attention

feature map to see if any activity is occurring, it also suggests that the amount of difference between the target and the distractors will affect search time. Other studies revealed correlations of visual search with the type of task and the visualization settings [31], showing variations in search efficiency depending on the similarity of targets and non-targets. The Guided search theory [32] suggests that an activation map based on both bottom-up and top-down information is constructed during visual search, meaning that attention is drawn to peaks in the activation map that represent areas in the image with the largest combination of bottom-up and top-down influence. Based on this sights, the bottom-up activation depends on feature categorizations, whereas the top-down activation is driven by the viewer's goals when looking to an image in search for the required visual information. Boolean maps theory [33] considers that visual search comprises two stages: selection and access. In this view, the visual system selects some elements of a scheme, excluding the others, and proceeds for a deeper analysis by accessing additional details of the select elements. Finally, the ensemble coding theory [34] brings the idea that low-level vision can generate a quick summary of how simple visual features are distributed across the field of view. More recent experiments concluded that visual search of precisely known features are influenced by the presence of visually similar distractors due to limitations in selection and masking [35].

All these researches on visual processing, summarized in [29], gives us significant cues about how to govern the user's visual attention and the importance of task and contextual setting. Top-down approaches to guide operators' visual attention have been tested by a model based on heuristic decision making, to foreseen user's decision strategies [36]. The current navigational information displays (e.g. ARPA, AIS, ECDIS) provide a large collection of features, each with several and distinctive visual properties (e.g. colour, orientation, size), resulting in a complex visual representation. Hence, we should simplify the visualizations in regards to the users' task by minimizing visual confusion, this means for instance, adjusting the electronic charts symbols, depending on planning or monitoring tasks. While planning, the navigator has time to assess all the chart features, to set the route and safe boundaries in accordance with his risk assessment. On the other hand, while monitoring, he is firstly concerned on avoiding hazards and to follow the route plan, which suggest that we have two major sets of information: dangers and positioning/navigation features. Therefore, features belonging to each of this sets should share a coherent visualization structure, that would evolve as the operator moves for additional search to find a detailed target.

Among several object attributes, colour, motion, orientation and size are strong guides for the deployment of attention [37]. For many of them, the presence of a property is more readily detected than its absence, which might be relevant for the detection of moving target among stationary [37]. Additionally, visual search efficiency increases as a function of target–distractor difference and decreases as a function of distractor–distractor difference [31]. To address clutter in overcrowded displays, several mitigation strategies can be applied by manipulating the features' appearance, spatial distortion and animation [38]. The selection by colour in a multiple-item display, where location and colour information are independent from each other and equalized, is mediated by location information [39]. Yet, the attention to location seems to be equally influenced by colour and location cues. A study on maps colours, dark colours benefits the contrast of the map viewer content and elements, therefore improving the perception of the map contents. It also shown that long wave-lengths colours shortens the viewer's reaction times in colour perception [40].

Visualization of real time multidimensional data generates complex representations, as it may be found in ARPA radar displays, and it is critically increased when combined with other information layers (AIS or ECDIS). New properties of the guidance cues can emerge from the deliberate or unintended combination of attributes, inducing positive or negative variations in the operators' visual attention. On the bridge, most of the available information is geo-referenced, and the trends confirm that more data is becoming easily available, such as aerial photos, textual information, routing and passage plans, weather and oceanographic data. Integration of this data demands further considerations over the cluttering effects. The combination of several distinct objects in the same presentation must be reassessed differently from the traditional selection of different layers, each with its own visualization properties. Automatic symbolization methods were developed to address the needs for layer's integrations minimizing any data loss [41, 42]. All the features must be contextually and coherently merged to support and guide individual's visual attention. For instance, when supporting the perception of close situations, displays should provide a clear and prioritized view of all the hazards and dangers, blending features like depth contour,

RADAR tracks and AIS information. Despite the different dimension of each feature, they are all hazards and this could drive the design of a common colour scales properties [43], to categorize their relative risk properties, e.g. time to closest danger.

One strategy to visualize large amount of information, has been using divided or several displays. This is more relevant when considering tasks involving search for low prevalence (LP) targets, i.e., that rarely occurs. Low target prevalence alters the behaviour of the operator and the implication of this phenomena are the viewer tendency to leave the search prematurely or to make motor or response errors. However it was found that no positive effect came by dividing up the display or by forcing the observer to slow down and correct errors [44]. In this view, it would be important to classify LP targets situations, like alarm cues, and study new designs to support their detection in time. Complementarily, it was demonstrated that simultaneous view is more appropriate than sequential view, further suggesting that the sequential view did not alleviate the divided attention problem, when we could suppose that sequential view would be more useful for monitoring tasks [45]. Viewers have an unconscious tendency to search for targets in novel locations in the display, as opposed to looking at locations that have already been examined [29]. Therefore, we could infer that this phenomenon could determine the size of displays used by operators.

Capacity limits of attention strongly affects the effectiveness of information visualizations, particularly the ability to detect unexpected information. Visual search experiments revealed that [46] search effectiveness can be increased by grouping, namely for oddball search, and reducing variety specially for demanding tasks. From these findings, we could again argue that same objects should be visualized differently depending on the user's task. It has profound implications in the way information is currently presented in bridge information's systems and subsequently in operator's effectiveness to extract information from them. Taking the example of ARPA displays, user's attention could be guided by grouping targets based on distinctive characteristics, like distance, CPA, TCPA or vessel type. Colour and flicker are attributes already used, but their effectiveness is reduced in congested displays, compelling operators to increase de RADAR scale and therefore losing the overall perspective. AIS data provides much more possibilities of manipulation due to the larger number of available dimensions, however the lowest integrity of this data recommends prudent evaluation of the integrated results, as they may cover-up erroneous data.

User's sense-making processes may induce cognitive biases in the process of visual perception, i.e., the type of representation may be subject to incorrect interpretations, particularly if the user is not familiar with the presented pattern [47]. Misinterpretations can emerge from clustering, completeness, anchoring and framing errors. One example that can be found in the bridge, is the representations of the same objects with different orientation schemes, such as AIS data in the ECDIS (north aligned) and in the radar display in head up mode. The mode error is found when the user unconsciously appropriates one of the visualisation schema, due to its greater perceived authority, to another which unwittingly does not fit. In what concerns the effects of interpretation of missing graphical data, it was found that higher degree of decision-confidence was achieved with the combination of emptiness and explanation [48]. Thus, rather than visualizing the last state or completing the data with some estimation form, it's better to provide a cue over the missing data. Additionally, map rhetorical styles influences the

user trust in the data and confidence in answering questions about the data, which means that different rhetorical designs can help achieve different persuasive goals [49]. Moreover, it was demonstrated that visual attention is influenced by semantics, so among the visual qualifiers we need to ponder on the possible object meanings and how they might guide visual search [50].

Experiments have demonstrated that the locus of attention is symmetrically distributed around the viewer's point of fixation, leaving the peripheral area less observed [51, 52]. On-board, when observing the ECDIS or RADAR displays, the focus of attention is usually the own vessel, therefore noting that user's attention is around that point, we should think in means to represent prioritized targets near that area, so they get a higher chance to be spotted. Moreover, viewers can resume an interrupted search much faster than they can start a new search, due to the unconscious perceptual predictions they make about the target based on the partial information acquired during the initial glimpse of a display [29]. Additionally, based on the current display, users' domain knowledge may give expectations about where certain data might appear in future displays, improving viewer's ability to locate important data. Therefore, we should challenge the possibility to provide RADAR data representations with minimum necessity to change scales.

5 Conclusions and Future Work

This study is an important contribution to provide new insights in the analysis and understanding of cognitive processes involved in maritime navigation, opening new possibilities for human systems interactions. Regarding visual attention, several key issues were identified, both derived from both top-down and bottom-up visual processes. Displays sizes, sequenced windows, zooming tools and scale changes can impair visual attention. Guided attention effectiveness' depends mostly on the perceived user' task and contextualization of the represented objects. Finally, integration of several information layers raises the possibility of masking information, providing erroneous visual guidance or misperception, due to unrestricted cluttering effects or conflicting visual features' properties.

We recognize that many other factors affect the navigation process and not all issues remain in the visualization domain, thus these results should be integrated in a broader perspective when addressing the conceptualization and design of new maritime navigation systems. This perspective of visualization over the bridge systems, has shown that, among all the described constraints, it is essential to perform a systematic contextual research of the maritime navigation function, trying to understand all the interactions and processes that happen in this work domain. The other categories will be addressed in future publications and the concluding findings will be merged into a conceptual framework to assess vessels navigation performance.

References

1. Schraagen, J.M., Militello, L.G., Ormerod, T., Lipshitz, R.: Naturalistic Decision Making and Macrocognition, p. 438. CRC Press, Boca Raton (2008)
2. Davenport, M., Risley, C.: Information Visualization: The State of the Art for Maritime Domain Awareness. MDA Corporation, Dartmouth (2006)
3. Taylor, H.A., Brunyé, T.T., Taylor, S.: Wayfinding and navigation: mental representation and implications for navigational system design. Rev. Hum. Factors Ergon. **4**, 1–40 (2008). doi:10.1518/155723408X342835
4. Peterson, K.: Three dimensional navigation. In: Ocean 2002 MTS/IEEE, pp. 1046–1050. IEEE (2002)
5. Belev, B.C.: Information characteristics of a man-operator for ensuring the vessel's safety of navigation. In: ENC-GNSS 2010, p. 8 (2010)
6. Chawla, P.: Human factor competencies for the future mariner. Int. Shipp. News 1–6 (2014)
7. Kopacz, Z., Morgaś, W., Urbański, J.: The changes in maritime navigation and the competences of navigators. J. Navig. **57**, 73–83 (2004). doi:10.1017/S0373463303002522
8. Wahlström, M., Hakulinen, J., Karvonen, H., Lindborg, I.: Human factors challenges in unmanned ship operations – insights from other domains. Procedia Manuf. **3**, 1038–1045 (2015). doi:10.1016/j.promfg.2015.07.167
9. Hollnagel, E., Woods, D.D.: Joint Cognitive Systems. Foundations of Cognitive Systems Engineering. Taylor & Francis Group, Boca Raton (2005)
10. Woods, D.D., Patterson, E.S., Roth, E.M., Christoffersen, K.: Can we ever escape from data overload? A Cognitive systems diagnosis. Cogn. Technol. Work **4**, 22–36 (2002). doi:10.1177/154193129904300310
11. Endsley, M.R., Kaber, D.B.: Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics **42**, 462–492 (1999). doi:10.1080/1463922021000054335
12. Kaber, D.B., Endsley, M.R.: The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. Theor. Issues Ergon. Sci. (2003). doi:10.1080/1463922021000054335
13. Hicks, J.S., Durbin, D.B., Morris, A.W., Davis, B.M.: A Summary of Crew Workload and Situational Awareness Ratings for U. S. Army Aviation Aircraft. Army Research Laboratory, Aberdeen Proving Ground (2014)
14. Bainbridge, L.: Ironies of automation. Automatica **19**, 775–779 (1983). doi:10.1016/0005-1098(83)90046-8
15. Lützhöft, M.H., Dekker, S.W.A.: On your watch: automation on the bridge. J. Navig. **55**, 83–96 (2002). doi:10.1017/S0373463301001588
16. Perrow, C.: Normal Accidents: Living with High Risk Technologies. Princeton University Press, Princeton (1999)
17. Vicente, K.: The Human Factor, Revolutionizing the Way People Live with Technology. Routledge, New York (2004)
18. Klein, G., Woods, D.D., Bradshaw, J.M., et al.: Ten challenges for making automation a "team player" in joint human-agent activity. IEEE Intell. Syst. **19**, 91–95 (2004). doi:10.1109/MIS.2004.74
19. Hutchins, E.: Cognition in the Wild. MIT Press, Cambridge (1995)
20. Antunes, P., Sapateiro, C., Zurita, G., Baloian, N.: Integrating spatial data and decision models in an e-planning tool. In: 16th CRIWG Conference on Collaboration and Technology, pp. 97–112 (2010)

21. Bowditch, N.: The American Practical Navigator: An Epitome of Navigation. US, National Imagery and Mapping Agency, Bethesda, Md (2002)
22. UK Ministry of Defence (Navy): BR 45(1)(2) Admiralty Manual of Navigation - Principles of Navigation 1 (2008)
23. Jul, S., Furnas, G.W.: Navigation in electronic worlds. In: CHI 97 Workshop SIGCHI Bulletin, pp. 44–49. ACM, New York (1997)
24. Darken, R.P., Peterson, B.: Spatial orientation, wayfinding, and representation. In: Stanney, K.M. (ed.) Handbook of Virtual Environments: Design, Implementation, and Applications, pp. 959–998. Taylor & Francis, London (2001)
25. Bjerva, T., Sigurjónsson, T.: Wayfinding by means of maps in real-world settings: a critical review. J. Navig. (2016). doi:10.1017/S0373463316000643
26. van Westrenen, F.: Development of a navigator model by the use of mental workload measures. In: Conference on Human Decision Making and Manual Control (1995)
27. Flach, J.M.: Complexity: learning to muddle through. Cogn. Technol. Work **14**, 187–197 (2012). doi:10.1007/s10111-011-0201-8
28. van Westrenen, F.: User centred interfaces in co-operative distributed systems (2004). doi:10.1145/1005220.1005230
29. Healey, C.G., Enns, J.T.: Attention and Visual Memory in Visualization and Computer Graphics. IEEE Trans. Vis. Comput. Graph. **18**, 1–20 (2012)
30. Treisman, A.M., Gelade, G.: A feature-integration theory of attention. Cogn. Psychol. **12**, 97–136 (1980). doi:10.1016/0010-0285(80)90005-5
31. Duncan, J., Humphreys, G.W.: Visual search and stimulus similarity. Psychol. Rev. **96**(3), 433–458 (1989)
32. Wolfe, J.M.: Guided Search 2.0 A revised model of visual search. Psychnomic Bull. Rev. **1**, 202–238 (1994). doi:10.3758/BF03200774
33. Huang, L., Treisman, A., Pashler, H.: Characterizing the limits of human visual awareness. Science **317**, 823–825 (2007). doi:10.1126/science.1143515
34. Ariely, D.: Seeing sets: representation by statistical properties. Psychol. Sci. **12**, 157–162 (2001). doi:10.1111/1467-9280.00327
35. Wyland, H.C., Vecera, S.P.: Goal-directed attentional selection: limitations from input variables, not imprecision. J. Exp. Psychol. Hum. Percept. Perform. **43**, 169–180 (2016). doi:10.1037/xhp0000287
36. Möhlenbrink, C., Manske, P., Kirlik, A.: An analysis of a ground traffic control decision support system based on the 3-step principle of heuristic decision making. Human factors a view from an integrative perspective. In: Proceedings of the HFES Europe Chapter Conference Toulouse (2012)
37. Wolfe, J.M., Horowitz, T.S.: What attributes guide the deployment of visual attention and how do they do it? Nat. Rev. Neurosci. **5**, 95–501 (2004). doi:10.1038/nrn1411
38. Ellis, G., Dix, A.: A taxonomy of clutter reduction for information visualisation. IEEE Trans. Vis. Comput. Graph. **13**, 1216–1223 (2007). doi:10.1109/TVCG.2007.70535
39. Vierck, E., Miller, J.: Precuing benefits for color and location in a visual search task. Percept. Psychophys. **70**, 365–373 (2008). doi:10.3758/PP.70.2.365
40. Vrenko, D.Z., Petrovič, D.: Effective online mapping and map viewer design for the senior population. Cartogr. J. **52**, 73–87 (2015). doi:10.1179/1743277413Y.0000000047
41. Toomanian, P., Harrie, A., Olsson, L.: Automatic symbolisation methods for geoportals. Cartogr. J. **49**, 153–163 (2012). doi:10.1179/1743277411Y.0000000028
42. Sun, S.: Symbolize map distortion with inscribed circles in polygons. Int. J. Cartogr. (2016). doi:10.1080/23729333.2016.1179863
43. Silva, S., Sousa Santos, B., Madeira, J.: Using color in visualization: a survey. Comput. Graph. **35**, 320–333 (2011). doi:10.1016/j.cag.2010.11.015

44. Kunar, M.A., Rich, A.N., Wolfe, J.M.: Spatial and temporal separation fails to counteract the effects of low prevalence in visual search. Vis. Cogn. **18**, 881–897 (2010). doi:10.1080/13506280903361988
45. Jun, E., Landry, S., Salvendy, G.: Exploring the cognitive costs and benefits of using multiple-view visualisations. Behav. Inf. Technol. **32**, 824–835 (2013). doi:10.1080/0144929X.2011.630420
46. Haroz, S., Whitney, D.: How capacity limits of attention influence information visualization effectiveness. IEEE Trans. Vis. Comput. Graph. **18**(12), 2402–2410 (2012)
47. Ellis, G., Dix, A.: Decision making under uncertainty in visualisation? In: Conference on IEEE Workshop on Visualization for Decision Making under Uncertainty, VIS2015 (2015)
48. Andreasson, R., Riveiro, M.: Effects of visualizing missing data: an empirical evaluation. In: Proceedings of the International Conference on Information Visualisation, pp. 132–138 (2014)
49. Muehlenhaus, I.: If looks could kill: the impact of different rhetorical styles on persuasive geocommunication. Cartogr. J. **49**, 361–375 (2012). doi:10.1179/1743277412Y.0000000032
50. De Groot, F., Huettig, F., Olivers, C.N.L.: When meaning matters: the temporal dynamics of semantic influences on visual attention. J. Exp. Psychol. Hum. Percept. Perform. **42**, 180–196 (2016). doi:10.1037/xhp0000102
51. Hillebrand, A., Wahrenberg, E., Manzey, D.: A new method to assess pilots' allocation of visual attention using a head-up display. In: Human Factors a view from an integrative perspective. In: Proceedings of the HFES Europe Chapter Conference Toulouse (2012)
52. Tse, P.U., Sheinberg, D.L., Logothetis, N.K.: Attentional enhancement opposite a peripheral flash revealed using change blindness. Psychol. Sci. **14**, 91–99 (2003). doi:10.1111/1467-9280.t01-1-01425

Gaps Between Users and Designers: A Usability Study About a Tablet-Based Application Used on Ship Bridges

Yemao Man[1(✉)], Margareta Lützhöft[2], Nicole A. Costa[1], Monica Lundh[1], and Scott N. MacKinnon[1]

[1] Division of Maritime Studies, Unit of Maritime Human Factors, Department of Mechanics and Maritime Science, Chalmers University of Technology, 412 96 Gothenburg, Sweden
{yemao.man,nicole.costa,monica.lundh,scottm}@chalmers.se
[2] Australian Maritime College, University of Tasmania, Launceston, TAS 7250, Australia
margareta.lutzhoft@utas.edu.au

Abstract. In recent years, Information Technology (IT) applications have been developed for the traditional shipping industry aiming at improving safety and efficiency, but the design vision is dominantly technology-centric without much attention being given to the user's actual work experience or needs. A usability study is an approach to uncover the gap between users and designers. Although there is substantial research on usability, the values and usability of the tablet-based applications used on ship bridges have rarely been investigated. In this paper, a tablet-based computer application was developed to provide decision support to river pilots for passing under bridges in the Gothenburg area. Experienced pilots were invited to participate in a simulated navigational environment. By exploring the potential barriers in using such novel applications on a bridge, this study aims to address the human factor concerns for future IT design and developments in the maritime domain.

Keywords: Human factors · Usability · Design · Ship navigation · Mobile application · Human machine interaction

1 Introduction

With the fast development of automation and digitalization, intelligent computerized systems and Information Technology (IT) solutions are contributing to a new wave of knowledge across many industries [1]. The shipping industry is not an exception. In recent years, many digitalized products, applications, and services have been introduced into the maritime domain, which provides shipboard users unprecedented opportunities to access data and information as well as a plethora of platforms assisting in decision support. One of the major ongoing developments in the maritime arena is E-navigation, which proposes to use electronic means to harmonize "collection, integration, exchange, presentation and analysis of marine information on board and ashore, enhance navigation and related services for safety and security at sea" [2]. Notwithstanding such

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_21

overarching development, the introduction of virtual reality (VR), augmented reality (AR) and other visualization techniques also provides new platforms to support training, retrofitting and design solutions [3–5]; the adoption of cloud computing in scenarios from on-demand engineering calculations to port management [6, 7] reveals great advantages in scalable computation and flexible resource management. In addition, increasing utilization of the big data concept oriented for the high volume of "sensor" data integration allows opportunities for ship performance monitoring and optimization to increase safety and efficiency [8, 9]. The emergence of autonomous systems, innovations to Location-Based Service (LBS), Wireless Sensor Networks (WSN) and the Internet of Things (IoT) in other industries [10, 11] have also contributed to the rapid growth and acceptance of autonomous intelligent ship systems [12–14]. Due to the intensive development of mobile technologies in the last decade, mobile devices such as smartphones and tablets have evolved into powerful portable computers that can process ubiquitous computing and allow distributed decision-making across industries. While there are still few mobile applications used for navigational purposes in the maritime domain, and changes to the regulatory environment are slow, the potential of mobile solutions might be infinite within the industry.

The shipping industry of today is at a critical point towards improvements in safety with the power of digitalization and IT while finding savings to ensure economic sustainability. Safety and sustainability often create economic conflicts. The Strategic Plan from IMO has recognized that "technological developments may have created new opportunities, but may also have negative consequences…if current technological advances continue without proper co-ordination there is a risk that the future development of marine navigation systems will be hampered through a lack of standardization… incompatibility between vessels and…complexity" [15]. Adding these technologies does not necessarily improve safety, especially if it propagates into rapid introduction in IT without full integration and downsizing the human presence in safety critical activities. It may seem intuitive to the IT service providers that digitalization can create value in retrofitting old ships, especially allowing navigators and ship engineers to access the data from vendor-specific analog devices in an easier and integrated way, but the trend of incorporating various sophisticated IT applications into the maritime domain is dominantly technology-centric. It has given less recognition to the human element and appropriate adaptation of the novel technological applications to the everyday work practice of shipboard users. Today the work on a ship's bridge is generally collective in a sense that the individual navigator needs to communicate and coordinate information to complete the tasks [16, 17]. While this group of users is not necessarily tech-savvy, the designers that develop state-of-the-art computer applications are not necessarily aware of the key success factors in certain navigational tasks either. When an accident occurs due to improper implementation of technology, end-users usually become the "culprits", blamed for these errors [18–20]. Usability is normally considered a second order activity in the engineer-centered world [21], so the consequence is that the users have to focus on how to use technology instead of on the tasks. Historically this is known as "computer-induced trouble" [22]. It has fostered the growth of the Human Computer Interaction (HCI) discipline in usability research [23], which in many cases today is still administered by IT professionals without end-user input [24]. It is important to evaluate

the usability of the novel systems with human factor practitioners and end-users, which has been proven to be an indispensable approach for assessing how helpful they could be and how satisfied the users should be [24, 25].

The aim of this study, thus, is to explore how new types of IT applications could influence a navigator's task performance, and probe how advanced technologies should be designed and organized around human workers to facilitate information exchange, coordination and decision-making. Although there is substantial research on usability of digitized applications, the values and usability of a novel tablet-based application used on a ship bridge for information coordination during inland navigation has rarely been investigated. In this paper, a tablet-based computer application was developed to provide real-time decision support to river pilots to optimize the estimated time of arrival (ETA) for passing under bridges in the Gothenburg area. Experienced river pilots were invited to participate in a simulated navigational environment and as the subjects of follow-up interviews. By exploring the potential barriers in using such novel mobile applications on a bridge, the study aims to address the human factor concerns for future IT developments in the maritime domain.

2 Mobile Application Design and Development

The mobile application of interest, GOTRIS (Göta River Information Services), was intended to create a digital decision support platform for river pilots on a ship's bridge to coordinate information of bridge openings and optimize ETA of passing under bridges in the Göta River of Gothenburg, Sweden. The information could be assembled from different transportation sectors (e.g., railway, shipping, and road) and parties of interest (e.g., harbours, bridges and weather monitoring stations). Considering it could be a standalone browser/server web application used on a ship bridge, IT professionals from a Swedish research institute specializing on applied IT research and development, developed the GOTRIS interface on an Asus Vivo Tablet running Windows 8.1. Tablets are usually equipped with a rich set of sensors that can allow data communication in an energy-efficient way (e.g., access to 4G data along the inland waterways while receiving real-time location updates via GPS). GOTRIS would enable the river pilot onboard to see the information from the administrations of bridges, weather forecast stations and other vessels nearby, and adjust the navigation schedule in relation to the bridge's opening time. The core function is to coordinate information from various stakeholders regarding water, road and railway traffic for an optimal bridge passing.

After login, the pilot could see several different tabs including Pilot (main tab), Weather, Meetings, Schedule, etc. The main tab view was designed to give essential information about the oncoming bridge and ETA (see Fig. 1). It displays the current position, the recommended speed of the ship if it needs to arrive at the bridge at the agreed time, the next bridge that the ship was going to meet, the prognostic meeting time (in red) and the confirmed time by the bridge administration (in green), and weather information related to the next bridge. The bridge administration played a central role in bridge opening management. They had to primarily check with schedules of trains and ETAs of ships sent from GOTRIS prior to confirming it on the platform or using

Very High Frequency (VHF) radio to contact the ship to hold its current position (see Fig. 1).

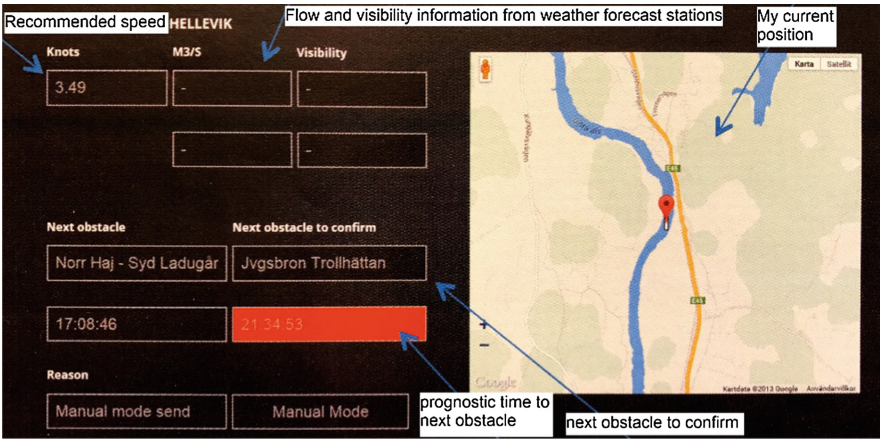


Fig. 1. The main tab view in the tablet-based application GOTRIS.

The meeting view outlined the predicted ship meeting trajectory estimated by GOTRIS (see Fig. 2) with respect to the next bridge's position. X-axis stood for the bridges that the vessel was going to meet chronologically and the Y-axis stood for the meeting time. As an example, Fig. 2 illustrated that the ship MISTRAL (in black) was going to meet the ship LOTTA (in green) at around 19:20 before arriving at the Vargön bridge.

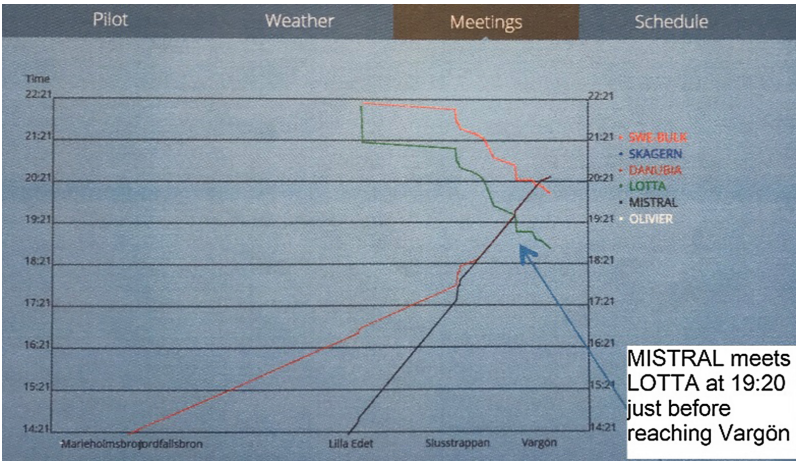


Fig. 2. The meeting tab view in the tablet-based application GOTRIS.

The schedule tab listed the exact meeting time estimated by GOTRIS (see Fig. 3). It provides a list overview of bridges and corresponding meeting time.

Fig. 3. The schedule tab view in the tablet-based application GOTRIS.

Overall, the essential information that GOTRIS could provide is the recommended speed of the vessel based on the confirmed bridge opening time and the dynamic position of the vessel in real time. The intention was to support the pilots to make decisions in speed control so that the vessel would neither arrive early and wait for the bridge to open (e.g., when the bridge-crossing train needs to pass first), nor miss the bridge opening time.

3 Method

Eight experienced pilots volunteered to participate in an experiment that examined navigating simulated ships in the inland waterways of Gothenburg region and participated in follow-up debriefing interviews. Their ages ranged from 33 to 53 years old with a mean age of 42 (SD = 10.5) years old. Previous experience as pilots varied between 6 to 20 years, with a mean period of 12.6 (SD = 4.3) years. They had been informed of the aim of the study and their rights to make their own decisions regardless of what the GOTRIS application recommended. Eight master mariner programme students were designated helmsmen paired with pilots in the scenario trials. Their task was to cooperate with the experienced pilots and execute navigation commands from the pilots. All participants signed a written consent form prior to the participation.

The scenario trials took place in a full-mission bridge simulator at Chalmers University of Technology in Gothenburg. There were total eight repeated rounds of a same scenario for the eight pairs of pilots and students. Each scenario lasted about 3 h. In each

run of the scenario, there was one pilot, one student, one observer, one IT professional, and one simulation assistant standby (acting as the bridge administrator). The tablet was located by typical navigational instruments. The pilot would navigate the ship with the student's assistance from Marieholmsbron (Marieholms Bridge) to the Götaälvbron (Göta River Bridge), but it was up to her/him when and how to use the GOTRIS programme for making schedules of bridge passing. They were encouraged to speak aloud their thoughts and intentions. The direct observation method was utilized by the experimenter to collect data regarding the physical and verbal aspects of the scenario, including individual actions (e.g., looking, touching, searching through the GOTRIS application), errors made, communications between individuals, the technology involved in the scenario and the system environment. Through direct observations, the experimenter was supposed to capture and understand the context within which the participants interacted. The method has been widely used, especially for the starting point of analysis efforts [26].

After the simulation run, the pilot was asked to fill in the Quality in Use Scoring Scale (QIUSS) questionnaire [27]. It required the participants to subjectively rate each dimension of usability (effectiveness, efficiency, safety and satisfaction) on a 5 point rating scale (0 = Low, 5 = High) based on the application's performance during the task under analysis. Then they would be interviewed with the following questions followed by an open-ended discussion on the GOTRIS application:

(1) What information do you think is useful from the tablet application?
(2) What information you need is missing in the application?
(3) How valuable do you think the tablet application is during your tasks?
(4) How much do you trust the application in terms of decision support?
(5) How do you want the design of the application to be improved for its usability?

4 Results

The average score concerning effective, efficiency, safety and satisfaction in QIUSS is presented in Fig. 4. The interviews revealed that they were slightly positive about this new IT service on a bridge in terms of usability, but the ways of using the service and the trust about the application varied depending on the navigation situation. The field notes taken during the scenario observations and subsequent interviews suggested that most pilots relied on their own experience rather than GOTRIS when the ship was approaching a bridge. One commonly observed behavioural pattern was that pilots used GOTRIS more frequently as long as the location of the ship was far away from the bridge. Once the ship was getting closer to the bridge, the pilots would change their reliance from GOTRIS to their own judgment as well as other conventional electronic nautical charts and VHF radio calls. Almost all pilots diverted their attention from these mobile devices when they were performing safety-critical bridge passing manoeuvres. Some of them thought the conventional communications between ship and the bridge administration via VHF was critically important in their final decision-making and maneuvering tactics.

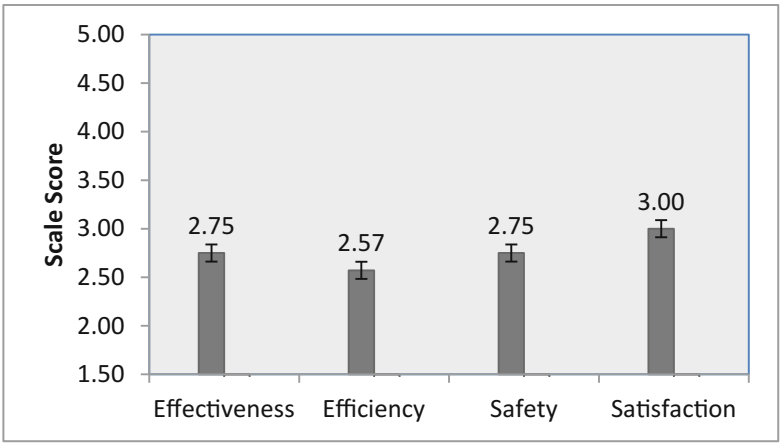


Fig. 4. The average score concerning effective, efficiency, safety and satisfaction in QIUSS.

Data reception delay was identified as a serious problem using such wireless services. It was observed that the ETA was calculated but not confirmed by the bridge administration after a long time delay, likely due to unstable network connectivity. However, no feedback was given to the pilot to inform of that connection failure, so the pilot had to use VHF to contact the bridge administrator about the bridge opening time. Consequently, the application was ignored most of the time in such scenarios. Additionally, the pilots reported that they sometimes found it difficult to understand the meeting view within the tablet or made much effort searching and filtering out the right data in the list entry under the schedule tab that matched their current situation.

In terms of the value proposition, all participants thought the tablet application was not absolutely needed in their everyday work practices. Some believed it was meaningful to have the ETA and recommended speeds in the system for the purpose of building equally trustable parties, such as the railway and river traffic. If a ship keeps her schedule but a train fails, the train should wait for the ship to pass under the bridge rather than the ship waiting for the train, which is the current reality. Others believed that the tablet could well be integrated in the electronic navigational charts in some ways or installed in a fixed position on a ship bridge for better usability.

5 Discussion

Methodologically, the combined approach of using observation, questionnaire and interview functioned well in collecting both objective performance data (e.g., behaviours and verbal communications) and subjective data (e.g., opinions), and furthermore lead researchers to uncover potential mismatches between the user's model and the designer's model of the interactive system of interest [28]. In this case, we recognized a clear gap between the "conservative" users who adhere to conventional ways to navigate the ship and "liberal" designers who embrace novel approaches to apply IT.

One motivation for GOTRIS was to provide a portable platform for ship pilots to coordinate information concerning bridge openings so that an optimal navigation plan could be achieved. From a technology perspective, the platform was designed as a mobile responsive web application populated with the state-of-the-art HTML5 technologies to enhance user experience. No participant seemed to appreciate the mobility convenience of the applications on a bridge. It was just "*like another toy*", said one participant. In addition, the problem of wireless data connection, though considered serious, is not a rare issue for internet-required tablet applications. Factors like bad weather may influence the cellular data connection performance, but the real problem is lack of engineering proper feedback to the user after such failures of data transmission delays, which increases the user frustration and decreases reliability of the software.

The score in QIUSS proved that the usability issues could also go beyond the technological implementation maturity. The main feature within the application was to recommend speed for the vessel, creating an ETA, giving strong decision support to the pilots to schedule navigational strategy way ahead of bridge passing, in order to deliver a more efficient and safer system performance. It was concluded that the usability of this function highly depends on the distance to the bridge. Most pilots considered the recommended speed as one "*formal thing of little value*". In a designer's world, ETA and recommended speed might be important system performance indicators to guide users' actions and moves. While this seems obvious, it does not necessarily accommodate the world of navigators in tasks like river piloting. As one pilot put in the interview, "*a pilot's job is to remind the captain how much time we still have*", it is important to understand the meaning of "time" and "speed".

An understanding of the work domain is a prerequisite to interpret the constraints in the context and display meaningful cues within the interface. It is valuable that designers can fathom ship handling as a non-linear system. When the ship is far away from the next bridge, the ETA and recommended speed could be valuable because it serves users a decision-making platform. However, as the ship approaches the bridge, the compliance of recommended speed or not would make a huge difference. "*It doesn't matter [if] it is 5 knots or 10 knots when you are far away but it will definitely matter when the meeting is imminent*.", said one participant. This was consistent with what we observed in scenarios that the pilots would start to turn from a "passive monitoring" work mode to an "active controlling" mode [29]. They kept looking out of the window and adjusting maneuvers. In the end, they no longer read the tablet to compare the recommended speed with the current speed at all, as it apparently could not support them with the desired assistance to pass the open bridge safely. The *in-situ* ship sense together with their past experience [30] governed their cognitive resources, enabling them to mentally calculate the needed speed and optimize navigating strategies. This was exactly like what one pilot described, "*the speed adjustment and ETA can be done within my brain*". So what is the value of the newly introduced technology on a ship's bridge?

This possibly explains why the rating is moderate between 2 and 3 in the scale of QIUSS, which indicates that the application can enable the users to achieve adequate performance and users can perform the tasks without hindrance but it does not provide sufficient assistance in practice [27]. What designers believe is useful in the system could be useless in the user's world.

The usability question is often a question of usefulness in its visualization solutions [31]. There were also lessons to be learned from this perspective. In the scenarios, it was observed that pilots expended considerable workload to understand the schedule list (see Fig. 3). The interviews also revealed the readability issues in understanding the view of meeting trajectory (see Fig. 2). These issues represent a typical gulf of evaluation in human computer interaction, i.e. the degree to which the artifact provides representations that can be directly perceived and interpreted in terms of the expectations and intentions of the user [32]. When a user starts to distinguish an element from a set of entities, he/she is essentially extracting elementary aspects of visualizations such as form, colour, texture, motion, and stereoscopic depth. If all aspects of the observed entities are similar (e.g., rows in a list), then none would pop up for users to find easily [33].

Nevertheless, GOTRIS was still considered valuable by the participants in the open-ended interview with respect to its pivotal role in information coordination for multiple stakeholders. The context for which GOTRIS was designed for is essentially a socio-technical system: the pilot was communicating with the helmsman to execute the maneuver strategy using the navigational instruments and coordinating with the bridge administration about bridge openings via the tablet or VHF radio calls. The decision-making process is complex and distributed per se. In order to provide decision support to the pilots, the design has to thoroughly understand the nature of their work, experience, and the contextual constraints. The nature of conventional piloting work is shifting to a Computer Supported Collaborative Work (CSCW) where collective efforts and knowledge transfer need to be managed and supported by the IT system. This might be the most important demand posing on the future IT developments aiming at decision support.

6 Summary and Conclusions

Digital applications are increasingly featured in the shipping industry. The tablet-based web application GOTRIS is one instance of these IT innovations to provide decision support to the river pilots to optimize the time for bridge passing in the Gothenburg area. In this paper, a simulator study was conducted to explore underlying gaps between maritime users and IT professionals beneath the GOTRIS program.

The results show that one main usability issue is related to the ship's distance to the bridge. The interviews reveal the limited value of the application as well as the gaps between users and designers. Finding the gaps is our final aim in an approach to obtain insights about software design. Software is more than a mediated device that users interact with. It also provides a *space* that designers create in which users' lives and experience get shaped [34]. Designers in the software industry may have freedom to consider a plethora of means to gather user requirements and adopt various engineering paths to build and decorate the *space*, but eventually this *space* will influence users' patterns of interactivity and nurture their knowledge [35]. Therefore, whether the *space* is compatible and harmonized with the user's actual work demands is vital in its usability assessment and value creation. The truism of complete user requirement elicitation is not panacea, but the designers need to focus the design vision to the world of the user

instead of merely accentuating technological innovations. A deeper understanding of the task, the user's work experience, the nature of his/her work and the context in which the user performs the task is crucial in designing usable products and creating value. Such vision is the key for the designers to understand what can influence usability and how it can be improved. One possible approach is to pose questions concerning the user's work experience during field observations. For instance, how do the pilots prioritize tasks and determine when to change speed? What are resources and constraints in the process? How can the interaction facilitate understanding and sense-making? When there are unexpected occurrences in the system (e.g., delays in data connection), how much fault tolerance can the system provide so that the user experience is the least compromised? These regards will shape the value of the software in a sense. If the design work stays at demonstrating technical solutions, then the product would risk becoming another ornamental bauble.

The bridge today is mainly about integration work in which seafarers are proactively constructing the workplace to meet the needs of tasks and duties [16]. With more computers and digital services introduced onboard in the future, conventional navigation is shifting to a CSCW, which poses great challenges on the HCI design in terms of resource integration and interface adaption for both individual and collective needs. The findings in this study suggest that what the shipping industry craves today is not likely another new form of interaction or a digital gadget, but a holistic system approach for interface design and system integration. That might be the ultimate gap on the system level that the community of users, designers and regulators need to address.

Acknowledgments. We gratefully acknowledge the support from RISE Viktoria and the GOTRIS Project Consortium.

References

1. Hermann, M., Pentek, T., Otto, B.: Design Principles for Industrie 4.0 Scenarios. In: 2016 49th Hawaii International Conference on System Sciences (HICSS) (2016)
2. IMO: E-Navigation. 2014. http://www.imo.org/en/OurWork/safety/navigation/pages/enavigation.aspx. Cited 6 March 2017
3. MacKinnon, S.N., Bradbury-Squires, D., Button, D.: Virtual reality based training improves mustering performance. In: 15th International Conference on Computer and IT Applications in the Maritime Industries—COMPIT 2016, Lecce, Italy (2016)
4. Lukas, U.F.v.: Virtual and augmented reality for the maritime sector—applications and requirements. In: IFAC Proceedings, vol. 43(20), pp. 196–200 (2010)
5. Mallam, S.C., Lundh, M., MacKinnon, S.N.: Evaluating a digital ship design tool prototype: designers' perceptions of novel ergonomics software. Appl. Ergon. **59**(1), 19–26 (2017)
6. Gentzsch, W., Purwanto, A., Reyer, M.: Cloud computing for CFD based on novel software containers. In: 15th International Conference on Computer and IT Applications in the Maritime Industries—COMPIT 2016, Lecce, Italy (2016)
7. Heilig, L., Voß, S.: A holistic framework for security and privacy management in cloud-based smart ports. In: 15th International Conference on Computer and IT Applications in the Maritime Industries—COMPIT 2016, Lecce, Italy (2016)

8. Jan, R.Ø., Prasad, P.L., Brage, M.: Big data in shipping—challenges and opportunities. In: 15th International Conference on Computer and IT Applications in the Maritime Industries—COMPIT 2016, Lecce, Italy. Technische Universtat Hamburg, Hamburg (2016)
9. Koga, S.: Major Challenges and Solutions for Utilizing Big Data in the Maritime Industry, p. 91. World Maritime University, Malmö (2015)
10. Man, Y., Ngai, E.C.H.: Energy-efficient automatic location-triggered applications on smartphones. Comput. Commun. **50**, 29–40 (2014)
11. Gubbi, J., et al.: Internet of Things (IoT): a vision, architectural elements, and future directions. Fut. Gener. Comput. Syst. **29**(7), 1645–1660 (2013)
12. Man, Y., et al.: From desk to field—human factor issues in remote monitoring and controlling of autonomous unmanned vessels. Procedia Manuf. **3**, 2674–2681 (2015)
13. Man, Y., Lundh, M., Porathe, T.: Seeking harmony in shore-based unmanned ship handling—from the perspective of human factors, what is the difference we need to focus on from being onboard to onshore. In: Stanton, N. et al. (eds) 5th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences, AHFE Conference, Krakow, Poland (2014)
14. Man, Y., Rylander, R.: Autonomous Safety on Vessels—An International Overview and Trends Within the Transport Sector. Lighthouse, Gothenburg (2016)
15. IMO: Strategy for the Development and Implementation of E-Navigation (2014). http://www.imo.org/en/OurWork/Safety/Navigation/Documents/enavigation/MSC%2085%20-%20annex%2020%20-%20Strategy%20for%20the%20development%20and%20implementation%20of%20e-nav.pdf. Cited 6 March 2017
16. Lützhöft, M., Nyce, J.: Integration work on the ship's bridge. J. Marit. Res. **5**(2), 59–74 (2008)
17. Lundin, J.: Talking About Work. Designing Information Technology for Learning in Interaction. Department of Informatics, University of Gothenburg, Gothenburg (2005)
18. Lützhöft, M.H., Dekker, S.W.A.: On your watch: automation on the bridge. J. Navig. **55**(01), 83–96 (2002)
19. Stanton, N.A.: Human error identification in human–computer interaction. In: Julie, A.J., Andrew, S. (eds.) The Human—Computer Interaction Handbook, pp. 371–383. Lawrence Erlbaum Associates, Mahwah (2003)
20. Grech, M.R., Horberry, T., Smith, A.: Human error in maritime operations: analyses of accident reports using the Leximancer tool. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 46(19), pp. 1718–1721 (2002)
21. Costa, N.: Human Centred Design for Maritime Safety: A User Perspective on the Benefits and Success Factors of User Participation in the Design of Ships and Ship Systems. Department of Shipping and Marine Technology, Chalmers University of Technology, Gothenburg (2016)
22. Landauer, T.: The Trouble with Computers Usefulness Usability and Productivity. MIT, Cambridge (1995)
23. Maxwell, K.: The maturation of HCI: moving beyond usability toward holistic interaction. In: Carroll, J.M. (ed.) Human-Computer Interaction in the New Millennium, p. 703. ACM/Addison-Wesley, New York (2001)
24. Grech, M.R.: In: Grech, M.R., Horberry, T.J., Koester, T. (eds.) Human Factors in the Maritime Domain. CRC Press, Boca Raton (2008)
25. Kjeldskov, J., Skov, M.B.: Creating realistic laboratory settings: comparative studies of three think-aloud usability evaluations of a mobile system. In: Conference: Human–Computer Interaction—INTERACT '03: IFIP TC13, Zurich, Switzerland (2003)
26. Stanton, N.A., et al.: Human Factors Methods: A Practical Guide for Engineering and Design. Ashgate, Uxbridge (2006)

27. Sherwood-Jones, B.S.: Quality In Use Scoring Scale (QIUSS) (2008). http://www.processforusability.co.uk/QIUSS/QIUSS.pdf. Cited 8 Feb 2017
28. Johnson, J., Henderson, A.: Conceptual models: begin by designing what to design. Interactions **9**(1), 25–32 (2002)
29. Porathe, T., Prison, J., Man, Y.: Situation awareness in remote control centres for unmanned ships. In: RINA Human Factors in Ship Design and Operation Conference, London, UK, pp. 93–101 (2014)
30. Prison, J., Lützhöft, M., Porathe, T., Ship sense—what is it and how does one get it? In: RINA Human Factors in Ship Design and Operation Conference, London, UK, pp. 127–130 (2009)
31. Tory, M., Moller, T.: Human factors in visualization research. IEEE Trans. Visual Comput. Graphics **10**(1), 72–84 (2004)
32. Norman, D.: Design of Everyday Things: Revised and Expanded. MIT, London (2013)
33. Ware, C.: Information Visualization: Perception for Design, 3rd edn, p. 536. Morgan Kaufmann, San Francisco (2012)
34. Winograd, T.: Bringing Design to Software. Addison-Wesley, Reading (1996)
35. Löwgren, J., Stolterman, E.: Thoughtful Interaction Design: A Design Perspective on Information Technology. MIT, Cambridge (2004)

Road and Rail: Driver Safety

ATV Safety in Agriculture: Injury, Illness, Analysis and Interventions

David Gilkey(✉), John Rosecrance, Hali Neves, and Elise Lagerstrom

Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO 80523-1681, USA
dgilkey@colostate.edu

Abstract. All-terrain vehicles (ATVs) have been implicated in over 14,000 deaths since 1982 and continue to be a major public health transportation related hazard. An estimated 78% of riders use ATVs for recreational purposes and the remaining 22% use the vehicles for various occupational applications. Recent data analysis reveals a 300% increase in work related ATV deaths. The agricultural community bares the greatest burden, experiencing 65% of all occupational related fatalities occurring on the farm or ranch. An examination of 1230 cases reveal that patterns exist between agent, host and environmental factors. Haddon's energy transfer injury theory provides an opportunity to develop human factors and ergonomic safety interventions. This presentation will include some of the historical background on the expanding popularity of ATVs, epidemiology of injury and illness related to loss of control events, case studies that demonstrate a range of interventions and application of human factors principles.

Keywords: All terrain vehicles · Quad bikes · ATV · Injury · Fatality · Agriculture

1 Introduction

The introduction of all-terrain vehicles (ATVs) in the 1970s was well received in the US increasing to an estimated 11 million units and 35 million riders in 2016 [1] The ATV Safety Institute (ASI) estimates that 78% of riders use their ATVs for recreational purposes and 22% use them for occupational applications. Lagerstrom et al. evaluated workers' compensation data on ATV related injuries from Montana for years 2007 through 2012. That data identified various occupations and industry types that were using ATVs in various operations. That study found that the vehicles were used predominately in agriculture but also in other industries such as public administration, construction, professional and technical services, administrative support services, utilities, manufacturing, mining, retail trade, finance, insurance, and food services [2]. Over the past four decades ATVs have flourished in popularity and their increased applications in occupational settings.

In the agricultural community the ATVs were seen as versatile equipment that bridged the gap between horse and pickup truck or tractor and horse [3]. All-terrain vehicles are uniquely designed and well suited for the landscape common to farms and

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_22

ranches. The vehicles are high-centered for ground clearance with low pressure tires that effectively grip travel surfaces, are mounted like a motorcycle, have a straddle seat, foot rests and are steered with handlebars [3]. The ATVs are manufactured in three common sizes to accommodate youth, juniors and adults users. Motor sizes and vehicle weights range from 90 cubic centimeters (cc) and less than 300 lbs. for youth models, 150 cc and 350 lbs. for juniors to 1000 cc and 800 lbs. for adult models that can travel at high speeds nearly 80 miles per hour [3]. The three-wheel vehicles that entered the market in 1970 were phased out in 1988 by decree and agreement between the industry and the Consumer Product Safety Commission (CPSC) as a result of the increasing injuries associated with ATVs.

2 Epidemiology

In 2010, the CPSC reported the adverse health impacts of ATVs peaked at an estimated 400,000 injuries per year resulting in 139,000 emergency room visits for medical care and nearly 800 fatalities experienced in 2008 [3]. The most recent data from CPSC indicated that 14,129 ATV fatalities had been investigated since 1982. In 2015, CPSC received 340 death notifications related to ATVs and estimated that emergency room visits declined to 97,000, see Table 1 [4]. The five states with the most deaths reported between 1982 and 2017 were Texas (780), California (703), West Virginia (725), Pennsylvania (697) and Kentucky (652), see Fig. 1 [5].

Evaluation of workers' compensation injury data from Montana revealed that gender and age were factors. Lagerstrom et al. found that 87% of those injured were male and 23% were 20 to 29 years of age. The next most frequently injured age group were 50 to 59 years of age comprising 20% of claims. These findings were contrasted to the CPSC 2015 data where injuries were highest among riders under 16 years of age making up 27% of emergency room visits. Whereas those between 16 and 24 years comprised 25% of the total and those 55 years and older group totaled only 7% of all visits [4].

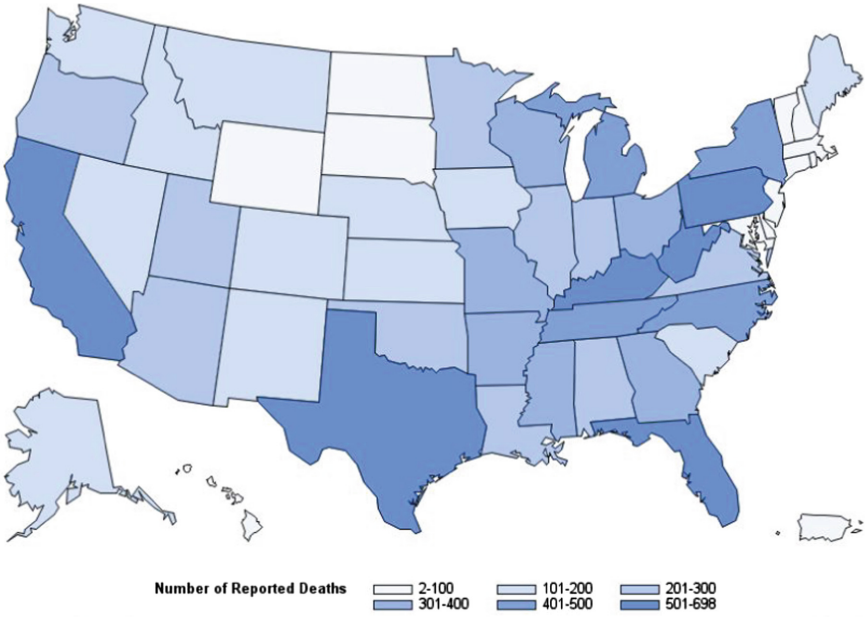
Lagerstrom et al. reported that 59% of all Montana injury claims were from the agricultural sector [2]. These data supported previous findings from Helmkamp, Biddle, March and Campbell that found 65% of fatalities among all occupational sectors were from agriculture. Their investigation evaluated workplace ATV deaths between 2003 and 2006 and identified Montana as the state with the highest number of fatalities in that time period. Injury and fatality rates among all occupational users is on a sharp increase of 300% between the years 1992 through 2007 suggesting a crisis is looming [6, 7]. As occupational applications of ATV have increased, there is great concern about the dramatic increase in deadly events. Effective interventions are needed to curb the growing number of incidents among workplace users. Investigators evaluated 1230 ATV fatality cases from the CSPC using an agent, host environment model to assess causation and opportunities for interventions [8].

Table 1. Reported ATV-related fatalities (by Year); ATVs with 3, 4, or unknown number of wheels; reported for the period January 1, 1982 through December 31, 2015

Year	Reported number of deaths	Difference since last update (12/31/2014[a])
Total	14,129	+512
2015	340	+340
2014	547	+162
2013	581	+34
2012	573	−1
2011	622	−4
2010	650	−4
2009	720	−1
2008	759	+1
2007	832	−1
2006	831	−1
2005	799	0
2004	747	−6
2003	651	+ 1
2002	547	−3
2001	517	−3
2000	445	−2
1999	396	−1
1998	253	+ 1
1997	240	0
1996	249	0
1995	200	0
1994	198	0
1993	183	0
1992	220	0
1991	230	0
1990	235	0
1989	230	0
1988	250	0
1987	264	0
1986	300	0
1985	250	0
1984	156	0
1983	85	0
1982	29	0

Source; U.S. Consumer Product Safety Commission: Directorate for Epidemiology/Divsion of Hazard Analysis
Note: Italics indicate that reporting is ongoing for the years 2013-2015
Note: The heavy line marks the change from death certificate mortality coding under the Ninth Revision of the international classification of Diseases (ICD-9) to coding under the Tenth Revision (ICD-10)

[a]Comparison is relative to counts of reported death, in previous year's report which was based on incidents reported to CPSC by 12/31/2014. Further review of historic records revealed information supporting adjustments to date of death, vehicle classification (*i.e.*, ATVs with 4 wheels for which the number of wheels was not previously determined or out of scope vehicles such as utility vehicles), and scope. Application of alternative search schemes identified records that could be further associated as multiple reports of the same incident



Source: U.S. Consumer Product Safety Commission: Directorate for Epidemiology/Division of Hazard Analysis. Graph generated using SAS®.
Note: This figure corresponds to the first two columns of Table 2. Reporting for 2013–2015 is ongoing, and Figure 1 does not include data for these years.

Fig. 1. Number of reported ATV-related fatalities by State (1982-2012)

3 Interventions

The ASI, in agreement with the CSPC, created a hands on, interactive rider course. The ASI has more than 2500 certified trainers nationwide that have trained nearly one million ATV owners [1]. The ATV safety training focuses on rider factors such as use of proper clothing, personal protective equipment, helmet use, pre ride check, starting, stopping, traveling up, down and across hillsides as well as managing obstacles. Additional interventions have been developed ranging from protective legislation and public health messaging to rollover protection devices. The agent, host and environment model provide an opportunity to consider strategies aimed at one or more aspect of energy transfer to mitigate or prevent the loss of control event and adverse health outcome. Training is directed at enhancing skills of the host while engineering devices are intended to improve agent performance. The transition of three to four wheel units was in fact, an engineering design change to improve stability of the ATV [3]. Unfortunately the benefits in safety derived through the introduction of the more stable quad bike were lost to the increasing size, power and speed of the ATVs [3]. In recent years more states have allowed ATVs on roadways resulting in more frequent auto-ATV collisions [8].

Legislation ranges widely from state to state and primarily is aimed at protection of youth and restricting ATVs from roadways. Between 2004 and 2014 the number of states allowing limited access to roadways increased from 22 to 35, more often among the

agricultural communities. Despite the decline in numbers of fatalities in recent years the incidence of road related injuries and fatalities is likely to increase with greater numbers of ATVs sharing roadways with other vehicles.

Fig. 2. ATV animal handling safety poster

Public health messaging has been used throughout agriculture communities to increase awareness of ATV related risks and hazards [9]. Investigators collaborated with extension agents and three local communities in Montana to develop an ATV safety tip sheet aimed at the common uses of ATVs on ranches. The four common job tasks included personnel transport, animal handling, fence mending and weed control. Posters and tip sheet were developed with community input and disseminated to all agricultural operations within McCone County, MT, see Fig. 2. Residents were surveyed about the effectiveness of the messages and relevance to their needs. Results from that investigation found that 97% had read the material, 93% felt that content was high quality safety and health information and 78% agreed that the information caused them to think about safety for themselves and/or workers and how to perform activities.

The development of rollover devices has been tested and found to reduce risk for crush injury [10–12]. Devices have become more common place in New Zealand and Australia where quad bikes are used widely in agricultural operations. The interest in stability and safety has led to the improved side by side, caged utility vehicles. Interventions aimed at the agent that reduce reliance on host skills and abilities provided added safety. Testing utility vehicles demonstrated that protected survival space is critical to reducing the probability for injury and that ATVs did not provide equivalent safety [13].

4 Future Studies

Future studies need to assess the effectiveness of training. While the ASI has trained nearly one million ATV owners in the US, an estimated 34 million riders may not have had training of any type. Exposure assessment using ergonomic instrumentation would be helpful to understand the demands and responses of the vehicle operator. Accelerometer and electromyography are useful tools in evaluating position and muscular response to stressors. While the data shows the newer caged utility vehicles are safer, seat belts must be worn and time is needed to determine if injury and fatality rates continue to fall. There remains over 10 million ATVs in the US being used and are not likely to be discarded. Market trends indicate that the new safer designs are gaining sales over the quadbike. The agriculture community is in need of multiple approaches to increase awareness and safety related to ATVs.

References

1. ATV Safety Institute: Industry background. (2016). http://www.atvsafety.org/downloads/INDUSTRY_BACKGROUND.pdf
2. Lagerstrom, E., Gilkey, D., Elenbaas, D., Rosecrance, J.: ATV-related injuries in Montana 2005–2012. Special issue - all-terrain (ATVs, quad bikes) and off-highway (ROVs, UTVs, SSVs, LSVs, LUVs, MUVs, XUVs). Safety **1**(1), 59–70 (2015). doi:10.3390/safety1010059
3. Government Accountability Office: ALL-TERRAIN VEHICLES, how they are used, crashes, and sales of adult-sized vehicles for children's use (2010) GAO-10-418 http://www.gao.gov/products/GAO-10-418

4. Consumer Product Safety Commission: 2015 Annual Report of ATV-Related Deaths and Injuries (2017). https://www.cpsc.gov/content/2015-annual-report-of-atv-related-deaths-and-injuries
5. Helmkamp, J.C., Biddle, E., Marsh, S.M., Campbell, C.R.: The economic burden of all-terrain vehicle related adult deaths in the U.S. workplace, 2003–2006. J. Agric. Saf. Health **18**, 233–243 (2012)
6. Helmkamp, J.C., Marsh, S.M., Aitken, M.E.: Occupational all-terrain vehicle deaths among workers 18 years and older in the United States, 1992–2007. J. Agric. Saf. Health **17**, 147–155 (2011)
7. Lagerstrom, E., Gilkey, D., Magzamen, S., Stallones, L., Rosecrance, J.: Understanding risk factor patterns in ATV fatalities: a recursive partitioning approach. J. Saf. Res. **59**, 23–31 (2016)
8. Consumer Federation of America: ATVs on roadways: a safety crisis (2014). http://www.consumerfed.org/pdfs/ATVs-on-roadways-03-2014.pdf
9. Lagerstrom, E., Hibiske, S., Gilkey, D., Rosecrance, J.: A case study: the development of safety tip sheets for ATV use in ranching. Safety, special issue - all-terrain (ATVs, quad bikes) and off-highway (ROVs, UTVs, SSVs, LSVs, LUVs, MUVs, XUVs). Safety **1**, 84–93 (2015). doi:10.3390/safety1010084
10. Franklin, R., Knight, S., Lower, T.: Mount Isa statement on quad bike safety. (2014). http://www.rrh.org.au/publishedarticles/article_print_2687.pdf
11. Lower, T., Trotter, M.: Adoption of quad bike crush prevention devices on Australian dairy farms. J. Agromed. **19**, 15–26 (2014)
12. Myers, M.: All-terrain vehicle safety-potential effectiveness of the quadbar as a crush prevention device. Safety **2**, 3 (2016). doi:10.3390/safety2010003. http://www.mdpi.com/journal/safety
13. Grzebieta, R., Rechnitzer G., Simmons K., Hicks D., Patton D., Mongiardini M., McIntosh A., Sherry D., Dal Nevo R.: The Australian Terrain Vehicle Assessment Program (ATVAP). In: Proceedings of the 24th International Technical Conference on the Enhanced Safety of Vehicles (ESV) Gothenburg, Sweden, June 8–11, 2015

Study of Spain's Vehicle Fleet Safety Level Based on the Results of Euro NCAP Tests

Oriol Vaquer, Núria Parera(✉), and Alba Fornells

Applus IDIADA, L'Albornar, PO BOX 20, 43710 Santa Oliva, Tarragona, Spain
{oriol.vaquer,nuria.parera}@idiada.com

Abstract. The Spanish Directorate-General for Traffic (DGT) and IDIADA have been working together with the aim of obtaining more data related to the vehicles registered in Spain and their safety. In order to do so, the DGT has updated the Spanish vehicle registration database from 2009 to 2016. In order to correct and homogenize the database, the data has been processed with an algorithm that processes all the information according to the type-approval number and the vehicle's generation type. On the other hand, all the Euro NCAP data from its foundation until 2016 has been uploaded to another database manually according to the vehicle's generation type. Having developed and updated both databases, a matching process was carried out unifying all the registered vehicles with the Euro NCAP results. The database has a sample of 7,499,257 vehicles from 2009 until 2016, of which 99.83% has been normalized according to vehicle commercial name.

Keywords: Safety · Euro NCAP · Database analysis · Star rating

1 Introduction

The passive safety protection of car occupants has improved over recent years. In order to implement safety measures, and objectively evaluate them and their efficiency a program has been developed, Euro NCAP. This program assesses the most sold vehicles in the European market by a set of crash tests in a way that is easily understood by consumers.

With the aim of carrying out broader road safety studies and relating them with the Spanish vehicle fleet, the Euro NCAP star rating has been included as a variable in the Spanish vehicles registration database of the DGT (Spanish Directorate-General for Traffic). One of the outputs of this new database is that it is possible to correlate the Euro NCAP star rating with injuries resulting from real accidents as established in the "Plan de Investigación en Seguridad Vial y movilidad 2013–2016 de la DGT" [1] .

Studies that correlate the Euro NCAP rating with accident injuries have been done before in Europe SARAC II [2], a study of the Advancement of Automotive Medicine [3] and the United States carried out by NHTSA [4]. This kind of project enables the assessment obtained from the real accident analysis results to offer a more realistic vision of vehicle safety and the possibility of monitoring vehicle design improvements over the years [5]. Having a methodology that assesses each vehicle registered with safety

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_23

level and at the same time knowing the potential risk of injury is very important for future investigations in road safety and for consumers' knowledge.

The project reported in this paper will make it possible to have a more homogeneous vehicle registration database with a new classification variable corresponding with the Euro NCAP rating. It will make it easier to study, develop and implement effective measures against accidents on Spanish roads.

2 Methodology

The creation of the database had the aim of unifying the vehicle registration data from the DGT database with high-quality criteria and optimizing vehicle classification. Moreover, the introduction of a new variable, the Euro NCAP rating, makes the database more valuable and a powerful tool in road safety investigation. To achieve both objectives, various steps were carried out: the original vehicle registration database was processed, then it was matched with a vehicle make, model and type list from IDIADA's database and finally it was matched with the Euro NCAP ratings.

First of all the DGT database was filtered with the aim of knowing which vehicles could be classified by the Euro NCAP rating. The methodology was as follows:

The vehicles introduced in the database were the ones registered until end 2013; the sample obtained was 30,672,620 vehicles.

1. As the Euro NCAP protocol mostly applies to cars, the vehicles that were not cars such as mopeds, motorcycles, vans, trucks, buses and other private vehicles were excluded from the database. The total database was reduced and the sample obtained was 22,058,535 cars.
2. The first Euro NCAP tests were carried out in 1997, due to this fact the vehicles registered before that date were deleted from the database obtaining a final sample of 18,133,231cars.
3. Finally, the study only included the cars registered from 2009, obtaining a final database with a sample of 4,684,440 cars.

The following figure represents the unification and assignation process that was carried out during the second stage of the methodology (Fig. 1).

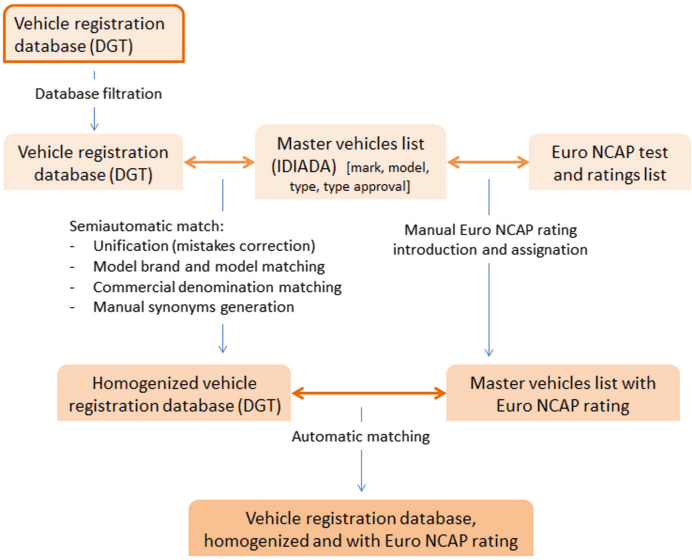


Fig. 1. Methodology schema

2.1 Unification of the Database

Once the sample of the database was reduced (4,684,440 cars), a search of pattern of errors and name duplications was carried out. A random sample of 200 vehicles from the total database was taken in order to find a pattern of errors.

This pattern was applied to the original database but many vehicles were not corrected, as the error was not identified in the original sample. For this reason, a matching procedure of the vehicles brand and model was applied to the database in order to homogenize and correct such errors from the original data.

2.2 Make and Model Matching

This part of the procedure consisted of matching the database unified from DGT (4,684,440 cars) and IDIADA's internal database (master vehicles list). The matching process included the following parameters: mark, model, type and type approval. The matching was carried out as follows:

- Type-approval number: using this parameter 83% of the database was corrected
- Vehicle type: using this parameter 14% of the database was corrected
- Commercial name: 1.8% of the database was corrected by this parameter.

1,2% of the database could not be homogenized. In order to solve this, wrong database parameters, such as the type-approval numbers or the vehicle type, were corrected manually; as a result 97% of the database was normalized.

The 3% that was not corrected was due to some vehicles in the database that were not linked to those parameters, because of that a procedure of normalization was developed for the commercial name and its synonyms.

2.3 Commercial Name Matching

Once the make and model matching was done, in order to normalize the vehicles that had the same commercial name algorithms were created in order to correct errors at the denomination of the commercial names that were not corrected in the previous phase.

The algorithm corrected the names by synonym association; the names were corrected by automatic itineration of the error detected. For example, the vehicles named ("Ibiza", "Ibiza TDI", "Ibiza 1.4") were included in the main commercial name of "Seat Ibiza". The tolerance applied was of 4–5 more characters in the commercial name of the vehicle.

Since not all names could be normalized in the previous step because of spelling mistakes, automatic correcting algorithms were created for each typical mistake detected. Moreover, other vehicle name errors were corrected manually.

2.4 Euro NCAP Rating Introduction and Assignation

A parallel process was carried out at the same time. It consisted of introducing for each vehicle its Euro NCAP rating obtained in the official tests. The list of Euro NCAP's rating was manually assigned to IDIADA's list in order to have all vehicles correlated with one Euro NCAP test and rating.

2.5 Euro NCAP and Vehicle Registration Database Match

Finally, the Spanish vehicle fleet database from the DGT, which was already normalized and homogenized, was matched with IDIADA's list, that had been already matched with the Euro NCAP ratings. A software application developed specially for this project, matched both. The matching of both lists achieved the aim of knowing the corresponding Euro NCAP rating for each Spanish vehicle of the fleet. The matching was by vehicle type and registration date.

2.6 Database Maintenance

New vehicles have been registered from beginning 2014 until end of 2015 and the database has been updated. These new registered vehicles have also been matched with Euro NCAP rating list, which have also been updated with new performed tests. A total of 2,058,076 vehicle registrations were added. This update was carried out following the same method.

3 Results

With this project, the national Spanish vehicle registration database has been homogenized and the Euro NCAP ratings have been linked to it. The current total number of vehicles in the database is 6,742,516.

3.1 Normalization of Vehicle Registration Database

Following the process defined above, 78% of vehicles in the database could be matched with the Master vehicles list using the homologation code (type approval), 19% could be normalized using the vehicle type, 1.4% were normalized with the commercial name and only 1.1% could not be normalized.

Regarding the commercial names homogenization, 99.87% of all vehicles makes and 98.89% of vehicle models could be normalized (Fig. 2).

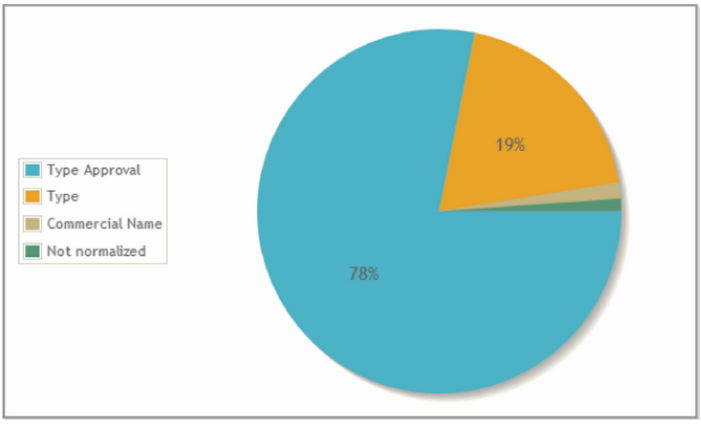


Fig. 2. Normalization of vehicle registration database

3.2 Euro NCAP Rating Assignation

Regarding the Euro NCAP rating association, 84% of the vehicle registration database could be assigned to a Euro NCAP rating using the homologation number (type approval), achieving an exact match. The other 16% of the vehicles have only been approximated to Euro NCAP rating using the first registration date (Fig. 3).

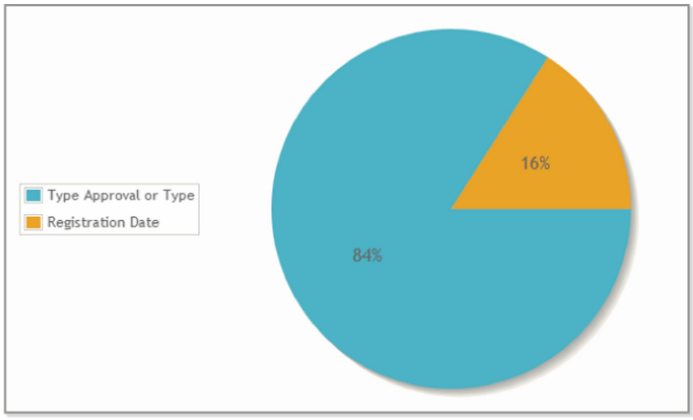


Fig. 3. Euro NCAP ratings assignation method

3.3 Database Study

The new database obtained after this matching process has been studied in order to obtain some conclusions about the Spanish passenger car fleet.

As can be seen in the following figure, 37% of the registered vehicles have been assessed by Euro NCAP before 2009, 52% have been related with Euro NCAP ratings from 2009 or later and 10% had no Euro NCAP association (Fig. 4).

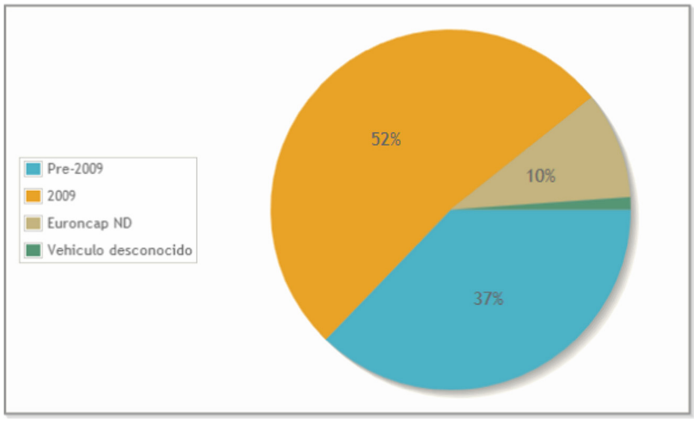


Fig. 4. Euro NCAP rating assignation regarding assessment

Regarding all vehicles which have an Euro NCAP assessment pre-2009, it was observed that 73% have 5 Euro NCAP stars in the adult rating, 24% of vehicles have at least 4 stars and nearly 100% have 3 or more stars.

Regarding the rating in all vehicles with Euro NCAP protocol after 2009, 83% of registered vehicles have 5 Euro NCAP stars and 94% have at least 4 or more stars (Figs. 5 and 6).

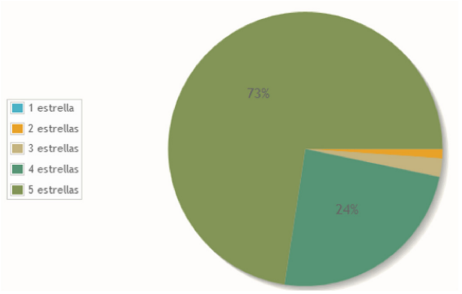


Fig. 5. Euro NCAP rating pre-2009 (adult rating)

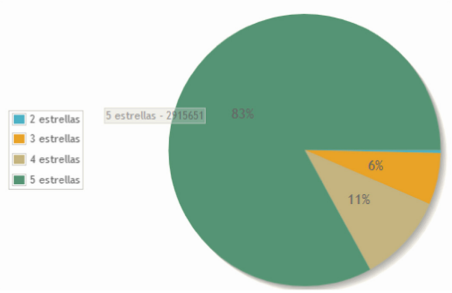


Fig. 6. Euro NCAP rating 2009 (overall rating)

4 Conclusions

The introduction of the Euro NCAP rating in the vehicle registration database from the Traffic General Directorate of Spain is the first step towards offering greater information about safety level to vehicle owners.

The results of this project make it possible to define an approach towards Spanish fleet safety and to identify weak points to try to find a more effective solution.

The obtained data is a good opportunity to carry out accidentology studies and identify the influence of vehicle safety level on accident severity and frequency.

On the other hand, there are many vehicles registered before 1997, which have not been assessed by Euro NCAP. These vehicles, whose safety systems are not at the same level as the new vehicles, are an important part of the Spanish fleet, and this should be taken into account. Therefore, the results from the 6,742,516 vehicles that were studied in this project should be carefully analyzed, as they are a small percentage of the total Spanish fleet. These results do not reflect the reality of Spanish road traffic.

Besides, Euro NCAP assessments are changing every year, becoming more challenging and because of these changes it is very difficult to analyze the evolution of the results over time. Consequently, it is necessary to know in depth the Euro NCAP tests and assessment procedures to correctly interpret the results.

The observed results show that there are more vehicles with higher safety ratings (five stars) after 2009 than before. Therefore, this suggests that the Spanish fleet is becoming increasingly safe.

References

1. Dirección General De Tráfico. Plan de investigación de seguridad vial y movilidad 2013–2016 (2013)
2. European Commission. Transport. Brussels (2003). http://ec.europa.eu/transport/road_safety/projects/doc/saracii.pdf (2012)
3. Segui-Gomez, M., Lopez-Valdes, F.J., Frampton, R.: An evaluation of the EuroNCAP crash test safety ratings in the real world. In: A l'Annual Proceedings/Association for the Advancement of Automotive Medicine. Association for the Advancement of Automotive Medicine, p. 281 (2007)
4. NHTSA. Correlation of NCAP performance with fatality risk in actual head-on collisions. Department of transportation (1994)
5. Lie, A., Tingvall, C.: How do Euro NCAP results correlate with real-life injury risks? A paired comparison study of car-to-car crashes. Traffic Inj. Prev. **3**(4), 288–293 (2002)

Effects of Actigraphically Acquired Sleep Quality on Driving Outcomes in Obstructive Sleep Apnea Patients and Control Drivers: A Naturalistic Study

Nazan Aksan[1(✉)], Robert Marini[1], Jon Tippin[1], Jeffrey Dawson[1], and Matthew Rizzo[2]

[1] University of Iowa, Iowa City, IA, USA
{nazan-aksan,robert-marini,jon-tippin,jeffrey-dawson}@uiowa.edu
[2] University of Nebraska Medical Center, Omaha, NE, USA
Matthew.rizzo@unmc.edu

Abstract. We examined the effects of sleep quality on next day driving outcomes in a 3.5-month naturalistic driving study of 67 OSA and 47 matched control drivers. Sleep quality measures included total sleep time and sleep fragmentation from actigraphy. The driving outcomes included average speed, lateral control, longitudinal control, distraction, attention to driving- and non-driving related tasks. Sleep quality affected next day's driving performance differently for OSA and control drivers. Better sleep quality was associated with better lateral and longitudinal control during highway driving for control drivers. The reverse was true for OSA drivers. Similar effects were also seen in terms of distractions and attention to the driving task. These effects suggest improved sleep leads to greater risky driving and 'activation' among OSA drivers. Collectively, the findings suggest investment in long-term monitoring of sleep quality in commercial vehicle drivers both with and without sleep disorders may help manage safety risks.

Keywords: Obstructive sleep apnea · Sleep quality · Naturalistic driving

1 Introduction

Meta-analytic studies indicate that Obstructive Sleep Apnea (OSA) is associated with increased crash risk [1]. Positive Airway Pressure (PAP), the standard treatment for OSA, appears to mitigate crash risk [2]. Processes that are associated with increased crash risk pre-PAP and decreased risk post-PAP are poorly understood [3–5]. One hypothesis is that sleep quality predicts driver attention and sleepiness which in turn affects driving safety measures such as improved vehicle control. Sleep quality as might be measured with indices of fragmentation, such as wake-minutes-after-sleep-onset (WASO), and adequacy of sleep as measured by total sleep time (TST) are critical for functioning irrespective of sleep disorders [6–13]. Importantly, several sleep surveys indicate that chronic partial sleep restriction, sleeping less than the minimum recommended 7 h of sleep per night, is the norm in the general population [12]. Furthermore, experimental studies show that restricting sleep to even 6.5 h for 14 consecutive days in non-sleep disordered healthy populations can produce performance decrements

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_24

comparable to total sleep deprivation or 24 h without sleep [7]. Importantly, experimental work indicates that there is large individual difference variation in ability to tolerate disruption to sleep and the sources of these differences are poorly understood [14, 15].

The goal of this study was to examine the day-to-day associations of sleep quality as measured by actigraphy with next day's driving outcomes in a 3.5 month naturalistic driving study in patients with OSA and matched control drivers. Recent work based on this sample showed that OSA patients continued to experience greater sleep fragmentation as measured by WASO, lower sleep efficiency despite adequate PAP treatment compared to controls. Furthermore, both OSA and matched control drivers without sleep disorders showed evidence of chronic partial sleep restriction [16]. In other work also based on this sample, we demonstrated that OSA drivers were less alert than controls both pre and post-PAP, and when vehicle speed was less than 45 mph, likely corresponding to city/residential driving environments, they showed less variability in speed and lateral control than controls suggesting lackluster reactivity and less risky driving profiles [5]. While these findings with regard to driving outcomes are consistent with the insidious onset of OSA, possibly associated with adjustment in risk tolerances over time, they also suggest that the effectiveness of PAP on improving sleep quality as well as driving outcomes may be limited when performance is observed in the context of chronic partial sleep restriction.

Because adaptation to both disease and chronic partial sleep restriction is already well-documented but poorly understood [7, 14, 15], it may be important to go beyond the description of general patterns within sleep, within driving or even the associations between sleep quality and driving performance. To that end, we adopted a complete individual differences perspective and quantified the within-person or within-driver bivariate associations between sleep quality metrics from a given night and driving outcomes observed the next day. This analytic framework makes minimal assumptions about what is adequate sleep for any given driver, what is safe with respect to driving outcomes. Rather, the analytic framework simply quantifies the magnitude of prediction from sleep quality to driving outcomes for each driver separately. The within-person associations essentially describe the extent to which changes around the habitual levels of sleep within a given driver vary with changes around the habitual levels of driving outcome measures for that same driver. We then asked whether the magnitude of these within-person or within-driver associations could be predicted by driver characteristics including disease status, cognitive functioning, and impulsivity.

2 Methods

2.1 Sample

Eighty-five OSA and 50 control drivers were recruited into a naturalistic driving study lasting 3.5 months. Control drivers were matched with OSA drivers at the group level on age within 5 years, education within 2 years, and distribution of gender, and county of residence for rural vs. urban driving. Patients met ICSD-2 clinical criteria for OSA [17] and had a Respiratory Distress Index (RDI) >15, while controls had no sleep

complaints and an RDI <5 as confirmed by overnight polysomnography. *Subject drop-out*: Twenty-one of these participants were excluded from the current report for a variety of reasons including: (a) participant's car was incompatible with the instrumented vehicle data acquisition system (IV-DAS, aka 'black-box'), or had technical issues that caused large drop out in either drive or video data (e.g. <20 days of useable driving data); (b) participants dropped out of study too early because of other commitments. The final sample in this report had 67 OSA (27 female; mean-age = 44.26; mean-education = 16.21 years) and 47 control (18 female; mean-age = 45.74 mean-education = 15.45 years) drivers. There were no systematic differences between those included or excluded in terms of disease severity, age, educational level (min p = 0.364). Group/disease status did not increase the likelihood of exclusion from this report (p = 0.744 Fisher's exact).

2.2 Study Procedures

The protocol called for observing OSA drivers with the IV-DAS for 2 weeks before beginning PAP-therapy and for 3 months after. Control drivers were evaluated on the same schedule to assure comparable data acquisition. Participants were given wrist actigraphy (Actiwatch Spectrum Plus, Philips Respironics, Murrysville, PA) throughout 3.5 month study period.

Procedures Pertinent to Driving Data Collection. IV-DAS contained four devices: an internal camera cluster, a GPS, OBD-II, and accelerometers. Two cameras were located beneath the rear view mirror. One pointed forward toward the road (i.e. driver's eye view). The other aimed at the driver face and upper body and car interior. Electronic drive files and associated video clips were transmitted to a remote server daily. Video data collection was triggered based on accelerometer exceedances (at least .35 g's) and a baseline data collection schedule [18]. Only data generated during trigger clips were used in this report. OBD speed, lateral, and longitudinal g's sampled at 1 Hz in each ignition on to off cycle were used in this report.

2.3 Measures

Driving Measures. In order to contextualize the electronic sensor data into likely-highway versus non-highway driving, each drive file was first segmented into consecutive sections where the OBD speed was below 45 mph or 45 mph and above. In each of these sections within a drive, mean OBD speed (in mph), standard deviation in lateral and longitudinal acceleration (in g units) were computed prior to producing averages across trips within a day as the primary vehicle-based performance measures. Clips were evaluated in 20-second segments in three broad domains: safety, exposure, and driver state [18] based on prior work [19–21]. Table 1 provides examples of coded dimensions used in this report. Coders were trained on sample clips illustrating the range of behaviors in each of the dimensions listed in Table 1 until they achieved the minimum inter-rater reliability of Kappa = .61. All coders were blind to actigraphy data. Percentage of 20-second segments in which the driver made driving related gaze movements and rate of

distractions (total number of distractions per 20-second segment) in trigger clips were computed for each driving day.

Table 1. Coded dimensions in 20-second segments from each video clip in two domains of interest

Driver state	Examples
Attention to the driving task	Driving related gaze (e.g., checking mirrors, scanning the road ahead)
Non-driving related attention	Gaze movements to side of the of the road, inside the vehicle
Distraction	Hygiene, cell phone, cognitive, in-vehicle, passenger, eating/drinking, smoking

Cognitive Function and Impulsivity. Extensive battery of cognitive function assessments tapping processing speed (Grooved pegboard, Useful Field of View, Trail-Making-Part-A), memory (Controlled Oral Word Association, Complex Figure-Recall, Auditory Verbal Learning-30 min Recall, Benton Visual Retention Test), visuospatial construction (Complex Figure-Copy, Line Orientation, Block Design), and executive function (Stroop Interference, Wisconsin Card Sort, Trail-Making-Part-B) were administrated prior to PAP and at the end of the study. Administration and scoring details for cognitive function assessments are described elsewhere [22, 23]. Following prior work, the raw scores from each test were standardized and averaged to form a composite scores quantifying participants relative standing in all of these four domains [24]. Participants also completed Zuckerman's Impulsivity scale, and Dula Dangerous Driving Index [25, 26]. The subscale scores were standardized and averaged to form a composite of Impulsive-Risky driving.

Sleep Quality. Actigraphy data were analyzed using the manufacturer's proprietary software set at low threshold [16]. In this report, we used Total Sleep Time (TST) and Wake Minutes After Sleep Onset (WASO) an index of sleep fragmentation (high values indicate restless sleep quality).

Data Reduction and Analysis. Each night's actigraphy data were linked with the following day's driving data from both video (e.g. relative frequency of driving related gaze movements) and electronic vehicle sensors (e.g. mean lateral control when speed < 45 mph). These daily measures were then used to compute Pearson correlations between the sleep metric and driving metric separately for each driver. The driver specific Pearson correlations were then r-to-z transformed and used as outcome measures in multiple regressions with disease status, cognitive function, and impulsivity as predictors.

3 Results

The within-person (or driver) correlations showing the effect of two sleep quality metrics, TST and WASO, on next day's driving outcome measures (both video and vehicle measures) are listed in the left-most column of Table 2. These within-person

correlations served as the outcome measure in multiple regressions and each driver's disease status, impulsivity, and cognitive function served as predictors. The table provides the standardized regression coefficients obtained from these models. The possibility of interactions among the predictors was also evaluated but these were non-significant in every case, hence are not shown in Table 2. Figures 1 and 2 depict the direction of effects from key findings in Table 2. For example, Fig. 1 shows the differences in the magnitude of the within-person associations between sleep quality metrics and next day's driving outcome separately for OSA and control drivers. The first contrast indicates that lateral control in high speed segments had positive associations with TST the night before for OSA drivers but this association was negative for control drivers. This effect indicates that greater TST predicts smaller standard deviations in lateral g's in high speed segments for control drivers consistent with more sleep being associated with improved lateral control. However, the reverse was true for OSA drivers. Similarly, following nights of greater WASO or sleep fragmentation, control drivers' showed larger variability in longitudinal g's in high speed segments consistent with poorer performance. However, the reverse was true for OSA drivers. Similar patterns were also

Table 2. Standardized regression coefficients for predicting the effects of sleep quality on driving outcomes given person-factors: disease status, impulsivity, and cognitive function

Outcomes:	Predictors:		
	Disease status	Impulsivity	Cognitive function
Within-person effect of TST on:			
Lateral control (<45mph)	0.07	0.07	−0.04
Longitudinal control (<45mph)	0.05	−0.01	−0.04
Average Speed (<45mph)	0.16	0.13	0.06
Lateral control (>=45mph)	0.22^{+}	0.02	−0.05
Longitudinal control (>=45mph)	0.00	−0.09	0.01
Average speed (>=45mph)	0.01	−0.15	−0.07
Distraction rate	0.28*	−0.08	0.07
Attention to the driving task	0.19^{+}	−0.02	0.23*
Non-driving related attention	0.32*	0.08	0.14
Within-person effect of WASO on:			
Lateral control (<45mph)	−0.05	−0.16	0.07
Longitudinal control (<45mph)	0.02	−0.22*	−0.02
Average Speed (<45mph)	0.02	−0.13	0.26*
Lateral control (>=45mph)	−0.18^{+}	−0.13	0.14
Longitudinal control (>=45mph)	−0.25*	−0.13	−0.10
Average Speed (>=45mph)	−0.08	−0.02	−0.14
Distraction Rate	−0.20^{+}	0.19^{+}	−0.01
Attention to the driving task	−0.01	−0.05	−0.09
Non-driving related attention	−0.06	0.01	0.05

$^{+}$ $p < .10$, * $p < .05$ or better

Abbrv. TST = Total Sleep Time, WASO = Wake Minutes After Sleep Onset

observed for driver state measures. Longer sleep time was associated with greater distraction rates and greater non-driving related attentional focus the next day for OSA drivers than control drivers. The direction of effects suggests that better sleep quality was associated with an increase in alertness and activity for OSA drivers.

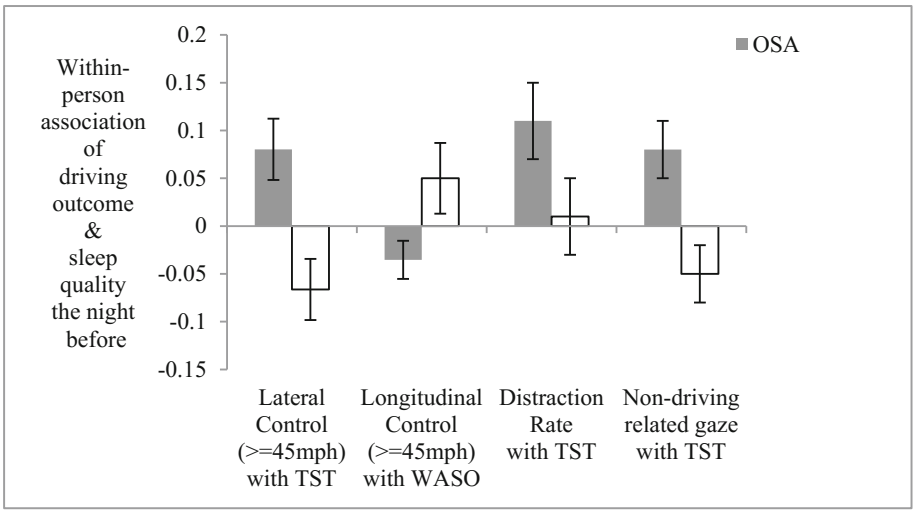


Fig. 1. Within-person associations of driving outcomes with sleep quality the night before on vehicle control and driver state variables for OSA and control drivers.

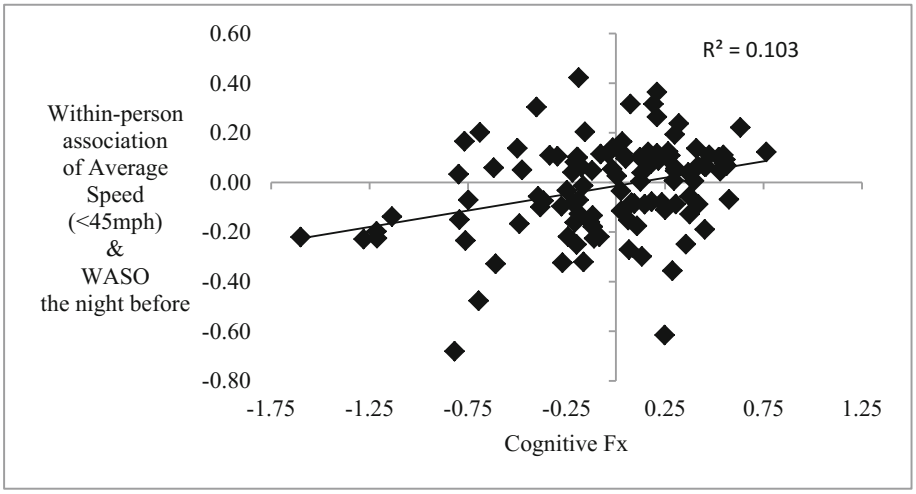


Fig. 2. Higher WASO was associated with lower average speed in low speed drive segments among drivers with low cognitive function. In contrast higher WASO was associated with higher average speed among drivers with high cognitive function.

Some of the models in Table 2 indicated cognitive functioning emerged as a predictor of these associations instead of disease. For example, Fig. 2 shows the effects of sleep fragmentation on average speed in low speed segments varied as a function of driver's cognitive functioning but not disease status. Drivers with higher cognitive function tended to show more positive correlations between fragmentation and average speed. In contrast to drivers with lower cognitive functioning tended to show negative correlations between fragmentation and average speed. Similar patterns were evident in the association between TST and attention to the driving task the next day. Longer sleep was associated with better attention to the driving task and this was particularly true for drivers with higher cognitive function. Impulsivity also predicted the effects of fragmentation on longitudinal control in low speed segments and distraction rate. For example, greater fragmentation was associated with a greater tendency to be distracted during trigger events but this was particularly true for more impulsive than less impulsive drivers.

4 Discussion

The findings showed that sleep quality affected next day's driving performance differently for OSA and control drivers. While longer and more restful sleep was associated with improved performance for control drivers such as lower variability in lateral and longitudinal g's in high speed segments, the reverse was true for OSA drivers. Similar patterns were also observed for driver state measures. Longer sleep time was associated with greater distraction rates and greater attention to non-driving related tasks the next day for OSA drivers than control drivers. These effects suggest improved sleep leads to greater risky driving and 'activation' among OSA drivers for whom good quality sleep is a rare, welcome experience despite adequate treatment with PAP [24].

When the effects of sleep on driving performance was similar for OSA and control drivers, cognitive function emerged as a predictor of the association between sleep quality and next day's driving outcome measures. For example, longer sleep was associated with better attention to the driving task particularly for drivers with higher cognitive function. Similarly, greater fragmentation was associated with lower average speed in low speed segments (city/residential driving) especially for drivers who scored low in cognitive functioning. Those effects suggest that cognitive function can mitigate the adverse effects of sleep quality on driving performance.

The findings together with earlier work from this same naturalistic study have implications for commercial vehicle industry in terms of managing safety risks. The effectiveness of PAP on improving sleep quality as well as driving outcomes may be limited especially when driving performance is observed in the context of chronic partial sleep restriction. The findings thus encourage long-term tracking of sleep quality, specifically fragmentation and total sleep time, in all drivers in addition to compliance with PAP among those with sleep disorders. The findings also suggest driver's cognitive function and impulsivity may be relevant to gauging how changes in sleep are likely to influence their driving performance.

Acknowledgements. The data collection was supported by RO1 HL091917, Real-World Driving Safety in OSA. We thank Sayeh Sabbegh, Michelle Nutting, Lacy Flanagan, Tara Ohrt, Courtney Waite, Nicholas Heisermann, Kelsey Thompson, Katherine Read, Mark Schall, Bryan Fiscus for their diligence and endurance in coding video clips for the study. The analyses presented in this paper were supported by Toyota Collaborative Safety Research Center.

References

1. Tregear, S., Reston, J., Schoelles, K., Phillips, B.: Obstructive sleep apnea and risk of motor vehicle crash: systematic review and meta-analysis. J. Clin. Sleep Med. **5**, 573–581 (2009)
2. Tregear, S., Reston, J., Schoelles, K., Phillips, B.: Continuous positive airway pressure reduces risk of motor vehicle crash among drivers with obstructive sleep apnea: systematic review and meta-analysis. Sleep **33**, 1373–1380 (2010)
3. Connor, J., Nortin, R., Ameratunga, S., et al.: Driver sleepiness and risk of serious injury to car occupants: population based case control study. Br. Med. J. **324**, 1125–1129 (2002)
4. Aksan, N., Dawson, J., Tippin, J., Lee, J.D., Rizzo, M.: Effects of fatigue on real-world driving in diseased and control participants. In: Proceedings of Ninth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, pp. 274–280 (2015)
5. Aksan, N., Marini, R., Dawson, J., Tippin, J., Rizzo, M.: Driving performance and driver state in obstructive sleep apnea: what changes with positive airway pressure. In: Proceedings of Tenth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design. (in press)
6. Banks, S., Dinges, D.F.: Behavioral and physiological consequences of sleep restriction. J. Clin. Sleep Med. **3**, 519–528 (2007)
7. Van Dongen, H.P., Maislin, G., Mullington, J.M., Dinges, D.F.: The cumulative cost of additional wakefulness: dose–response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep **26**, 117–129 (2003)
8. Belenky, G., Wesensten, N.J., Thorne, D.R., Thomas, M.L., Sing, H.C., Redmond, D.P., Balkin, T.J.: Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose–response study. J. Sleep Res. **12**, 1–12 (2003)
9. Goel, N., Rao, H., Durmer, J.S., Dinges, D.F.: Neurocognitive consequences of sleep deprivation. Semin. Neurol. **29**, 320 (2009)
10. Krueger, P.M., Friedman, E.M.: Sleep duration in the United States: A cross-sectional population-based study. Am. J. Epidemiol. **169**, 1052–1063 (2009)
11. Akerstedt, T., Nilsson, P.M.: Sleep as restitution: An introduction. J. Intern. Med. **254**, 6–12 (2003)
12. National Sleep Foundation. https://sleepfoundation.org/sleep-polls-data/sleep-in-america-poll/2005-adult-sleep-habits-and-styles
13. Short, M.A., Banks, S.: The functional impact of sleep deprivation, sleep restriction, and sleep fragmentation. In: Bianchi, M.T. (ed.) Sleep Deprivation and Disease: Effects on the Body, Brain and Behavior, pp. 13–26. Springer, New York (2014)
14. Van Dongen, H.P.A., Baynard, M.D., Maislin, G., Dinges, D.F.: Systematic inter-individual differences in neurobehavioral impairment from sleep loss: Evidence of trait-like differential vulnerability. Sleep **27**, 423–433 (2004)

15. Van Dongen, H., Maislin, G., Dinges, D.F.: Dealing with inter-individual differences in the temporal dynamics of fatigue and performance: Importance and techniques. Aviat. Space Environ. Med. **75**, A147–A154 (2004)
16. Tippin, J., Aksan, N., Dawson, J., Anderson, S., Rizzo, M.: Sleep Remains disturbed in obstructive sleep apnea patients treated with positive airway pressure: A 3-month cohort study using continuous actigraphy. Sleep Med. **24**, 24–31 (2016)
17. Kushida, C.A., Littner, M.R., Hirshkowitz, M., et al.: Practice parameters for the use of continuous and bi-level positive airway pressure devices to treat adult patients with sleep-related breathing disorders. Sleep **29**, 375–380 (2006)
18. Aksan, N., Schall, M., Anderson, S.W., Dawson, J.D., Tippin, J.T., Rizzo M.: Can intermittent video sampling capture individual differences in naturalistic driving. In: Proceedings of Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, pp. 135–141 (2013)
19. Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J., Ramsey, D.J.: The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. Washington D.C, National Highway Traffic Safety Administration (2006)
20. Neale, V.L., Dingus, T.A., Klauer, S.G., Sudweeks, J., Goodman, M.: An overview of the 100-car naturalistic study and findings. In: 19th International Technical Conference on the Enhanced Safety of Vehicles, Washington, D.C. (2005)
21. Wierwille, W.W., Ellsworth, L.A.: Evaluation of driver drowsiness by trained raters. Accid. Anal. Prev. **26**, 571–581 (1994)
22. Lezak, M.D., Howieson, D.B., Loring, D.W.: Neuropsychological Assessment, 4th edn. Oxford University Press, New York (2004)
23. Anderson, S., Aksan, N., Dawson, J., Uc, E., Johnson, A.M., Rizzo, M.: Neuropsychological assessment of driving safety risk in older adults with and without neurologic disease. J. Clin. Exp. Neuropsychol. **34**, 895–905 (2012)
24. Dawson, J., Uc, E., Anderson, S.W., Johnson, A.M., Rizzo, M.: Neuropsychological predictors of driving errors in older adults. J. Am. Geriatr. Soc. **58**, 1090–1096 (2010)
25. Dula, C., Ballard, M.E.: Development and evaluation of a measure of dangerous, aggressive, negative emotional, and risky driving. J. Appl. Soc. Psych. **33**, 263–282 (2003)
26. Zuckerman, M.: The psychobiological model for impulsive unsocialized sensation seeking: a comparative approach. Neuropsychobiology **34**, 125–129 (1996)

Integrating Traffic Safety in Vehicle Routing Solution

Iyad Sahnoon[1,2], Mohamed Shawky[1,3(✉)], and Abdulla Al-Ghafli[1]

[1] Traffic and Patrol Directorate, Abu Dhabi Police, Abu Dhabi, United Arab Emirates
isahnoon@gmail.com, m_shawky132@hotmail.com
[2] Tatweer, Abu Dhabi, United Arab Emirates
[3] Ain Shams University, Cairo, Egypt

Abstract. Traditionally, travel time was considered as the main travel cost in Vehicle Routing Solutions. This paper aims to develop a new travel cost function in the route selection process that integrates the crash risk cost estimation with the travel time cost in a comprehensive routing system. The proposed methodology estimates the cost of both the delay time and crash risk per road segments and intersections in the roadway network. The crash risk was predicted by applying Negative Binomial Regression model with Empirical Bayes function for fatal and injury crashes, separately. Two values of travel cost were estimated one based on the travel time and the other based on the crash risk probability. Five alternatives of combined sharing ratio between the two cost values are investigated to find a single cost value. A case study area was selected from Abu Dhabi roadway network. The optimum route has been decided based on network analyst module in Arc-GIS software package. The results showed that the integrated traffic safety element has an impact on the route selection where at many cases the optimum route has been changed based on the sharing percentage of the crash risk cost in the total journey cost.

Keywords: Routing solution · Travel cost estimation · Delay time cost · Crash risk · Crash cost · Sustainable transportation

1 Introduction

Traffic safety and crash risk exposure has an increasing importance in the sustainability of the transportation system. Traffic safety can play a significant role in advising the road users to select a proper trip route. Currently, most of the route choice modules rely on the estimated travel time or delay as a main variable for travel cost estimation. In many cases, the selected route is not the preferable one for the road user due to other aspects. For more sustainable transportation system, the traffic safety concept should be integrated in the routing selection problem. This concept is powerful to focusing the problem of road safety in the daily lives of people. In addition, implementing traffic safety in routing solution can be utilized as a tool to encourage people for using more safety transportation modes such as public transportation. Especially with the increasing usage of mobile applications in getting information regarding traffic conditions.

Integrated traffic safety in routing selection has been extensively used for hazard material trucks problems only. However, each road user has a degree of risk to be

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_25

involved in a traffic crash given some factors related to the human, environment, traffic and road conditions. A more comprehensive approach for sustainable vehicle routing is achieved when the economic, environmental, and social perspectives are considered during the travel. The crash risk can be expressed as cost value that will be added to the traditional travel time cost. Then, the optimum route can be defined based on the total minimum cost.

This paper aims to develop a new travel cost function based on integrating two cost values of travel time and crash risk involvements. Accordingly, a single travel cost value is estimated for each trip route. Then, the results of the routes considering the new travel cost will be tested and compared to the route that has been selected based on the traditional routing selection considering delay time only to verify its applicability. This approach helps to improve the concept of sustainable transport that is facing many challenges during the trip route process towards safer roads without conflicting the driver's economical interest while executing his/her trip. In this study, the traffic safety related factors that are utilized in the traffic safety cost function include; historical traffic crashes, road characteristics, crash emergency response time, and automated speed enforcement. The roadway network was divided into segments and the costs per segment are estimated. Accordingly, the best route is selected based on the minimum total travel cost value considering all possible alternative routes using a GIS-based environment. It is worth mentioning that the routing algorithm that is implemented in this study is based on the well-known Dijkstra's algorithm in finding shortest path which is ran within the ArcGIS Desktop application.

2 Literature Review

Delay time is considered as the main concern of the individual's travel trip. One of the main reasons of the traffic delay is the traffic congestion where the literature defines two types of traffic congestions on roads: recurrent when it depends on relevant traffic flow; non-recurrent when it depends on external events such as crashes. In many cases, individuals consider the delay risk as money lost (e.g. individual time value and fuel consumption). Delay costs were estimated by multiplying the value of hourly person travel by the average passenger vehicle occupancy rate which was 1.58 occupants in Florida [1]. Schrank and Lomax estimated the time cost of the traveler by $14.60 per hour [2]. Accordingly, traffic congestions caused an additional 4.2 billion hours for travel, resulting in consumption of additional 2.9 billion gallons of fuel corresponding to a congestion cost of $78 billion in USA.

Pradhananga et al., discussed a methodology that aimed to minimize the sum of the population-based (i.e. risk to exposed population) and congestion-based risk cost (i.e. delay) from truck transport in Japan [3]. The problem was solved using Ant Colony System-based algorithms and was compared to the case where each objective is to be taken solely. The comparison showed that the model provides a better alternative to the conventional population-based model as it gives compromised optimal solution avoiding routes that causes large increase of the congestion-based cost. Regarding crash risk integration in routing solution, few studies were found in the literature where the

very most of them were applied on dangerous goods routing application as case studies where the impact of the hazardous material release in case of a crash involvement was assessed on the surrounding population (i.e. human or environment) in addition to the inclusion of other travel costs such as the travel time and direct transport costs [2–13].

Some researchers addressed the relationship between the traffic congestion and road crashes [14, 15] in risk assessment process. Conca et al. analyzed the inter-relationships between frequency of accidents and traffic flow (i.e. Level of Service, LOS) in vehicle routing problem considering a traffic safety element [16]. The case study was applied on the routing of dangerous goods focusing on the reasons that lead to the leakage of dangerous materials. The developed cost function is a combination of travel time costs (i.e. direct and indirect costs) and the risk due to crashes. This new approach of risk assessment integrated the vehicle density, number of lanes, speed, vehicle occupancy, length of the studies roadway segments, weather conditions, and cost of human loss in the cost estimation formula. The results showed the trend of travel cost per unit length in kilometer with and without the safety risk cost at different road traffic densities (vehicle/hour/lane). Two factors were considered; the surrounding population density per kilometer and traffic free flow speed. At high population density areas, the cost of transport considering the traffic safety risk has a minimum value at LOS (C) while for the less densely populated areas the variation of travel cost per kilometer between one LOS to another is minimum. Also, the reduction in free flow speed has relatively lower travel cost.

Samanlioglu [17] developed a multi-objective location routing model in the Marmara region of Turkey. In the mathematical model, three objectives were considered: minimizing the total transportation cost of hazardous materials and fixed costs of treatment, disposal and recycling; minimizing total transportation risk to population along the route; and minimizing total risk to population around the treatment centers. A weighted formulation was developed and computed using CPLEX software to find effective solutions to the problem. The results answered questions related to the locations of these sites, as well as the routing of hazardous waste and waste residues to and from these sites, taking into consideration the technological compatibility issues of wastes and treatment centers, and minimum and maximum capacity requirements of these centers.

Inanloo and Tansel proposed a framework to evaluate transportation networks in regards to their suitability for dangerous goods shipments [18]. Four criteria were considered in development of the network assessment tool: Health risks due to exposure after accidental releases, Delay costs, Trucking expenses, and Proximity to vulnerable areas (e.g. schools and hospitals). ArcGIS software using Python programming were employed in this study as the framework of the analysis. In specific, the delay cost was estimated by multiplying the delay time caused by an accident to the dollar value of travel time delay.

The literature review showed a lack of applying the traffic safety risk assessments in the general vehicle routing solution of individual drivers. For more sustainable transportation network, the traffic safety element should be integrated in the total travel cost function which can change the public's perception regarding the traffic safety when they plan their trips. The current study attempts to develop a new travel cost function to be

applied for individual drivers routing problem. This function will integrate both delay cost and safety risk cost in a weighted form using a GIS-based environment.

3 Case Study

A part of roadway network of Abu Dhabi (AD) city, UAE, is selected as a case study area in the current research to implicate the proposed methodology. The required data for the sample road network were collected and prepared through different tools and software packages including ArcGIS 10.4 and Excel spreadsheets. The employed data include historical traffic fatal and injury crashes for the period from January 2010 till March 2016 (i.e., about six years) for the study area, traffic counts for the road segments and intersections at the peak hour period, road characteristics (i.e. number of lanes, posted speed limit, horizontal curvature, etc.), automated speed enforcement existence, and average crash emergency response time obtained from historical crash data during the same period at each road segment and intersection.

The road network database was created populating each roadway segment and intersection with the described data linking the geo-geometric feature with attributes. It is worth mentioning that a huge effort was put in preparing the road network database to be usable in this study with an extensive manual data collection for some site characteristics (e.g. from Google earth and field visits). Figure 1 shows the case study area where the proposed methodology is applied. It includes a total of 2916 road segments and 52 intersections.

Fig. 1. Case study area

4 Methodology

Figure 2 shows the proposed methodology framework of developing the integrated travel cost function and routing selection solution. The two main components of the total travel cost function are travel time cost (i.e. delay) and crash risk cost.

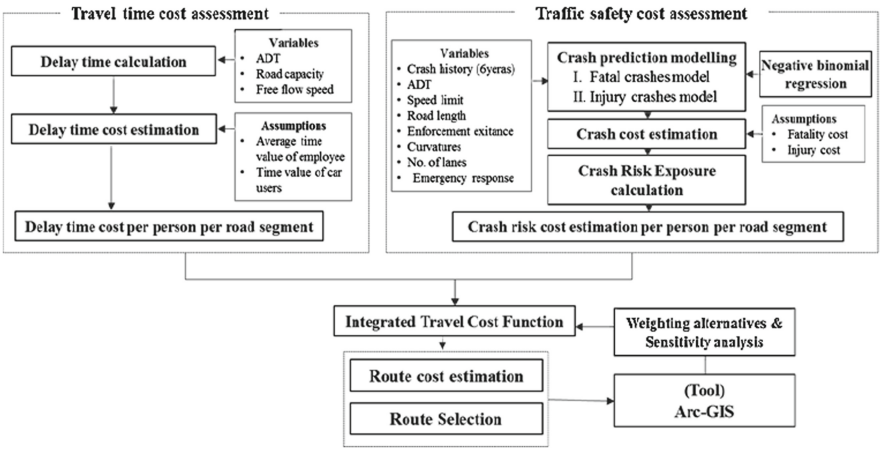


Fig. 2. The proposed methodology of integrated travel cost function

4.1 Travel Time Cost Assessment

The general time delay function was applied to estimate the delay time due to traffic congestion on each road segment and intersection. Traffic counts during morning peak period (i.e. 6:00–9:00 am time) were considered in this study since they represent the worst-case scenario of the traffic condition during the working day. Reducing the delay time is a primary goal of all road users where it directly affects their daily activities in a great manner. Travel time was calculated according to the Bureau of Public Roads function where it assumes that the travel time has a nonlinear relationship with volume/capacity ratio [19]. Then, the delay time is calculated as the difference between the total trip travel time and the free flow travel time. Equations 1 to 4 show the free flow travel time, total trip travel time, and delay time, respectively, for a road segment *(i)* and Eq. 4 shows the total delay time of a certain road corridor *(a)* including the delay at the intersections *(int)*. It is worth mentioning that the delay times at a single intersection were estimated to be the same at all approaches for the sake of simplicity in this study.

$$FFT_i = 3.6 * X_i / S_i \tag{1}$$

$$TT_i = FFT_i \left[1 + \beta \left(\frac{v_i}{c_i} \right)^{\alpha} \right] \tag{2}$$

$$DT_i = (TT_i - FFT_i) \tag{3}$$

$$DT_a = \sum_{i=1}^{i} (TT_i - FFT_i) + d_{int} \tag{4}$$

where;

FFT_i: Free flow travel time for segment (i) in seconds
X_i: Road segment (i) geometric length in meters
S_i: Road segment (i) posted speed limit in km/hr
TT_i: Road segment (i) total travel time in seconds
v_i: Road segment (i) traffic volume in vehicle per hour (peak condition)
c_i: Road segment (i) capacity in vehicle per hour

α and β are constants (the value of α is usually 4 and that of β is 0.15, depending on the roadway type)

DT_i: Delay time on road segment (i) in seconds
DT_a: Delay time on the entire road corridor (a) in seconds
D_{int}: Delay time at intersection in seconds

The delay unit cost is derived from the annual average income per capita in Abu Dhabi based on the published data by the local statistics center in the city. An average value of AED 55 is used to represent the individual average hourly unit cost during the working days. This unit cost value is multiplied by the delay time to find the delay cost value at each road and intersection for the study area. Figure 3 shows a sample of the network with the normalized delay cost values in UAE Dirhams (AED). Four classes/color-codes were used to represent the delay risk on the network. As shown, most of the at-grade urban intersections experience more time delay and hence more delay risk on the drivers.

Fig. 3. Part of the road network with delay cost values (color-coded)

4.2 Crash Risk Cost Assessment

Modeling of Vehicle Crash Prediction. The proposed vehicle crash risk cost estimation was initiated by predicting the average annual crash involvements per driver for each roadway segment and intersection taking into consideration the historical fatal and injury crash data. The Negative Binominal (NB) regression model was utilized for crash prediction as per the recommendation of the well-known Highway Safety Manual (HSM-2010) [20]. Two crash prediction models were developed for two types of crash data; fatal crashes (i.e. crashes resulted with at least one fatal case) and injury crashes (i.e. crashes resulted with at least one injury case regardless of its level). Several independent variables were tested to investigate their significance in the crash prediction models. The dependent variables were set to be the fatal crash and injury crash frequencies, each at a time. The independent variables are: traffic counts for the road segments and intersections during the peak hour period, road characteristics (i.e. number of lanes, posted speed limit, length, horizontal curvature existence, automated speed enforcement existence), and average crash emergency response time obtained from historical crash data during the same period for each road segment and intersection. Recall that the historical crash records cover the last six years in the study area. The NB model can be formulated as shown in Eq. 5:

$$Ln\ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \tag{5}$$

Where; Y is the dependent variable (i.e. number of crashes); X_1, X_2, … X_n are the predictor variables; and β_0, β_1, β_2, …, β_n are the regression coefficients.

The SPSS statistical software package was employed to estimate the model parameters using the customized negative binomial with log link function with the option to estimate the dispersion parameter rather than setting it to the system's default value. It accounts for the over-dispersion that is found in the crash data and quantifies an over-dispersion parameter. Several trials with different alternatives of the tested variables were conducted to determine the contributing ones at a significant level of 90% (p-value ≤ 0.1) which satisfy the author's research purposes. Table 1 shows the results of the parameters estimate for both types of crash models. This table shows that the average crash emergence response time, number of lanes that represents the road category are significant variables in the two model types. While, the value of the posted speed limit and the existence of horizontal curves are significant for fatal crashes model while the automated speed enforcement exitance is significant for injury crashes model.

After using the NB regression to predict the future average annual traffic crash involvements on the study area, the HSM recommends combining the predicted crashes with the observed ones based on the fact of that the observed crashes are not the only base to the safety assessment of the road segment. This method is called the Empirical Bayes (EB) to estimate the number of expected fatal and injury crashes (N_{exp}). The Empirical Bayes (EB) formula is shown in Eq. 6. The result is determined by how much is given to the crash predicted on similar road segments which are represented by a weighted factor (w) during the analysis period (n) as shown in Eq. 7.

$$N_{exp} = w * \left(N_{predicted}\right) + (1 - w) * \left(N_{observed}\right) \tag{6}$$

$$w = \frac{1}{1 + \frac{\sum_{n=1}^{n} N_{predicted,n}}{k}} \tag{7}$$

The weight value depends on the strength of the predicted crash and how different may be traffic safety levels on a specific road segment or intersection compared to the average represented by the crash prediction model. An important parameter in the weight factor is the over dispersion parameter (k) that is used in the calibration of crash prediction model. This factor has the dimension of inverse of unit length. It provides an indication of the statistical reliability of the model. The closer the over dispersion parameter is to zero, the more statistically reliable the crash prediction model.

Table 1. Results of parameters estimate for crash models

Model type	Parameter estimates							
	Parameter	*B*	Std. Error	Hypothesis test		Exp (B) (odd ratio)	95% Wald confidence interval for exp (B)	
				Wald Chi-square	Sig.		Lower	Upper
(I) Fatal crashes model	(Intercept)	−7.156	1.0069	50.514	0.000	0.001	0.000	0.006
	Speed limit	0.024	0.0123	3.959	0.047	1.025	1.000	1.050
	No. of lanes	0.498	0.2265	4.839	0.028	1.646	1.056	2.566
	Avg. response time	0.239	0.0421	32.290	0.000	1.270	1.170	1.380
	Curvature	0.616	0.3886	2.516	0.011	1.852	0.865	3.967
(II) Injury crashes model	(Intercept)	−1.859	0.1821	104.204	0.000	0.156	0.109	0.223
	No. of lanes	0.387	0.0494	61.538	0.000	1.473	1.337	1.622
	Avg. response time	0.185	0.0125	218.434	0.000	1.203	1.174	1.233
	Enforcement	0.278	0.1311	4.513	0.034	1.321	1.022	1.708

Crash Risk Cost Estimation. To estimate the crash risk cost for both fatal and injury crashes, the cost resulted from each type should be included in the estimation function. Based on a study that was prepared by the Department of Transport in AD in year 2010, the cost of a single fatal crash is AED 2.5 million UAE Dirhams and the cost of a single injury crash is around 0.5 million UAE Dirhams. Each cost value is multiplied by the crash rate of each severity type to obtain a single crash risk cost (CRC) at each roadway segment and intersection. The total crash risk cost is calculated as shown in Eq. 8.

$$CRC_i = CR_F * Cost_F + CR_I * Cost_I \tag{8}$$

Where; CRC_i is the crash cost at road segment (i), CR_F is the fatal crash rate on segment (i) calculated previously, $Cost_F$ is the cost of a fatal crash, CR_I is the injury crash rate at road segment (i), and $Cost_I$ is the cost of an injury crash. The crash rate (CR) is calculated as per Eq. 9 where the value is reflected on single vehicle driver or passenger since he/she is the affected person of the crash involvement.

$$CR_i \frac{N_{\exp(i)}}{ADT_i \times 365\, days\, x\, L \times V_{occ}} \tag{9}$$

Where; ADT_i is the average daily traffic of the road segment (i), L is the segment length, V_{occ} is the average vehicle occupancy and is assumed to be 1.4 passenger per vehicle based on an observational survey during the morning peak period at the study area. Note that the above formulas are applicable to the intersection crash rates calculations except that no length measures are used in this case. At the end, the crash risk cost (CRC) is multiplied by the probability of the crash occurrence at each road segment and intersection since the crash has a value of certainty to happen. The probability of crash occurrence (P_{crash})at road segment (i) is calculated by dividing the total crashes (i.e. fatal and injury types) at each segment (i) by the total of all crashes of the study area as shown in Eq. 10. The final form of crash cost estimate (CCE) for roadway segment (i) or intersection (int) is shown in Eq. 11. Figure 4 shows a sample of the network with the crash risk cost values in UAE Dirhams. Based on the probability of crash occurrence per roadway facility the corresponding crash cost estimate per vehicle occupant is shown. One explanation of the low values of CCE is that the number of fatal and injury crashes on the urban network set of Abu Dhabi is relatively lower than the ones existed on the rural highways where the speeding behavior is much noticed as a main contributing factor in the crash severity. However, since the speed limits on the urban streets are relatively lower with many speed control devices implemented at different roadway facilities the probability of fatal and severe injury crashes occurrence is small.

$$P_{crash,i} = \frac{N_{tot\ crash,i}}{\sum_{i=1}^{i} N_{tot\ crash,i}} \tag{10}$$

$$CCE_{i\ or\ int} = CRC_{i\ or\ int} * P_{crash,i\ or\ int} \tag{11}$$

The next section of this study shows the results of each discussed travel cost function (i.e. delay and crash risk) as well as the integrated cost function that combines both costs in a weighted form to test their sensitivity on the outputs.

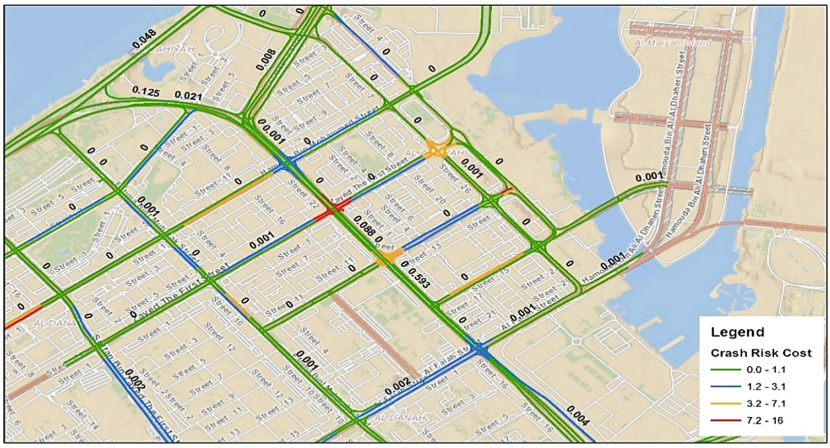


Fig. 4. Part of the road network with crash risk cost values (color-coded)

5 Integrated Travel Cost Function

The developed models for the delay time and crash risk costs are integrated. Different weighted travel cost function alternatives were investigated to test the change in the route selection. The alternatives represent the public interest on one travel cost over the another or equally likely while planning for their journey. Five different weighted alternatives were investigated to test the integrated travel cost function's sensitivity as follows:

- Alternative (1): 20% of delay cost and 80% of crash risk cost
- Alternative (2): 40% of delay cost and 60% of crash risk cost
- Alternative (3): 50% of delay cost and 50% of crash risk cost
- Alternative (4): 60% of delay cost and 40% of crash risk cost
- Alternative (5): 80% of delay cost and 20% of crash risk cost

Network Analyst tool in ArcGIS software package was utilized to implement the routing activity within the selected study area. This tool is based on finding the shortest travel route given a certain travel impedance using Dijkstra's algorithm. The results show optimal routes between the selected origin-destination pairs considering minimum delay cost and crash risk, separately, as well as the integrated total travel cost function that combines the proposed two travel costs.

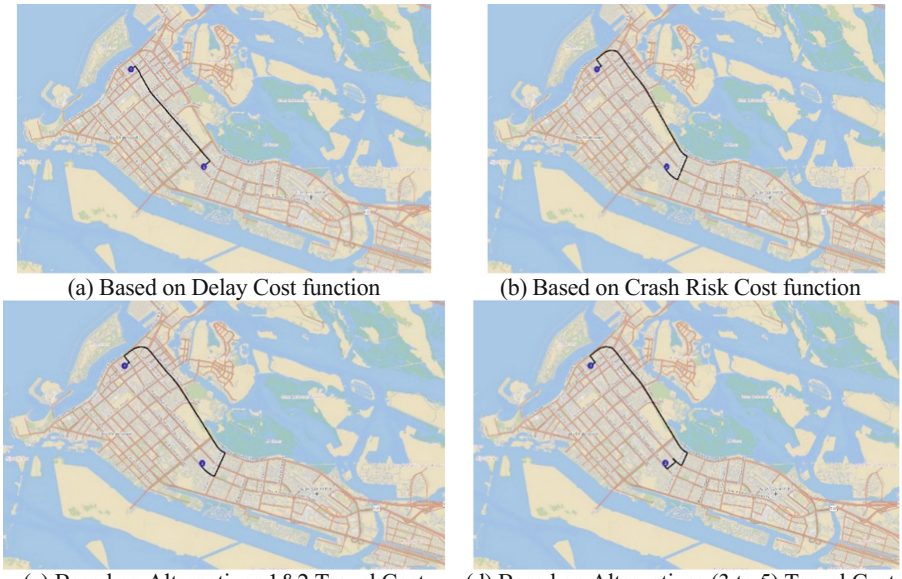
The output results include the optimum route in a map for each case and sharing percentage alternative. Table 2 represents three selected origin-destination pairs within the study area with different route lengths to test the proposed methodology and find the optimal routes considering each weighted alternative. The table compares the total trip cost in (UAE Dirhams) per selected travel cost function and the corresponding trip length. Due to the paper size limitation, one route (i.e. route 3) is selected as a sample for map representation when selecting different travel cost function. The maps shown in Fig. 5 represent the different route decision solutions for each travel cost function for the selected origin-destination points of route 3. The route solutions for delay and crash risk travel costs are different since the selection criteria for both may conflict from the driver's perspective if considered separately. Crash risk cost based-route is approximately 4 km longer compared to the one generated based on the delay travel cost function and this is logic based on the authors experience of the nature of both conflicted travel cost functions and the local driver main interest of minimizing the trip's travel time. The safest route decision is not always the shortest one.

The integrated travel cost functions (i.e. alternatives 1–5) show different route selections based on the weighted alternatives of the two individual travel costs (Fig. 5a and b). The route solution for alternatives 1 and 2 (Fig. 5c) with 80% and 60%, respectively, more weight for traffic crash risk cost shows approximately similar route choice compared to the other alternatives as shown in (Fig. 5d), however, it is less. Note that considering any of the combined travel cost alternatives has a total length that is very close to the one resulted by the delay cost function only for routes (1) and (2) while the case is the opposite for route (3) where the crash risk cost appears to be a leading part in then function. This conclusion proves that the proposed methodology has a strong evidence to select routes that have approximately total trip lengths that are similar to the ones for the shortest path that considers the delay time only. Instead of considering the delay time only in the route

Table 2. Estimated cost (UAE Dirhams) and total route length (km) for three selected routes

Route	Parameter	Delay cost	Crash risk cost	Share ratio (delay cost/crash risk cost)				
				Alternative 1	Alternative 2	Alternative 3	Alternative 4	Alternative 5
1	Estimated cost (AED)	14.2	27.05	35.2	29.9	27.3	24.7	19.4
	Selected route length (Km)	11.4	15.5	11.4	11.4	11.4	11.4	11.4
2	Estimated cost (AED)	10	39	41.3	36.07	30.3	26.5	18.9
	Selected route length (Km)	8.5	12.3	8.3	8.7	8.7	8.7	8.7
3	Estimated cost (AED)	16	51.6	46.3	39.7	36.2	32.3	24.6
	Selected route length (Km)	7.6	11.4	11.4	11.4	11.5	11.5	11.5

selection. On the other hand, the lowest trip travel cost value was found while considering the delay time objective in the route decision for the studied route cases (1 to 3).



(a) Based on Delay Cost function

(b) Based on Crash Risk Cost function

(c) Based on Alternatives 1&2 Travel Cost function

(d) Based on Alternatives (3 to 5) Travel Cost function

Fig. 5. Route decision solutions based on all travel cost functions for route 3

6 Conclusion

Economic and safety are considered as two main components of sustainable transportation system. Currently, travelers use are using applications that rely on minimum travel time (i.e. minimum cost) to find the best route for their trips. The main purpose of this study is to develop an integrated travel cost function that takes into consideration the economic and traffic safety perspectives in vehicle route selection. The travel delay and crash risk costs are the representatives of the economic and traffic safety parts of the problem, respectively.

The delay cost is assessed by calculating the delay time at each road segment and intersections in a roadway network of the selected case study. The general time delay function was applied and the average hourly income of travelers was used to estimate the delay cost. The risk cost to be involved in a traffic crashes was estimated after developing two negative binomial regression models that predict the crash frequency of two types; fatal and injury crashes at each road segment and intersection. Crash frequency data during last six years was employed in the modeling process. The results showed that posted speed limit, number of lanes (i.e. road category), average emergency response time and existence of speed enforcement devices are significant variables in estimating crash involvement risk at a significant level of 90%. Based on the cost value of the fatal and injury crashes, the risk cost of crashes at each road segment was calculated.

Arc-GIS software was utilized as a framework to apply the developed model of the integrated travel cost for a selected study area. The road network was built and populated with the needed data with an extensive effort and time. The network analyst tool that is built in the software was used to calculate the optimal routes for selected origin-destination points on the network taking each developed travel cost as the travel impedance for route selection.

To integrate the delay and crash costs in a single function, different alternatives for sharing percentages were applied and a sensitivity analysis was conducted. Three cases were presented as examples of the routing outputs for different alternatives. The results confirmed the applicability of integrating traffic safety element in routing decision where at most of the cases the total trip length that is very close to the one based on the delay cost function. This conclusion does not conflict with the driver's interest of finding the shortest path for his/her trip. However, after integrating a sharing percentages, the results showed that in many cases the minimum delay time route is the route with the minimum total estimated travel cost.

The findings of this research highlighted the importance of including the traffic safety perspective in journey planning process. This conclusion lead to more sustainable transportation system and put the road safety as a social problem all the society members have a role for improving safety level. Extra recommendation can be made based on this study by introducing many travel criteria beside the delay in the routing computerized applications to allow the road users to select their trip costs depending on their interest and cultural background of the community. Of course, other route selection criteria can be added to the proposed applications such as the tolls, speed cameras avoidance, and others.

References

1. Florida Department of Transportation (FDOT) (2011). http://www.dot.state.fl.us/planning/trends/tc-report/behavior082611.pdf
2. Schrank, D.L., Lomax, T.J.: The 2007 urban mobility report. Texas Transportation Institute, Texas A&M University, Texas (2007)
3. Pradhananga, R., Taniguchi, E., Yamada, T., Qureshi, A.G.: A risk based vehicle routing and scheduling problem with time windows in hazardous material transportation. In: Transportation Research Board 92nd Annual Meeting (No. 13-1865) (2013)
4. Bianco, L., Caramia, M., Giordani, S.: A bilevel flow model for hazmat transportation network design. Transp. Res. Part C Emerg. Technol. **17**(2), 175–196 (2009)
5. Cassini, P.: Road transportation of dangerous goods: quantitative risk assessment and route comparison. J. Hazard. Mater. **61**(1–3), 133–138 (1998)
6. Chakrabarti, U.K., Parikh, J.K.: Route evaluation for hazmat transportation based on total risk–a case of Indian state highways. J. Loss Prev. Process Ind. **24**(5), 524–530 (2011)
7. Dadkar, Y., Jones, D., Nozick, L.: Identifying geographically diverse routes for the transportation of hazardous materials. Transp. Res. E Logist. Transp. Rev. **44**(3), 333–349 (2008)
8. Fabiano, B., Currò, F., Palazzi, E., Pastorino, R.: A framework for risk assessment and decision-making strategies in dangerous good transportation. J. Hazard. Mater. **93**(1), 1–15 (2002)
9. Guo, X., Verma, M.: Choosing vehicle capacity to minimize risk for transporting flammable materials. J. Loss Prev. Process Ind. **23**(2), 220–225 (2010)
10. Rahman, A., Fiondella, L., Lownes, N.E.: A bi-objective approach to evaluate highway routing and regulatory strategies for hazardous materials transportation. J. Transp. Res. Forum **53**(1), 7–22 (2014)
11. Saat, M.R., Werth, C.J., Schaeffer, D., Yoon, H., Barkan, C.P.L.: Environmental risk analysis of hazardous material rail transportation. J. Hazard. Mater. **264**, 560–569 (2014)
12. Yang, J., Li, F., Zhou, J., Zhang, L., Huang, L., Bi, J.: A survey on hazardous materials accidents during road transport in China from 2000 to 2008. J. Hazard. Mater. **184**(1–3), 647–653 (2010)
13. Zografos, K.G., Androutsopoulos, K.N.: A decision support system for integrated hazardous materials routing and emergency response decisions. Transp. Res. C Emerg. Technol. **16**(6), 684–703 (2008)
14. Shefer, D., Rietveld, P.: Congestion and safety on highways: towards an analytical model. Urban Stud. **34**(4), 679–692 (1997)
15. Quddus, M.A., Wang, C., Ison, S.G.: Impact of traffic congestion on road accidents: a spatial analysis of the M25 motorway in England. Accid. Anal. Prev. **41**(4), 798–808 (2009)
16. Conca, A., Ridella, C., Sapori, E.: A risk assessment for road transportation of dangerous goods: a routing solution. Transp. Res. Procedia **14**, 2890–2899 (2016)
17. Samanlioglu, F.: A multi-objective mathematical model for the industrial hazardous waste location-routing problem. Eur. J. Oper. Res. **226**(2), 332–340 (2013)
18. Inanloo, B., Tansel, B.: A transportation network assessment tool for hazardous material cargo routing: weighing exposure health risks, proximity to vulnerable areas, delay costs and trucking expenses. J. Loss Prev. Process Ind. **40**, 266–276 (2016)
19. BPR: Traffic Assignment Manual. U.S. Department of Commerce, Bureau of Public Roads (1964)
20. AASHTO: Highway Safety Manual. Amirian Association of State Highway and Transportation Officials (AASHTO) (2010)

Structural Study of the Side Impact Test with a Fully-Instrumented Pole

Jacob Sanchez(✉), Melcior Malivern, Núria Parera, and Fernando Bernal

Applus IDIADA, L'Albornar, Santa Oliva, PO BOX 20, 43710 Tarragona, Spain
{jsanchez,melcior.malivern,nuria.parera,
fernando.bernal}@idiada.com

Abstract. The aim of the project was to fully instrument the pole which is used in the side impact test with triaxle load cells. This project aims to obtain more information from the vehicle structure performance and behavior during the crash test. A crash test was done to validate the acquisition system implemented at the pole and an in-depth analysis of the data obtained was carried out. A comparison was made between the vehicle absorption mechanisms and structural elements involved. The full instrumentation of the pole makes it possible to obtain a complete view of the interaction between the vehicle and the pole and identify the most critical parts of the vehicle structure and their performance. Also, this study helps us to define the structural parts of the vehicle that need to be improved and the aspects that need to be taken into account to guarantee the safety of the occupants.

Keywords: Crash test · Analysis · Passive safety · Load cells

1 Introduction

The side impact test, also known as the pole crash test, represents a lateral accident against a post or a tree. To carry out this test a vehicle is placed on a sled and is launched against a pole (rigid cylinder) at an angle. In order to obtain more information to better develop the safety of the vehicles, a project was carried out to implement instrumentation in the pole test.

The accident data published by the DGT from 1998 to 2015 in Spain [1] was taken into account to see the effect of developing vehicles with higher safety measures; Fig. 1 below shows the evolution of the frontal, lateral, rear and front-lateral accidents during that period.

Figure 1 shows the comparison between the different accident scenarios and it can be seen that the lateral accidents are the third most common kind of accident in Spain. Although there has been a reduction of lateral accidents from 2006, more improvement to vehicle safety needed to be made.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_26

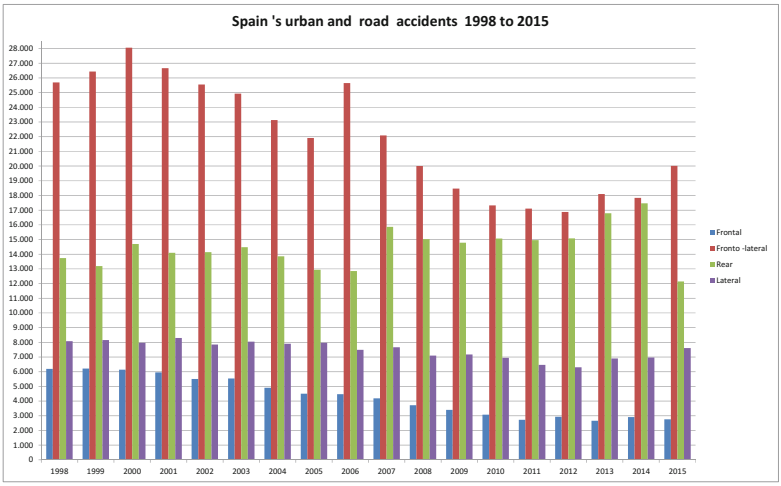


Fig. 1. Road accidents in Spain from 1998 to 2015

From 1998 to 2015 lateral accidents represented 14.1% of total accidents with victims, while frontal accidents were 8.1%, rear were 27.4% and frontal-lateral accidents were 41.5%, as shown in Fig. 2.

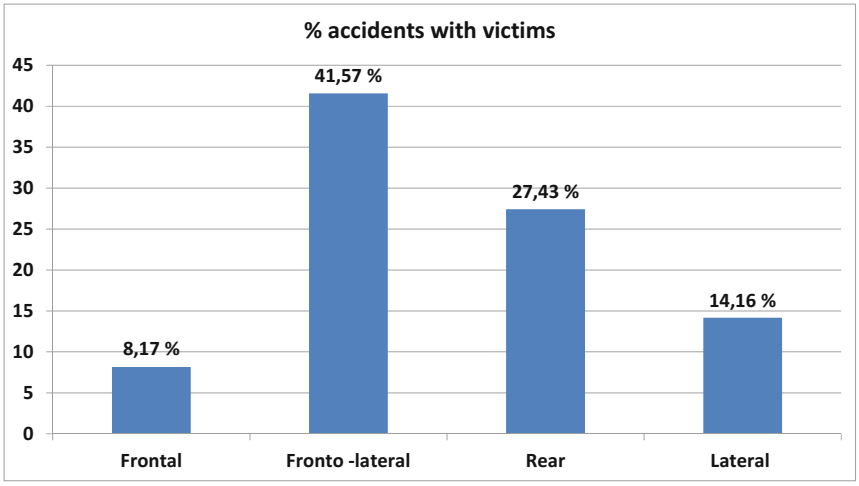


Fig. 2. Percentage of accidents with victims

Although this kind of accident is the most common, the fatalities have been reduced due to the implementation of safety systems such as the lateral airbag or curtain airbag, as shown in Fig. 3.

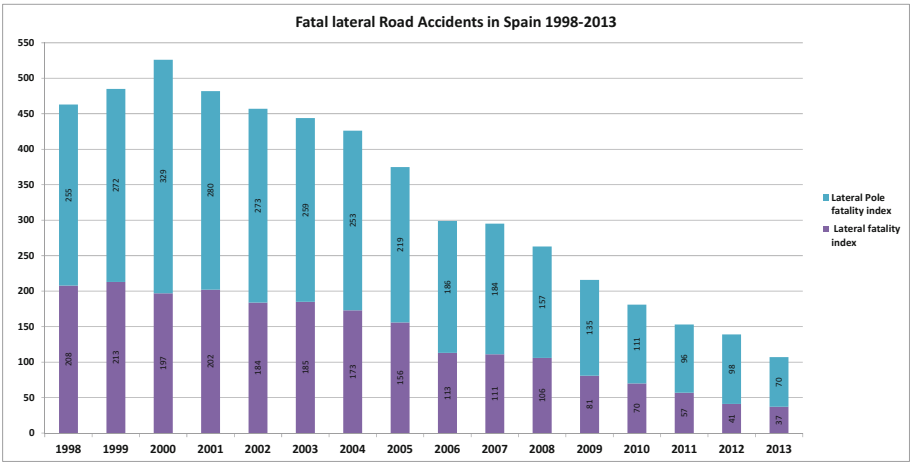


Fig. 3. Fatalities in lateral road accidents

In Fig. 3 it can be seen that the lateral pole crash has a higher percentage of fatalities. The fatality rate for all accidents is 3.4% while in lateral accidents alone it is 1.3%. If we focus our study on lateral accidents, the death rate in accidents against a post or a tree is 8.6%.

2 Pole Test Regulation

The pole crash test represents a lateral accident against a post or a tree. The test is carried out with a vehicle that is placed on a sled and is launched at 19.88 mph (32 km/h) to the pole at an angle of 75º. As shown in Fig. 4.

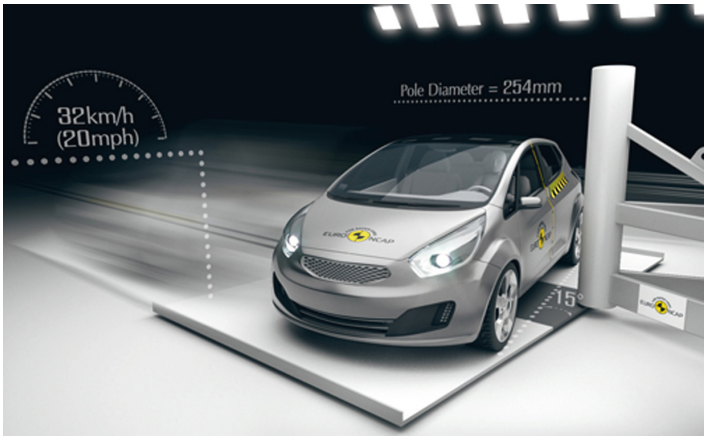


Fig. 4. Euro NCAP crash test pole preparation

The pole crash test is stipulated by the European regulation UNR135, the American regulation FMVSS214 and the Australian regulation ADR 85/00. The NCAP tests that include the pole crash test are: Euro NCAP, U.S NCAP, Latin NCAP, KNCAP and ANCAP.

The structural study and the behavior of the vehicle parts that are in contact with the pole are important to improving vehicles in this kind of test. In order to help our clients to develop and improve their vehicles, IDIADA's Passive Safety department considered instrumenting the pole crash test with load cells. Doing so, more data will be obtained and a better analysis will be made, as there will be better knowledge of which part of the vehicle interacts with the pole.

3 Pole Design

In order to design the new pole with the load cells implementation, several aspects were considered. First of all, it was taken into account that the design of pole and load cells must fulfill the test requirements and also the regulation requirements, such as dimensions and characteristics.

Following the Euro NCAP protocol [2], the pole is defined as a rigid vertical metal structure which begins no more than 102 mm above the lowest point of the tires on the striking side of the test vehicle when the vehicle is loaded and extending at least 100 mm above the highest point of the roof of the test vehicle. It is also indicated that the pole is 254 ± 3 mm in diameter and set off from any mounting surface, such as a barrier or other structure, so that the vehicle will not contact such a mount or support at any time within 100 ms of the initiation of the vehicle to pole contact. The instrumented pole design is shown in Fig. 5.

Fig. 5. Pole design

A new tool design has been made in order to protect the load cell from the impacts. The tool design has a half-moon shape; each load cell of the pole will have this tool, making it possible to change one load cell without dismantling all the pole's load cells.

Furthermore, the design of the pole support was done. As can be seen in Fig. 6, a metallic plate has been designed to reliably fix the pole to the block wall. Also, a rigid structure has been designed with the aim of preventing fractures at the pole's structure.

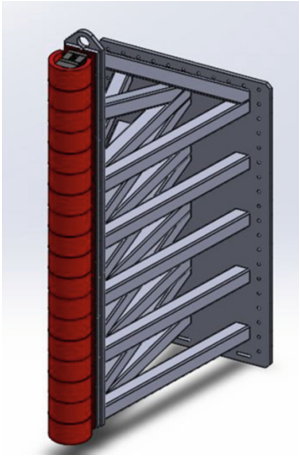


Fig. 6. Pole structure and support

4 Construction of the Instrumented Pole

The construction of the pole was carried out and the result is shown at the following figure:

At the following picture is shown the load cells positioned at the pole and the half-moon shape tool that protects them (Figs. 7 and 8).

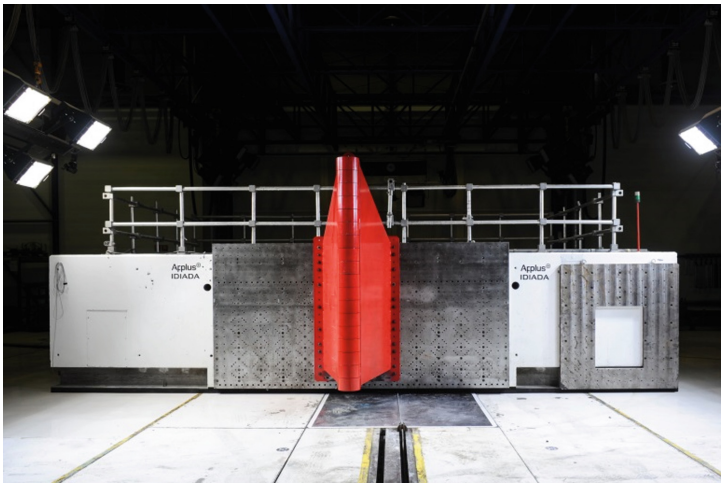


Fig. 7. Pole structure and support

Fig. 8. Pole's load cells and half-moon shape tool

5 Methodology

The data registered by the load cells and the vehicle sensors has been filtered following the regulation SAE J211. In this study it has been applied a filter of 180 CFC for acceleration data and a filter of 60 CFC for force data. The DIAdem software has been used to do the calculations and filtrations of the signals.

Filtering the signals modifies the data but at the same time it allows to have a better reading of the signal due to the elimination of data form the upper peaks and lower peaks. As a result, the obtained signal is clearer and it shows a better results tendency with less noise.

6 Validation Test and Forces and Energy Analysis

A test was carried out in order to validate the instrumentation of the pole. First of all the theoretical kinematic energy was calculated as follows, where the mass of the vehicle was 2067.5 kg and the velocity of the vehicle was 32 km/h (8.88 m/s).

$$Ec = 1/2 * 2066 * (8.88)^2 = 81515.73\,J = 81.515\,kJ \quad (1)$$

The second step was to analyze the accelerations of the vehicle with the accelerometers signal as it can be seen in the Fig. 9. This signal was filtered at 180 CFC following the regulation SAE J 211.

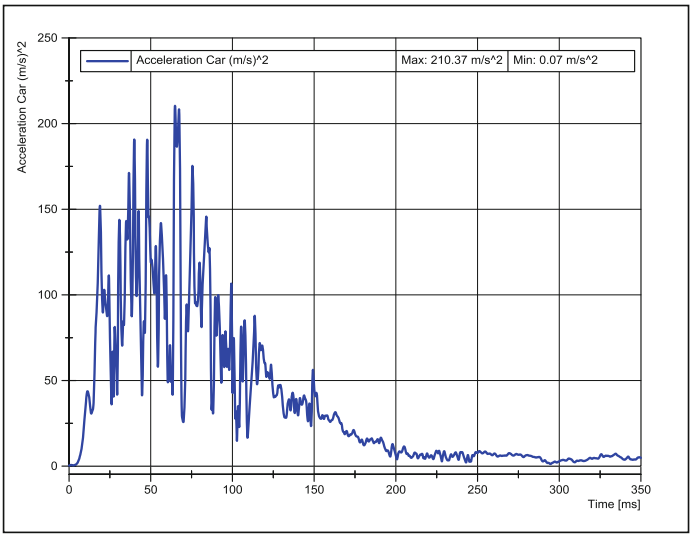


Fig. 9. Accelerations of the vehicle and the pole

In the Fig. 9, is the result of the acceleration of the B-Pillar at right part of the vehicle. The result shows a maximum of 210.37 m/s^2 (21.45 g). At the 130 ms it can be seen that the impact with the pole is almost over, this result is validated by the velocity calculation.

The next step was to calculate the velocity form the accelerometers signals. In Fig. 10 is shown the velocity of the B-pillar during the crash test.

The maximum velocity value was 8.98 m/s that correspond at the real initial velocity of the crash test. This result is obtained by using a tool that measures the velocity of the vehicle at the millisecond zero (offset). This offset is applied at the calculation of the

velocity. At 130 ms approximately the impact with the pole was completed, due to the fact that the vehicle reaches 0 m/s and then obtains a negative velocity. Once the vehicle has impacted the pole, the vehicle is almost stopped afterwards, then vehicle goes backwards and also a small rotation can appear, making the velocity of the vehicle negative.

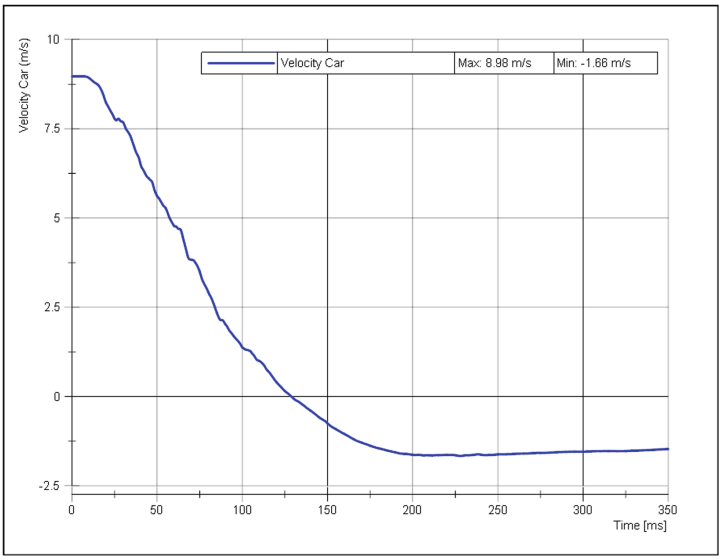


Fig. 10. Velocity of the vehicle at the crash test

To calculate the forces involved at the crash test. The pole's forces are calculated by a sum of the load cells signals in x, y and z directions. The force of the vehicle is calculated by the following formula:

$$F = m * a. \tag{2}$$

The results are shown at the following figure.

The results that are shown in Fig. 11 show a maximum force value from the pole of 284.93 kN and 354.68 kN from the vehicle. The signal of the pole has been filtered at 60 CFC following the regulation SAE J 211. At 75 ms the high force represented by the b-pillar signal represents a vehicle structural part that has broken and the sensor reads the noise but it is not a real force of the test, due to this fact the signals are filtered to avoid noise and false peaks. The false peaks are identified by checking the sub-frame videos where the entire vehicle's major structural parts performances can be seen during the test.

The difference between the pole signal and the vehicle is due to the fact that the vehicle's analysis was done at the B-pillar of the right side and not at the vehicle's center of mass (Tunnel). In order to obtain better results the calculations should be done at the vehicle's center of mass. Due to the deformations that the tunnel suffers during the crash

test it was not possible to assess the test at the vehicle's tunnel. Although, this variation of calculation point the signals from both forces follow the same path and are similar.

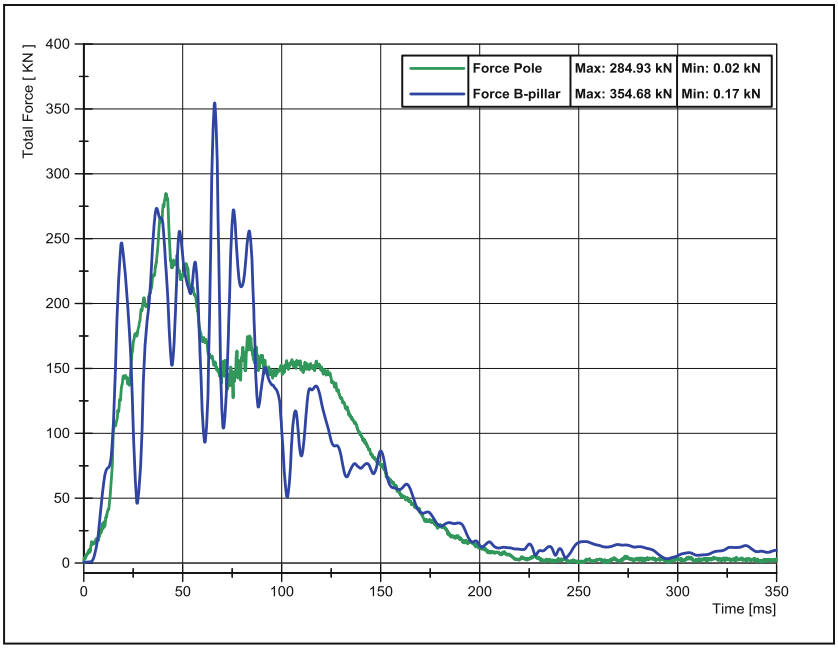


Fig. 11. Right B-pillar and pole forces

Finally the energy (work) calculation was done following the next equation:

$$E_{(t)} = \int_0^t F_{(t)} v_{(t)} d_{(t)}. \tag{3}$$

The results can be seen at the following figure:

The results shown at Fig. 12, validates the good function of the pole acquisition system. The energy signals almost are the same but Due to the different movements that the vehicle has after the maximum intrusion, it can be seen that the signal of the B-pillar keeps a constant value at the end while the Pole energy decreases. The maximum energy value registered by the pole was 85.66 kJ and the energy maximum value of the B-pillar was 89.70 kJ.

The first calculation that was made at the beginning of the analysis showed as a result 81.515 kJ. There is a difference of 8.185 kJ between the real energy (work) and the theoretical one. This difference is due to the results are from the other side of the impact (right) and do not have the same interaction as the left side of the vehicle and also the signals were filtered in order to reduce the noise.

The gap between the pole and the B-pillar is because an amount of energy was dissipated by material deformation, heat and friction [3, 4] that was not taken into account in this analysis. Also, the analysis has been done with the resultant signals form

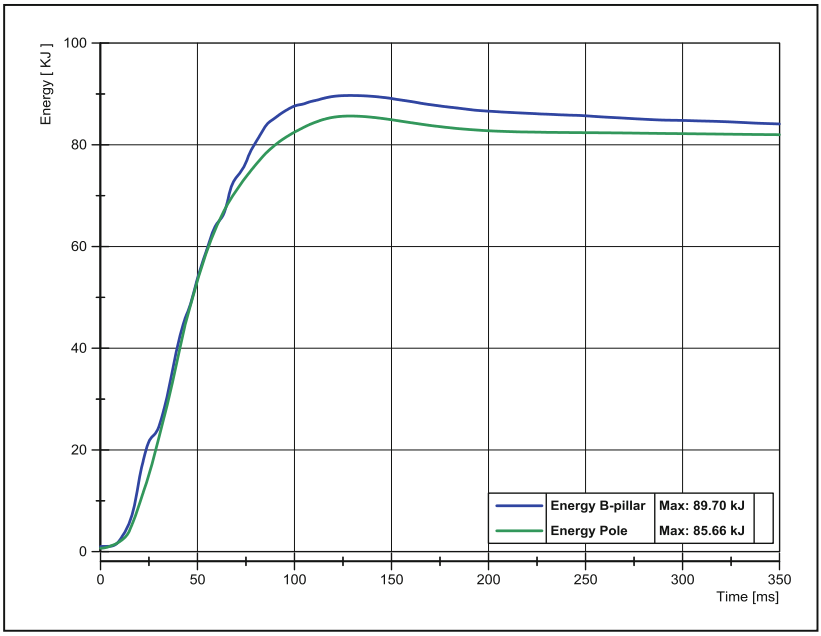


Fig. 12. Right B-pillar and pole forces

the x, y and z axes, but even so the small rotation of the vehicle was not entirely calculated. The rotation happens in the lasts milliseconds where the energy left is really small and does not have much influence at the crash test analysis due to the intrusion has happened a long time ago and the vehicle is not in contact whit the pole anymore.

7 Conclusions

The analysis showed that an instrumented pole is an important improvement and an important tool for the crash test analysis. The acquisition system has been validated and the results obtained were expected. Also, the signals from the pole were easier to assess due to the fact that they had less noise.

The next step for the project is to analyze each element and see how interacts with the load cells of the pole and how this affects at the vehicle structure.

References

1. Dirección General de Tráfico: http://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/publicaciones/anuario-estadistico-accidentes/
2. Oblique pole side impact testing protocol: http://www.euroncap.com/en/for-engineers/protocols/adult-occupant-protection/
3. Kostina, A., Iziumova, A., Plekhov, O.: Energy dissipation and storage in iron under plastic deformation (experimental study and numerical simulation). Frattura ed Integrità Strutturale: Annals 2014: Fracture and Structural Integrity: Annals 2014 8 (2014)
4. Kapoor, R., Nemat-Nasser, S.: Determination of temperature rise during high strain rate deformation. Mech. Mater. **27**(1), 1–12 (1998)

Neurobehavioural Evaluation of Rehabilitation Programs for Dangerous Drivers

Pavel Řezáč[1(✉)], Veronika Kurečková[1], Petr Zámečník[1], Daniel Shaw[2], Milan Brázdil[2], Kristína Czekóová[2], Beáta Špiláková[2], and Miguel Salazar[2]

[1] CDV - Transport Research Centre, Lisenska 33a, 636 00 Brno, Czech Republic
pavel.rezac@cdv.cz
[2] CEITEC MU, Masaryk University, Kamenice 753/5, 625 00 Brno, Czech Republic

Abstract. Study aim to advance understanding of empathy in different driver samples, and to identify ways of modifying such social behaviour in dangerous drivers. This is achieved by evaluating empathy – related driver – rehabilitation programs with a novel neurobehavioral technique. Statistics show that almost 85% of road – traffic accidents in the Czech Republic are caused by the drivers themselves, and approximately 63% of these accidents are due to dangerous driving behaviour (e. g. Speeding, aggressive driving). Developing an effective rehabilitation requires an objective method of evaluation. Our rehabilitation program for dangerous drivers consists of empathy induction. Many studies have revealed associations between reduced empathic awareness and various forms of antisocial behavior. Further research also shows that empathic skills can be enhanced during group therapy. Empathy and anti – social behaviour are associated with specific patterns of brain function and structure. Present study focuses on way, how to evaluate the drivers 'rehabilitation program. We used for this evaluation MRI data combining functional (e.g. partial least squares) and effective connectivity analyses (e.g. dynamic causal modelling) with measures of structural covariance. Combining these multidimensional network level analyses of neuroimaging data with measures of trait empathy and behavioral performance on socio – emotional tasks allow us better comprehensive into brain relationships.

Keywords: Rehabilitation programs for offender drivers · Evaluation of rehabilitation programs · Dangerous drivers introduction

1 Introduction

Statistics show that the drivers themselves cause almost 85% of road-traffic accidents in the Czech Republic, and approximately 63% of these accidents are due to dangerous driving behaviour (e.g., aggressive driving, speeding). Therefore, traffic-safety programs must attempt to modify the roots of such behaviour. Indeed, current European initiatives follow the primary principle behind the goals for DRIVER education (GDE) matrix - that norms, beliefs, and attitudes at the highest psychological level influence driving attitudes, behaviours and decisions at the lower behavioural level. The projects ANDREA [1] and SUPREME [2], for example, both promote the use of rehabilitation programs focused on attitudinal change (e.g., beliefs concerning risky driving), and

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_27

GADGET [3] suggests that a primary role of driver rehabilitation programs should be behavioural adaptation.

High-risk driving behaviours have been associated with personality types [4] and dimensions [5], various demographic characteristics [6], and attitudes [7]. Whilst it is straight forward to acquire cognitive and psychological profiles via self-report, distortions are inevitable; if respondents assume that their replies have an influence on allowed to drive again, social desirability bias will mask true program efficacy. Clearly, then, more objective and comprehensive methods of evaluation are required [8]. In particular, an objective assessment of attitudinal change would present a more appropriate measure of driver-rehabilitation programs.

Developing an effective drivers' rehabilitation program requires an objective method of evaluation, out this has proven to be a particularly problematic task. Typically, such evaluations have involved an assessment of either recidivism rates or accident figures in the period following rehabilitation. While some studies report the positive effects of therapeutic rehabilitation (e.g. [9, 10]), others report no additive effect over basic skills training [11, 12]. These discrepancies likely reflect the inadequacy of recidivism rates and/or accident frequency as criteria by which to measure the success of rehabilitation programs [11] acquire such data is influenced strongly by factors beyond the control of investigators, including the frequency of checks and the willingness of authorities to impose penalties. Further, various confounding factors influence the risk exposure of driver samples; for example, licenses can be revoked for different lengths of time after completing rehabilitation program. Behaving in a socially acceptable manner requires us to modulate our own behaviour according to that of others. This demands an awareness and understanding of others' emotional state - that is, empathy. Empathy, then, be considered a proxy for pro-social behaviour; it provides the affective and motivational basis behind compassion and altruism. Indeed, many studies have revealed associations between reduced empathic awareness and various forms of antisocial behaviour (e.g., conduct disorder, violence; [13–15]). Furthermore, dangerous drivers expressed greater self - (i.e. egocentric) compared with other-regarding preferences. This was reflected in different patterns of brain connectivity; specifically, among neural systems involved in feelings of guilt for acting counter to social values [16–19], feelings of compassion towards others([20, 21], and moral decision-making [22]. The implications of these preliminary results are two-fold: First, they indicate that the rehabilitation of dangerous drivers may benefit from empathy-relater training; secondly, they suggest that when combined with behavioural data, neuroscientific techniques can offer an alternative means with which to evaluate driver-rehabilitation programs that overcomes the pitfalls of existing methods. Research has shown at empathic skills can be enhanced, simply by training individuals to recognize better the emotional states in themselves and others (for a review see [23]). Empathy Induction is one form of empathy training, and has proven to be highly effective in increasing empathic awareness and expression. Typically, it involves focusing attention on the emotions felt by someone else (e.g., a transgressor's attention directed towards the distress felt by the victim). This often involves role taking, the role of a real or fictional figure and is encouraged to focus on the target's dysphoric state. The effectiveness of such training has been demonstrated in a wide variety of contexts (e.g., [24–26]), but most importantly in crime prevention (e.g., [27,

28]), for instance, modified male attitudes towards rape after they heard descriptions of rape from the victim's perspective. Given its success in altering attitudes and fostering prosocial behaviour, it is conceivable that Empathy Induction would help to deter driving behaviour that risks the safety of others. This hasn't been directly tested. Therefore, we propose to incorporate Empathy Induction into driver- rehabilitation programs. Furthermore, we will introduce neuroscientific techniques to evaluate objectively the effectiveness of this novel program; specifically, we will assess the degree to which Empathy Induction is capable of enhancing behavioural and neural indices of empathic awareness in dangerous drivers. In our early study we try to find how can we realize drivers' rehabilitation program and can we evaluate this program.

2 Methods

We collected from previous studies based on desktop analysis. We analysed drivers' rehabilitation programs from abroad, mainly from Germany and Austria. Most of them have decades of experience with the implementation of rehabilitation programs and willingly shared it. At the top of the list of acknowledgements are Bettina Schützhofer, one of the coordinators and supervisors of the Austrian rehabilitation programs; Gregor Bartl, the creator of alles-fuehrerschein.at; Ralf Risser of the Viennese research institution FACTUM; Jürgen Merz of the institute TÜV SÜD; and Harald Meyer of the institution PRO-NON. After this first part of analysis, we establish drivers' rehabilitation program to the Czech Republic. We also try to find the best way how to evaluate this new program. We analysed possibilities of evaluation, which were mentioned on the traffic psychology expert panel. Then we made national expert panel, which were consisted of eight human factors in transportation psychologists, neuroscientists, traffic psychologists. We find the appropriate way to evaluate traffic offenders' progress after attending the rehabilitation programs.

3 Results

We develop structure of rehabilitation programs. We concern on changing drivers behavior mainly thru empathy induction. Rehabilitation program consist of 5 group meetings. First meeting is dedicated to (introduction, presentation of the program, getting to know the participants, formalities and administration). On the second meeting, we solve (analysis of the offenses — serious moving violations or the accidents of participants, or offenses that were the direct cause for the revocation of a driver's license.). Third meeting deals with alcohol and substance abuse. Fourth meeting has a topic (perceived risks in driving, victims of risky driving behaviour). On the last meeting we focus on (stress and fatigue behind the wheel, emotions behind the wheel, etc.), creation of personal strategies to avoid risky behaviour behind the wheel, completing the course, the closure of group topics). As we mentioned we used mainly techniques for Empathy Induction. First technique is "recapitulation of the traffic offense". We use questions like (What led to the offense for this verbal technique what exactly happened? What were the consequences of the offense? What will the consequences be for the rest

of your life? Who is affected by the offense? How else could the whole situation have concluded? Which particular dangerous behaviors are related? How often does the participant exhibit such behavior? How could the situation have been prevented?). Second technique which we use is "Letter to Victims." Participants are asked to recall the most serious accident they caused or an offense in their driving history. Then, they should think of who their risky behavior in this situation hurt or could have harmed. Their task is to write a letter to that person. After the letter is finished, it is their job to try, to the maximum extent possible, to empathize with the victim, and write an imaginary response. After this part we ask participants (What was it like to write the letter? What was the hardest part? What was surprising? Which letter they wrote was better - the letter from the offender, or from the victim?). If participants wish, they can read all or part of the letters, although no one should be to force to do it. In case the only real victims of the risky behavior would be the participants themselves, we ask them to imagine that one of their loved ones is in the role of the victim and might be hurt because the participants endangered themselves. It is appropriate to leave space for all participants in the group to comment and reflect upon the testimony of others. These two techniques lead to encourage traffic offenders' empathy.

The second part of our analysis concerns on evaluation of drivers' rehabilitation program. We use two adjacent 3T functional magnetic resonance scanners, which we combine in "hyperscanning" protocols, whereby functional neuroimaging perform on pairs of dangerous and safe drivers simultaneously as they interact on socioemotional tasks. This will allow us to specify the neural signatures underlying cognitive and affective empathy during real world social contexts. To track the flow of information between the brains of individual drivers, we apply brain-to-brain Granger- causality mapping, inter-brain multi-voxel pattern analysis and classification analyses (e.g., [29]). Developing these novel techniques will allow us to capture the flow of information between brains (i.e. neural "resonance"). Within empathy-related brain systems (see [30]). Given that empathy and antisocial behaviour are associated with specific patterns of brain function and structure, we will also analyse MRI data by combining functional - (e.g. partial least squares) and effective-connectivity analyses (e.g. dynamic causal modelling) with measures of structural covariance. Recently, by combining structural and functional analyses in this way, our group have been able to reveal new insights into the relationship between brain function and structure. Combining these multi-dimensional network-level analyses of neuroimaging data with measures of trait empathy and behavioural performance on socio-emotional tasks would allow a comprehensive investigation into brain-behavioural relationships in the context of driver empathy.

4 Discussion

Our early findings show the best practice to evaluate drivers' rehabilitation programs is by using hyper scanning. Hyperscanning could be performed on pairs of dangerous and safe drivers, allowing for a pre-rehabilitation between subject comparison to specify differences in behavioural and neurophysiological indices of empathic awareness between these driver groups. Differences in behavioural performance on the interactive

tasks will be described in detail, along with differences in the brain systems involved in cognitive and/or affective empathy and the patterns of functional connectivity among them. This research will provide new insights into the neural circuitry underlying two components of empathy during real social interactions. We also want to analysis how this influences driving behaviour. Early findings shows that aggressive drivers will express less affective empathy, and will engage the insula and ACC less than conscience drivers during interaction prior to rehabilitation. This novel evaluative technique will permit the assessment of whether or not specific empathy-related rehabilitation enhances behavioural and/or neurophysiological measures of cognitive and/or affective empathy in dangerous drivers. To achieve the ultimate goal of reducing road-traffic accidents in Czech Republic, the results of this evaluation will be used to update the content of current driver rehabilitation programs in the Czech Republic and serve as a guideline for rehabilitation practices elsewhere throughout Europe.

Good way how to evaluate participants' progress could be basic psychodiagnostic assessment and behavioural measures of empathy. Neuropsychological tests will be administered prior to hyperscanning, providing additional measures for the between subject dangerous vs. safe-driver and within subject pre-, post-rehabilitation comparisons. This will involve a comprehensive neuropsychological battery (including an assessment of cognitive abilities - e.g. memory, executive function). An examination into the nature of interactions among cognitive and affective personality systems associated with empathic awareness, using the Personality Styles and Disorder Inventory (PSDl; [31]); an assessment of trait empathy with the Interpersonal Reactivity Index (IRI; [32]), permitting separate measurements of the cognitive and affective components of empathy. An assessment of impulsivity measuring using the novelty seeking scale of the Temperament and Character Inventory (TCl; [33]); and the ability of drivers to self-regulate their emotions, using The Action Control Scale (ACS-90; [34]) are good to use.

Several well-validated experimental paradigms designed to measure various aspects of socio-emotional capacities. In our eyes, the interactive hyperscanning protocols are useful. These include the Multifaceted Empathy Test (MET; [35]); the Perspective-taking and Director Task [36]; and the Emotional Egocentricity Bias paradigm. Each task should be modified such that members of a pair undergoing scanning will take turns in playing the role of the protagonist or the experimental subject. This will allow us to administer each paradigm within an interactive content. Behavioural and neurophysiological changes should be measured before and after rehabilitation in a sample of dangerous drivers, and once in control (safe) drivers.

Acknowledgements. This article was produced with the financial support of the Czech Science Foundation (GACR), project: Improving driver – rehabilitation programs: A neurobehavioural evaluation of Empathy Induction (16-18261S), on the research infrastructure acquired from the Operation Programme Research and Development for Innovations (CZ.1.05/2.1.00/03.0064).

References

1. Assailly, J.P., Chatenet, F., Hatakka, M., Keskinen, E., Willmes-Lenz, G.: EU-project ''Andrea''. In: Bartl, G. (ed.) Kuratorium für Verkehrssicherheit (2002)
2. Kallberg, V.P.: SUPREME-European Best Practice in Road Safety (2008)
3. Hattaka, M., Keskinen, E., Gregersen, N., Glad, A., Hemetkoski, K.: Results of All-Project GADGET. WP3, European Commission, Brussels (1999)
4. Ulleberg, P.: Personality subtypes of young drivers. Relationship to risk-taking preferences, accident involvement, and response to a traffic safety campaign. Transp. Res. Part F **4**(4), 279–297 (2001)
5. Sarma, K.M., Carey, N.R., Kervick, A.A., Bimpeh, Y.: Psychological factors associated with indices of risky, reckless and cautious driving in a national sample of drivers in the Republic of Ireland. Accid. Anal. Prev. **50**, 1226–1235 (2013)
6. Iliescu, D., Sârbescu, P.: The relationship of dangerous driving with traffic offenses: a study on an adapted measure of dangerous driving. Acci. Anal. Prev. **51**, 33–41 (2013)
7. Nordfjærn, T., Jørgensen, S.H., Rundmo, T.: An investigation of driver attitudes and behaviour in rural and urban areas in Norway. Saf. Sci. **48**(3), 348–356 (2010)
8. Phillips, R.O., Ulleberg, P., Vaa, T.: Meta-analysis of the effect of road safety campaigns on accidents. Acci. Anal. Prev. **43**(3), 1204–1218 (2011)
9. Follmann, W., Heinrich, E., Corvo, D., Muehlensiep, C., Zimmermann, C., Klipp, S., Duenkel, F.: Psychologische Rehabilitations-und Therapiemassnahmen fuer verkehsauffaellige Kraftfahrer. Bergish Gladbach (2008)
10. Masten, S.V., Peck, R.C.: Problem driver remediation: a meta-analysis of the driver improvement literature. J. Saf. Res. **35**(4), 403–425 (2004)
11. Elvik, R., Høye, A., Vaa, T., Sørensen, M.: The Handbook of Road Safety Measures. Emerald Group Publishing Limited, Bingley (2009)
12. Ker, K., Roberts, I.G., Collier, T., Beyer, F.R., Bunn, F., Frost, C.: Post-licence driver education for the prevention of road traffic crashes. Acci. Anal. Prev. **37**(2), 305–313 (2005)
13. Eisenberg, N., Eggum, N.D., Di Giunta, L.: Empathy-related responding: associations with prosocial behavior, aggression, and intergroup relations. Soc. Issues. Policy Rev. **4**(1), 143–180 (2010)
14. Jolliffe, D., Farrington, D.P.: Empathy and offending: a systematic review and meta-analysis. Aggress. Violent Behav. **9**(5), 441–476 (2004)
15. Marshall, L.E., Marshall, W.L.: Empathy and antisocial behaviour. J Forensic Psychiatry Psychol. **22**(5), 742–759 (2011)
16. Zahn, R., de Oliveira-Souza, R., Bramati, I., Garrido, G., Moll, J.: Subgenual cingulate activity reflects individual differences in empathic concern. Neurosci. Lett. **457**(2), 107–110 (2009)
17. Zahn, R., Moll, J., Paiva, M., Garrido, G., Krueger, F., Huey, E.D., Grafman, J.: The neural basis of human social values: evidence from functional MRI. Cereb. Cortex **19**(2), 276–283 (2009)
18. Green, S., Ralph, M.A., Moll, J., Stamatakis, E.A., Grafman, J., Zahn, R.: Selective functional integration between anterior temporal and distinct fronto-mesolimbic regions during guilt and indignation. Neuroimage **52**(4), 1720–1726 (2010)
19. Wiech, K., Kahane, G., Shackel, N., Farias, M., Savulescu, J., Tracey, I.: Cold or calculating? Reduced activity in the subgenual cingulate cortex reflects decreased emotional aversion to harming in counterintuitive utilitarian judgment. Cognition **126**(3), 364–372 (2013)
20. Decety, J., Michalska, K.J., Akitsuki, Y., Lahey, B.B.: Atypical empathic responses in adolescents with aggressive conduct disorder: a functional MRI investigation. Biol. Psychol. **80**(2), 203–211 (2009)

21. Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., Grafman, J.: Human fronto-mesolimbic networks guide decisions about charitable donation. Proc. Natl. Acad. Sci. **103**(42), 15623–15628 (2006)
22. Kahane, G., Wiech, K., Shackel, N., Farias, M., Savulescu, J., Tracey, I.: The neural basis of intuitive and counterintuitive moral judgment. Soc. Cognitive Affect. Neurosci. **7**(4), 393–402 (2012)
23. Feshbach, N.D., Feshbach, S.: Empathy and education. In: Decety, J., Ickes, W. (eds.) Social Neuroscience of Empathy. The MIT Press, Cambridge (2009)
24. Haegerich, T.M., Bottoms, B.L.: Empathy and jurors' decisions in patricide trials involving child sexual assault allegations. Law Human Behav. **24**(4), 421–448 (2000)
25. Beadle, J.N., Sheehan, A.H., Dahlben, B., Gutchess, A.H.: Aging, empathy, and prosociality. J. Gerontol. Series B **70**(2), 213–222 (2015)
26. Barraza, J.A., Zak, P.J.: Empathy toward strangers triggers oxytocin release and subsequent generosity. Ann. N.Y. Acad. Sci. **1167**(1), 182–189 (2009)
27. Berg, D.R., Lonsway, K.A., Fitzgerald, L.F.: Rape prevention education for men: the effectiveness of empathy-induction techniques. J. Coll. Stud. Develop. **40**(3), 219–234 (1999)
28. Foubert, J.D., Perry, B.C.: Creating lasting attitude and behavior change in fraternity members and male student athletes: the qualitative impact of an empathy-based rape prevention program. Violence Against Women **13**(1), 70–86 (2007)
29. Anders, S., Heinzle, J., Weiskopf, N., Ethofer, T., Haynes, J.D.: Flow of affective information between communicating brains. Neuroimage **54**(1), 439–446 (2011)
30. Astolfi, L., Toppi, J., Fallani, F.D., Vecchiato, G., Cincotti, F., Wilke, C.T., Babiloni, F.: Imaging the social brain by simultaneous hyperscanning during subject interaction. IEEE Intell. Syst. **26**(5), 38–45 (2011)
31. Kuhl, J., Kazén, M.: PSSI - Inventář stylů osobnosti a poruch osobnosti. Testcentrum, Praha (2002)
32. Davis, M.H.: Measuring individual differences in empathy: evidence for a multidimensional approach. J. Personal. Soc. Psychol. **44**(1), 113–126 (1983)
33. Cloninger, R., Przybeck, T.R., Svrakic, D.M.: The Temperament and Character Inventory (TCI): A guide to its development and use. Center for Psychobiology of Personality, Washington University, Louis, MO (1994)
34. Kuhl, J.: Action versus state-orientation: psychometric properties of the action control scale (ACS-90). In: Kuhl, J., Beckmann, J. (eds.) Volition and Personality: Action Versus State Orientation, pp. 47–56. Hogrefe Huber, Seattle (1994)
35. Dziobek, I.: Comment: towards a more ecologically valid assessment of empathy. Emot. Rev. **4**(1), 18–19 (2012)
36. Samson, D., Apperley, I.A., Braithwaite, J.J., Andrews, B., Bodley Scott, S.E.: Seeing it their way: evidence for rapid and involuntary computation of what other people see. J. Exp. Psychol. **36**(5), 1255–1266 (2010)

Different Learning Conditions and Their Impact on Hazard Perception Training

Jana Hilz(✉), Sarah Malone, and Roland Brünken

Saarland University, Campus A4 2, 66123 Saarbrücken, Germany
{j.hilz,s.malone,r.bruenken}@mx.uni-saarland.de

Abstract. In the present study, we applied the 4C/ID model to the hazard perception domain and investigated the effectiveness of two of its components. A computer-based learning environment with animated traffic scenarios and supportive and procedural information was developed. We employed a 2 by 2 factorial design with the between-subjects factors supportive and procedural information. The results indicate that the two information types seem to foster different components of hazard perception performance: either speed or accuracy.

Keywords: Novice drivers · Hazard perception · Computerized training · 4C/ID-Model

1 Introduction

Especially within the first period of solo driving, young novice drivers are more likely to be involved in an accident than experienced drivers. Beside age-associated risk factors like for example, unfavorable attitudes, experience-related risk factors such as poor hazard perception skills contribute to their high crash involvement.

One option to address this issue is to improve driving instruction via computer-based self-learning applications. Research has demonstrated that computer-based trainings (CBTs) can be effective tools for teaching safety relevant skills, like for example hazard perception [1]. Hazard perception is the ability to anticipate potentially dangerous situations on the road [2]. This ability can be measured with Hazard Perception Tests (HPTs). Common HPTs require participants to respond as quick as possible to upcoming hazards depicted in real traffic video-based scenarios. Improving hazard perception ought to be an objective of a training intervention, as it is positively related with crash risk and negatively with driving experience [2].

Training interventions aimed at improving hazard perception skills differ considerably regarding length and contentual aspects. Whereas several trainings [3] last approximately half an hour, other interventions take several days [4] or even weeks [5]. The most trainings can be completed within 90 min.

Hazard perception trainings also differ regarding the presentation of traffic scenarios and the type of practice and test items used. Most common in learning environments are dynamic scenarios filmed in real traffic. In some trainings participants are instructed to comment on hazards depicted in traffic scenarios and to judge their own driving

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_28

performance. Another task format are localization tasks in which participants are required to indicate hazards via mouse click. Often, the applied learning tasks are similar to the simple button-push tasks used in common HPTs.

Existing trainings also differ regarding their didactical design. A strategy that has not been used until now for designing a computerized application for driver training is the *Four-Components Instructional Design (4C/ID) Model* [6]. This design approach was developed to facilitate learning of complex skills and has proven to be successful for example in air traffic control, computer programming or medical education. Since hazard perception is a complex cognitive skill the 4C/ID model should be applicable to this domain as well.

The 4C/ID model states *learning tasks*, *supportive* and *procedural information*, and *part-task practice* as crucial elements for well-designed learning environments aimed at teaching complex skills. Key components of this approach are authentic learning tasks that provide learners with meaningful whole-task experiences. Essential information is presented in two ways: as *supportive information*, which promotes schema acquisition, and as *procedural information*, fostering automatization. As a first step, supportive information is given for example, as a text, which helps learners to perform the non-recurrent (non-routine) aspects of a learning task (e.g. rating the risk potential of a specific traffic constellation) that are related to problem solving and reasoning. Non-recurrent aspects refer to skills that will be performed schema-based after the training. Procedural information on the other hand, relate to recurrent (routine) aspects of a learning task (e.g. monitoring the area around a crosswalk while approaching) and is presented not until then it is needed to accomplish the task. Recurrent aspects relates to skills that will be performed rule-based after the training. Part-task practice is the fourth component of the 4C/ID model and pertains to additional practice of routine aspects that learners need to develop to a high level of automaticity (e.g. glance across the shoulder when turning right).

The aim of the present study was to apply the 4C/ID model to the hazard perception domain and to investigate the effectiveness of two of its components; supportive and procedural information. Animated traffic scenarios were presented and learners were supported either by two types of learning information or only by one information type (either supportive or procedural information). Textual explanations regarding different types of possible traffic hazards were presented before learners watched a corresponding animated traffic scenario (supportive information). Visually offered highlighting of hazard cues and corresponding auditory voice-over were presented as part of the animation (procedural information).

According to the 4C/ID model it is hypothesized that the availability of supportive and procedural information leads to better learning outcomes than learning with only one type of learning information or no information at all. If they receive supportive information, participants are expected to identify more hazard cues and to respond faster to them than attendants learning without that information. Participants provided with procedural information will detect more hazards and will react quicker to them than trainees learning without that type of information.

2 Method

2.1 Design

To assess which kind of learning information is most suitable for teaching hazard perception skills, we employed a 2 by 2 factorial design with the two between-subjects factors supportive information (with vs. without) and procedural information (with vs. without). The central dependent variables for learning effects were the number of correct identified hazards (accuracy) and participants' reaction times.

2.2 Participants

Ninety-one participants (55 male; $M_{age} = 16.9$, $SD = 2.58$) were randomly assigned to four groups. All of them held no driver license and had no driving experience at all.

2.3 Material

Prior Knowledge Test. To assess participants' prior knowledge a test with five multiple-choice items out of the German driving theory test was applied.

Hazard Perception Training. A CBT with nine learning tasks (see Table 1) to teach hazard perception skills was developed. Each learning task started with showing a picture of the driving scene from the car driver perspective and mentioning the topic of the learning unit. Thereafter, the program provided text-based descriptions of hazardous events and information, which safe drivers use to timely identify hazardous cues and to deal with them properly (supportive information). Participants that were not provided

Table 1. Topics and brief descriptions of the learning units within the training program

	Topic learning unit	Scenario description
1	Stops (Tram, Bus, Cab)	Approaching tram station; pedestrian steps into drivers lane and car ahead brakes (urban)
2	Cyclists	Car ahead brakes because of cyclist ahead (urban)
3	Lane change	Truck ahead suddenly merges into drivers lane (highway)
4	Sharp bends	Slow vehicle (hidden) in sharp bend to the right (rural)
5	Intersection with traffic lights	Approaching intersection; traffic lights turning from green into yellow; car ahead suddenly stops (urban)
6	Entrances/gateways	Vehicle in front enters gas station; vehicle ahead is braking (urban)
7	Crosswalks	Crosswalk at the exit of a traffic circle; pedestrians crossing (urban)
8	Turning left	Car ahead indicates to turn left; traffic in the opposing lane has right of way (urban)
9	Overtaking	Being overtaken by another car that suddenly goes back into lane because of oncoming traffic (rural)

with supportive information (SI) received alternative text-based learning information, which was unrelated to the hazardous event of the learning unit. Minimum learning time was system-paced. Participants had to spend a minimum of 90 s on the supportive information. Thereafter, a continue button came up and attendants could go on with the task or dwell upon the supportive information. Having read the supportive information the participants watched the corresponding animated traffic scenario. As the 4C/ID model states procedural information (PI) as one crucial factor for learning complex skills half of the participants received additional information while watching the animation. Hazardous areas or objects were signalized and a corresponding auditory voice-over was provided to the recipients. Attendants not provided with PI did not get any hint while watching the animated scenarios. Participants could scrutinize the hazard scenarios only once.

Having completed the learning task, participants had to show off in a second animated traffic scenario what they had learned. They were required to respond as quickly as possible to upcoming hazards by pressing the space bar.

Factual Knowledge and Transfer Test. After the learning session the participants had to complete a second multiple-choice test, which consisted of eight items (also obtained from the formal driving test). This test covered issues partly addressed by the learning tasks and was administered to quantify factual learning.

In addition a transfer test, containing six animated traffic scenarios (with a total of eight hazards included) was applied to assess whether the participants were able to transfer their newly acquired skills to other traffic situations. Two out of the six scenarios contained no hazard cues and were included as distractors. Five presented hazardous situations called for near transfer and one for far transfer. Near transfer scenarios were similar to scenarios that had been trained previously and the far transfer scenario, was not similar to those that had been trained. Transfer test items required the participants to press a button as fast as possible as they became aware of emerging hazards within the animated traffic scenarios. For each correctly solved item, they obtained a point (maximum score of eight).

2.4 Procedure

Participants were tested in groups up to a maximum of 15 individuals in the computer laboratory of the department of education at Saarland University or in schools and sports clubs that offered suitable rooms for testing. After a short welcome attendants were randomly assigned to one of the four learning conditions. The assessment was entirely computer-based. At first, participants had to indicate demographic information (e.g. age, gender, education level, potential driving experiences etc.) and had to work through a factual knowledge test applied to assess their driving related prior knowledge. Following this, they had to complete the nine learning tasks with a minimum time for learning restricted by the system. Items were presented in random order. Finally, participants were required to complete a posttest (factual knowledge test and transfer test). Test sessions lasted approximately 60 min and attendants received five Euros for their participation.

3 Results

To assess their prior knowledge attendants completed five multiple-choice items out of the German driving theory test. An ANOVA showed no significant differences between the four groups ($F_{3,87} = 1.12$; $p \geq .05$). Moreover, participants had low test scores, which means they had rather few driving related prior knowledge ($M = 1.29$; $SD = .89$).

After each learning task, participants had to demonstrate in an animated traffic scenario what they had learned by pressing a response key as soon as becoming aware of an upcoming hazard. If they pushed the button within a predefined time window, which started at the time the hazard first becomes visible and stopped before the point of time when a collision would be likely to occur, they received a point for their answer (accuracy). The time between hazard onset and key press was also recorded as a dependent variable (reaction time).

Descriptive data (see Table 2) for the z-standardized reaction times demonstrate the more learning information participants received (supportive and procedural information on the one hand and on the other hand supportive information is more extensive than procedural information) the faster they reacted to upcoming hazards. As it was expected that the more extensive the provided learning information is the better are the learning outcomes, a one-way ANOVA was conducted. The analysis revealed a significant difference for the reaction times between the four groups ($F_{3,87} = 4.54$; $p = .005$; $\eta_p^2 = .14$). This result was confirmed by a significant linear trend across the four conditions ($F_{1,87} = 13.18$; $p < .001$; $\eta_p^2 = .13$). To examine whether both factors contribute to this trend an ANOVA was conducted. The analysis revealed a significant main effect for the factor supportive information ($F_{1,87} = 9.82$; $p = .002$; $\eta_p^2 = .10$). No main effect for the factor procedural information ($F < 1$) and no interaction effect ($F < 1$) could be found. If they were provided with supportive information, participants responded faster to emerging hazards in the traffic scenarios.

Table 2. Descriptive results for the performance in the training

	SI and PI	Only SI	Only PI	No SI, no PI
	$n = 24$	$n = 21$	$n = 23$	$n = 23$
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Reaction time[a]	−.44 (.77)	−.23 (1.03)	.04 (.70)	.51 (1.16)
Accuracy[b]	8.67 (.76)	8.29 (.90)	8.65 (.49)	8.26 (.92)

[a]z-standardized; [b]max. = 9; SI supportive information; PI procedural information

A significant linear trend for the dependent variable accuracy could not been demonstrated ($F < 1$). An ANOVA for accuracy revealed no main effect for the factor supportive information ($F < 1$), but indicated a main effect for procedural information ($F_{1,87} = 5.53$; $p = .021$; $\eta^2 = .06$). No interaction effect could be demonstrated ($F < 1$). If they were presented with procedural information, participants were more likely to identify hazard cues.

Following the learning session participants completed an eight items multiple-choice test. This test captures topics from the learning tasks and was administered to quantify factual learning. Additionally, attendants were required to complete a hazard perception test with six animated traffic scenarios aimed at testing near and far transfer of learning.

No significant differences between the four groups were found for the performance in the factual knowledge test ($F < 1$). Moreover, no significant group differences could be demonstrated for the accuracy ($F < 1$) and the reaction times ($F < 1$) in the transfer test; both for near and far transfer.

4 Discussion

The 4C/ID approach states four components as crucial for learning environments aiming at complex skills acquisition. The aim of the present study was to apply the 4C/ID model to the hazard perception domain and to investigate if supportive and procedural information are necessary to teach hazard perception skills.

According to the 4C/ID approach, it was assumed that participants would benefit from supportive information regarding their performance in the test items. This hypothesis could be confirmed, partially. Trainees that learned with supportive information responded faster to upcoming hazards but did not identify more hazard cues than participants who were not provided with that information type.

Additionally it was presumed, that those participants who received procedural information would perform better in the test than those who were not provided with that type of information. This hypothesis could be affirmed in part. If procedural information was presented, attendants detected more hazard cues in the test, but did not react faster to them as if that type of learning information was not included in the learning program.

Results indicate that the two types of information seem to foster different components of hazard perception performance: either speed or accuracy. As supportive information included many facts about any components that can turn into a hazard in a distinct traffic situation, they might have prepared the learners to anticipate risks as soon as they encountered such a traffic situation in a test item. Already expecting a hazard, the participants were able to react immediately when a hazard cue effectively emerged. However, procedural information might have guided learners' attention to those areas on screen where a hazardous situation was occurring (because the same areas had been highlighted in the learning scenarios), which has probably enhanced the number of hazard cues that were detected.

Contrary to our expectations, groups did not differ regarding their performance in the factual knowledge and transfer test. It might be that one practice item per topic and a total learning time of 45 min is not enough to elicit transfer of learning.

5 Conclusion and Practical Implications

The results of the present study support the assumption that the 4C/ID model is applicable to the hazard perception domain. Moreover, the results support the assumption that computer-based learning environments aimed at hazard perception skills acquisition

should provide supportive and procedural information. At least if the target group for the training intervention are beginner drivers.

In a next step this information can be used to construct an adaptive training program. Within this training program learning information can be reduced stepwise with an increase in trainees procedural knowledge.

Acknowledgments. The present study was part of the research project "eLearning-Unterrichtskonzepte-Fahranfänger" funded by the German Federal Highway Research Institute (BASt; Grant: FE 82.0564).

References

1. Petzoldt, T., Weiß, T., Franke, T., Krems, J.F., Bannert, M.: Can driver education be improved by computer based training of cognitive skills? Accid. Anal. Prev. **50**, 1185–1192 (2013)
2. Horswill, M.S., Taylor, K., Newnam, S., Wetton, M., Hill, A.: Even highly experienced drivers benefit from a brief hazard perception training intervention. Accid. Anal. Prev. **52**, 100–110 (2013)
3. McKenna, F.P., Horswill, M.S., Alexander, J.L.: Does anticipation training affect drivers' risk taking? J. Exp. Psychol. **12**(1), 1–10 (2006)
4. Isler, R.B., Starkey, N.J., Sheppard, P.: Effects of higher-order driving skill training on young, inexperienced drivers' on-road driving performance. Accid. Anal. Prev. **43**(5), 1818–1827 (2011)
5. Allen, W., Park, G., Terrace, S., Grant, J. (eds.): Detecting transfer of training through simulator scenario design: a novice driver training study. In: Proceedings of the Sixth International Driving Symposium on Human Factors in Driving Assessment, Training, and Vehicle Design; 2011; Lake Tahoe, CA
6. van Merriënboer, J.J.G., Kirschner, P.A.: Ten steps to complex learning. A systematic approach to four-component instructional design. Lawrence Erlbaum Association, Mahwah (2007)

Effects of Using a Cell Phone on Gaze Movements During Simulated Car Driving: Hand-Held and Hands-Free Conditions

Bethânya Graick Carizio[1,3(✉)], Gisele Chiozi Gotardi[2,3], Juliana Cristina de Ângelo[1,3], Paula Fávaro Polastri[3], Fábio Augusto Barbieri[4], Luis Carlos Paschoarelli[1,5], and Sérgio Tosi Rodrigues[1,3]

[1] Graduate Program in Design, Faculty of Architecture, Arts, and Communication, UNESP - Univ. Estadual Paulista, Bauru, SP, Brazil
bethanya.carizio@yahoo.com.br, jul.angel@hotmail.com, paschoarelli@unesp.com.br, srodrigu@fc.unesp.br

[2] Graduate Program in Management and Informatics in Health, UNIFESP - Univ. Federal de São Paulo, São Paulo, SP, Brazil
gcgotardi@gmail.com

[3] LIVIA - Laboratory of Information, Vision, and Action, UNESP - Univ. Estadual Paulista, Bauru, SP, Brazil
paulafp@fc.unesp.br

[4] MOVI-LAB – Human Movement Research Laboratory, UNESP - Univ. Estadual Paulista, Bauru, SP, Brazil
barbieri@fc.unesp.br

[5] LEI - Laboratory of Ergonomics and Interfaces, UNESP - Univ. Estadual Paulista, Bauru, SP, Brazil

Abstract. The aim of this study was to evaluate the effects of using a cell phone, operated in handheld (HAN) or speaker mode (SPK), on driver's gaze movements during a driving simulation. Ten novice and ten experienced drivers had their line-of-gaze kinematics measured while driving. In the cell phone use conditions, participants answered easy, everyday life questions. Results revealed that, as compared to a control (CON) condition, the number of visual fixations and vertical gaze variance were significantly increased and relative visual fixation time was significantly reduced during SPK and HAN conditions. Additionally, horizontal gaze variance was significantly increased under the SPK condition, as compared to CON condition. Novice and experienced drivers presented similar gaze behaviors. In short, talking on the cell phone disturbed attention due a decrease in eye's stability. Increased number of fixations with shorter durations tended to hamper the ability to obtain relevant visual information for safe driving.

Keywords: Car driver · Cell phone · Driving simulator · Distractor · Ergonomics

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_29

1 Introduction

Research on driving while using a mobile phone is growing in many countries; however research on the dangers of using cell phones while driving has not been consistent [1]. In Brazil, scientific research initiatives regarding this subject are not common.

Currently, the Brazilian transportation legislation does not allow the use of cellphones while driving, however the use of cellphones in the hands-free mode is permitted. Research indicates that both modes of cell phone use can have a comparable risk of accident occurrence [2] and frequently causes significant damage to drivers and pedestrians [1, 2]. The risks associated with talking on a cell phone is of great interest in transportation research as it is the most common task performed by drivers while operating a vehicle [3].

It is important to investigate all possible factors that can contribute to a drop in a driver's performance while operating a vehicle in order to prevent possible accidents and to improve transportation safety. In this context, the objective of the following study was to evaluate the distracting effects of cell phone use by analyzing the gaze behavior of drivers as they perform the tasks of answering and talking on their cell phones (handheld and hands-free or speaker mode) and identifying how these tasks are affected by driving experience while driving in simulated conditions.

1.1 Cell Phone Use Distraction on Driving and Traffic Safety

Distracted driving is a behavior that puts at risk the safety of drivers, passengers and pedestrians. A report conducted by the *National Highway Traffic Safety Administration* (NHTSA) indicated that, in 2014, there were a total of 29,989 fatal accidents in the United States caused by driver error, 2,955 of these fatalities were caused by distracted drivers (10% of all fatal accidents). Within these distractions, 13% were attributed to cell phone use [4].

An increase in cognitive load as a result of using a cell phone and operating a vehicle can interfere with the cognitive process of identifying and interpreting information relevant to driving conditions and surroundings, increasing the risk of accidents [5]. The principal use of cell phone for North American drivers is for conversation [3]. This activity requires visual and manual operations in order to place and receive phone calls or to end a conversation. A driver diverts a greater amount of processing resources in order to complete the tasks related to cell phone use, instead of concentrating entirely on the operation of a vehicle [1], yet studies show that drivers who talk on a cell phone while driving are four times more likely to be involved in traffic accidents [6].

In Brazil, in 2007, the city of São Paulo registered 237,000 citations to drivers talking on cell phones. It was the fourth highest ranked traffic violation for that year [7]. In Rio de Janeiro, the number of citations given to drivers for talking on their cell phones while driving reached nine million in 2008 [8].

Many people believe that talking on their cell phone while in speaker mode is less dangerous than holding the device in their hand to talk; however, research shows that the ability of a driver is highly affected in both cases [9]. Drivers can use up to 9% of

driving time handling and securing the cell phone, both in handheld and in speaker mode [10].

According the evidence stated above, it is clear that the cell phone is a distracting device for drivers, independent of how or what it is used for. Determining gaze behavior patterns under high cognitive demands allows a deeper understanding of the requirements of the cognitive system and the amount of attention involved. The present study examines the performance of dual tasks (driving and talking on a cell phone), proving quantitative data for clarifying the issues under debate.

1.2 Gaze Behavior

In the context of traffic safety, studies have shown that of the accidents caused by human error, more than 90% of them are due problems in the acquisition of visual information [11, 12]. Patterns of eye movements and acquisition of visual information that accompany static activities such as reading have been studied since the early 20th century, but it is only from the creation of the head-mounted eye trackers in the 1980s that it has been possible to study dynamic tasks such as walking, driving and practicing sports, among other activities [13].

Eye movements and their relation to visual attention have been widely studied in the area of sports and ergonomics, with an increase in research in automotive areas. The most studied gaze behaviors are visual fixation, smooth pursuit, saccadic and vestibulo-ocular reflex. Visual fixation allows for the processing of detailed information retrieved from its environment; it is a behavior in which the eyes remain relatively stationary on some aspect of the environment. The generated image is formed in the foveal region of the eyes. Smooth pursuit behavior is a combination of slow, continuous eye movements in which the ocular velocity is analogous to the velocity of the object it is tracking [13]. Saccadic is a rapid ocular movement that brings information from a new part of a scene onto the fovea. The vestibulo-ocular reflex is composed of compensatory and conjugate eye movements induced by head movements. These movements stabilize the eyes, maintaining the object of interest in the foveal region during head movements [14].

2 Method

Twenty automobile drivers between 18 and 40 years of age, male and female, participated in the study. The participants were equally divided into two groups of novice and experienced drivers. The group of novice drivers were characterized as individuals with less than 15,000 km of total driving experience, while those in the experienced group were participants of more than 30,000 km of total driving experience. It has been indicated that after a distance of 30,000 km of driving practice, drivers learn to use peripheral vision more effectively [15]. Kinematics of the eye measurements were taken of each participant as they drove in a driving simulator, at speeds between 80–120 km/h, in daylight, on a linear trajectory, in low level traffic.

A Head-Mounted Eye Tracker is a system used to analyze the movements of the eyes, with an eye-head data connection option, made by Applied Sciences Laboratories

(ASL, model H6) (Fig. 1). Its operation consists of detecting the position of the pupil and the reflection on the cornea through a tiny, almost infra-red beam of light that reflects on to the surface of the cornea and represents the position of the head in the image. This system consists of two micro cameras: the first camera records the subject's eyes (the image is reflected on the inside of a display coupled to a fixed support attached head of the user) (Fig. 2a) and the second camera directly records the scene that the user is viewing (Fig. 2b). In the first camera, the relative position of the pupil and a reflection in the cornea to a light source that is applied coaxially to the axis of the camera lens are identified in real time by an ASL software, based on their different levels of contrast.

Fig. 1. Head-mounted eye tracker. Source: [16]

Fig. 2. a. Video image of the eye registered by the Head-Mounted Eyetracker; b. Image of the X and Y axis (points of view on the screen); c. *Cockpit* (Logitech). Source: [16]

The centroids of the corneal reflection and of the pupil are automatically calculated and their coordinates, horizontal and vertical (in relation to the video) are utilized in order to define the line of sight in relation to the optical system of the apparatus. Based on the position of line of sight and the information derived from the calibration procedure, the system generates an indicator that is added to each of the frames of image from the camera scene (Fig. 2b), in order to indicate the position of the subjects gaze at the same time [13]. All of the data is recorded and translated into numerical format. The sampling frequency was 60 Hz.

For this experiment, a cockpit designed by Logitech was used, as illustrated in Fig. 2c. The driving simulator used was designed by City Car Driving (version 1.0). The scenario used in this study was a four lane highway with two lanes for each side of direction, the amount of traffic at 10% (the minimum amount of traffic on the highway) with a sunny morning setting. The systems presented were synchronized to allow the integrated analysis of the data. In addition to software from each system, software from Matlab (The Matworks Inc., 2010 - version 7.10.0.499) and SPSS software (SPSS Inc., 2008 - version 17.0.1) were used for statistical analysis.

The first step was to apply a test in order to measure the visual acuity, of which all should have normal or corrected vision. Afterwards, each participant was allowed a period of time to become familiarized with the cell phone (1 min), the cockpit (1 min) and the driving simulator (8 min).

Once familiarized with the equipment, calibration of the system to the eye movements was initiated. After the calibration was complete, the experiment was initiated. The experiment consisted of three different research driving simulation conditions: (CON) control, only driving; (HAN) driving and talking on a handheld cell phone, and (SPK) driving and talking on the cell phone in the hands-free speaker mode. The HAN and SPK conditions were randomized, while CON condition was always the first task performed. Each different driving simulation condition had duration of 4 min each. All of the participants were instructed to drive at a velocity between 80 e 120 km/h for each task, in order to maintain each simulation as close to reality as possible. In the CON condition, the participants only drove, with their eye behavior and head movements recorded. Using an external camera, the behavior of each participant was recorded during the simulation. In the second simulation test, HAN or SPK condition tasks were initiated. In the HAN condition, the participants were asked to drive naturally and when the cell phone rang (attached to the left-hand side of the steering wheel), they were to pick up the cell phone in their right hand, without removing the left hand from the steering wheel. The driver was then asked a series of random questions such as: What is your address? What do you plan to do on your vacation? The SPK condition was similar to the HAN, other than the driver would talk in the hands-free speaker mode, leaving the cell phone attached to the support next to the steering wheel. Four conversation scripts were elaborated (as to not repeat the same one with each participant).

To analyze and process the data (at intervals of 60 s per task), routines were created in Matlab in order to calculate the following dependent variables: number of fixations (units), relative time of fixations (%) (the percentage in relation to the total time of 60 s in which the participant remained fixating), horizontal variance of gaze position (cm) and vertical variance of gaze position (Table 1). The data for each variable were submitted to a group (experienced and novice drivers) by condition (CON, HAN and SPK) analysis of variance, with repeated measures in the last factor. There was only one trial for each condition. The significance level was set at 0.05 for all analyzes. Pairwise comparisons (*post-hoc*) were performed using Turkey's LSD test. Adjustments of probability by Bonferroni and corrections of degrees of freedom of Greenhouse-Geisser were applied as necessary [17].

3 Results and Discussion

The objective of this study was to evaluate the distracting effects of cell phone use on drivers' visual information acquisition during simulated driving. More specifically, this study aimed to identify both the effect of the tasks of answering and talking on a cell phone (in hands-free and speaker mode) and the effect of driver's experience on driving.

Table 1. Means and standard errors of total number of fixations (units), relative fixation time (%), horizontal gaze variance (cm), and vertical gaze variance (cm) of experienced and novice groups, with significance values for effects of group, condition, and group by condition interaction. *ns* = non-significant ($p > 0.05$)

Dependent variable	Group	Condition	M	SE	Group effect (*p*)	Condition effect (*p*)	Group x condition interaction (p)
Number of fixations (units)	Experienced	CON	104.700	10.180	*ns*	<.001	*ns*
		HAN	131.300	9.334		F(2.0,35.2) = 14.16	
		SPK	132.400	8.647		$\eta^2 = .44$	
	Novice	CON	104.000	10.180		HAN > CON (.001)	
		HAN	120.500	9.334		SPK > CON (.001)	
		SPK	125.900	8.647			
Relative fixation time (%)	Experienced	CON	84.456	3.575	*ns*	.002	*ns*
		HAN	82.491	2.910		F(2, 36) = 7.34	
		SPK	82.072	4.447		$\eta^2 = .29$	
	Novice	CON	87.258	3.575		HAN < CON (.021)	
		HAN	76.743	2.910		SPK < CON (.010)	
		SPK	76.680	4.447			
Horizontal gaze variance (cm)	Experienced	CON	82.214	14.101	*ns*	.003	*ns*
		HAN	76.460	23.742		F(1.6, 29.6) = 8.09	
		SPK	196.685	38.267		$\eta^2 = .76$	
	Novice	CON	56.670	14.101		SPK > CON (.005)	
		HAN	116.149	23.742			
		SPK	148.297	38.267			
Vertical gaze variance (cm)	Experienced	CON	60.860	14.893	*ns*	.005	*ns*
		HAN	114.301	23.996		F(1.5, 27.2) = 7.57	
		SPK	165.635	33.398		$\eta^2 = .0.30$	
	Novice	CON	56.837	14.893		HAN > CON (.009)	
		HAN	124.907	23.996		SPK > CON (.003)	
		SPK	135.128	33.398			

Studies with driving simulators in Brazil are relevant due to the mandatory use of these devices in automobile driving schools since the beginning of 2016. Resolution No. 543 of DENATRAN, dated July 15, 2015, established a deadline for the implementation of the simulators until December 31, 2015. Currently, individuals qualifying in category B (driving for light vehicles) must carry out, at least, 25 h of class training, of which five hours are in driving simulator [18].

The proposed method for the present study guarantees the participants, despite of differences in levels of experience between groups, faced similar requirements in the performance of task conditions. Criteria for assessing motor vehicle driving experience [15], based on an estimated life time distance traveled, allowed for a clear differential between the groups. The groups were gender balanced, and all participants were young adults with normal vision. Participants were able to perform tasks successfully; any trial resulting in failure (collisions) were removed from the analysis. Thus, it was possible to determine attention and cognitive demands, through gaze behavior variables, during task performance between the conditions (CON, HAN and SPK) and groups (experienced and novice), as shown in Table 1.

The present results indicated that the use of the cell phone while driving, as compared to the control condition, is associated with a greater number of fixations and with a reduction in fixation relative time (Table 1). Thus, the acquisition of visual information patterns characterized by fixation variables in the control condition is compatible with participant's lower attention demand during task performance. Conversely, the increased pattern of spatial distribution of line-of-gaze (horizontal and vertical variance) and the increase in number of fixations and its respective decrease in relative fixation duration can demonstrate the constraints imposed by the task of talking on a cell phone while simultaneously driving a vehicle with or without manipulating the cell phone (HAN and SPK conditions), resulting in a greater cognitive workload in both tasks. Overall, the results showed that using mobile phones while driving in a car driving simulator significantly alters participants' spatiotemporal gaze profile, while participants' experience does not affect this profile.

Quantifying the location of driver fixations when interacting with a cell phone is essential for understanding how cell phone use can affect the performance and driver's attention during simulated driving. In this study, the comparison between the ideal situation of driving a vehicle without interference from the use of a cell phone (control), and situations with the use of a cell phone in the participant's hand or in the hands-free speaker mode, showed changes in the acquisition of information and visual attention on a highway, manifested through the dynamics of eyes (Table 1). Using the cell phone, regardless of the type of use, significantly increased the number of fixations, relative to the control condition. The relative fixation time (which expresses the sum of the times of each fixation) was significantly reduced from the control condition to the on-hand or hands-free speaker conditions. From the point of view of the spatial distribution of gaze during the task, the horizontal (SPK condition) and vertical (HAN and SPK condition) variances of the gaze on the highway increased significantly during the use of the cell phone.

An analysis by the National Highway Traffic Safety Administration (NHTSA) in 2011 in the United States corroborates the findings of this study. The report shows

research into the effects of distraction on driving, along with the use of three types of mobile phones: mobile phones, mobile phones in the hands-free mode, and cell phones integrated into the on-board computer of the car. The experimental conditions were the control condition, in which the participants only drove, and the conditions in which they answered the cell phone, read and wrote message and talking using the device. The reported study had a naturalistic nature, in a real driving situation. Two hundred and four drivers were recruited; the results showed that participants used the cell phone 10.6% of the time they were driving. Through security critical risk analysis and variance analysis, it was shown that performing visual and manual tasks, such as answering the phone, entering numbers or letters and reading messages on a handheld cell phone (holding in hand) degrades performance and reduces the time of gaze fixation on the highway, increasing the risk of accidents [19].

In the present study, the relative fixation time in the CON condition indicated that participants fixated on the highway (on simulators screen) an average of 86% of the total time. During HAN and SPK conditions, this percentage dropped to 80 and 79%, respectively; that is, drivers were interested in visual aspects located elsewhere 20 to 21% of the total driving time. This reduction of the relative fixation time in the two conditions when participants were using cellular phone was statistically significant, with no differences between the HAN and SPK conditions. This data is an important indication of the inattention in the HAN and SPK conditions, arising from the use of the cell phone. Since fixation is the gaze behavior in which the best information processing and attentional engagement occurs [13], a reduction of this period gives evidence that the probability of worse behavioral responses has increased. This is particularly true if there is an increase in the requirements of the driving task, such as an increase in speed, when the need for quick responses due to unexpected stimuli, night and rain contexts, among others.

The data presented in this study indicate that both task conditions (HAN e SPK) distracted the gaze away from the roadway for much longer periods of time than that of the control condition. In both groups, there was increased risk to the driver. A 2001 study in Japan presented similar results regarding to the deleterious effect of attention due to performing dual tasks (driving and talking on a cell phone held in hand). The Japanese study was conducted using 50 participants and it investigated the distracting effects of using a cell phone (in hand and hands-free mode) and listening to music while driving. The amount of time that a participant's gaze was directed towards the cell phone, the speaker and the radio was measured, while driving at the same time. The average fixation time towards the cell phone held by the hand (hand-held) while driving, was significantly higher compared to the amount of time participants looked at the phone in speakerphone mode. The fixations towards the cell phone in speaker mode were significantly higher than the fixations on the radio. These results indicated that the drivers, when holding the cell phone in hand, fixate less on the highway and more on the device than when using the cell phone in the hands-free mode or listening to the radio [20]. The present study partially corroborates the findings of Ishida and Matsuura [20], since it only shows the effect of using a cell phone on the duration of the fixations, but does not find differences between its use in the in-hand and hands-free mode.

McCarley et al. [21] aimed to evaluate the effects of distraction on eye movements and visual coding while detecting changes in the road environment, and simulating a natural conversation on a cell phone in the hands-free mode. For the control condition, participants would watch a video with scenes of traffic in an urban environment and had to detect the appearance of a geometric figure in the scene pushing a button. In the dual task situation, participants watched a video and simultaneously talked with another person (who was in another room) by means of a microphone, simulating the conversation of a cell phone in the hands-free mode. The conversation was casual with themes about television entertainment. The results indicated that the cognitive workload imposed by casual conversation may have made it difficult to detect changes in the traffic scene. There was an increase in the number of fixations of visual gaze on the highway and decrease of the durations of fixations in the dual task in relation to the control task. These findings are similar to the increase in the number of fixations and the reduction of the relative fixation time observed in the present study.

McCarley et al. [21] observed visual perception failure during the dual task, making it difficult to visualize the geometric figures in the scene. The short durations of fixation during the conversation made very little time available for adequate coding of the visual information by the fovea, not allowing higher levels of processing. The data from this experiment indicated that the foveal analysis (image processing) in visual search becomes more prone to error when the duration of fixation is decreased [22]. The study by McCarley et al. [21] corroborates the findings of the present study, since there were more fixations for the HAN and SPK condition in relation to CON, with relative durations of fixations with respect to the total time available when compared to CON. Both studies suggest that dual tasks, such as driving and talking on a cell phone (in-hand or hands-free mode) concomitantly, present a greater demand of attention, and increasing the potential of vehicle failure, such a missing a traffic sign or causing an accident.

In summary, present findings show a relationship between the changes in information processing - due to an increase in attentional demand due to the use of cellphones in conjunction with the driving an automobile and the respective changes in the behavior of the eyes. These indicatives are in line with the results of the present study.

4 Conclusions

This article presented an experimental protocol to investigate dual tasks involving driving an automobile in a driving simulator while participants interacted with a cellular phone. The respective perceptual and cognitive demands were characterized by data obtained from an eye tracking system. In general, the results showed that, due to the stronger cognitive and attention demands, the number of fixations on the highway and gaze variance (horizontal and vertical) increased and the relative fixation time decreased, as compared to the control condition (no conversation). Effects of experience on driving were not observed, resulting in similar gaze behavior of both groups. In short, talking on the cell phone disturbed attention due a decrease in eye's functional stability. Increased number of fixations with shorter durations tended to hamper the ability to obtain relevant visual information for the safe driving. Areas of study such as

Ergonomics can effectively assist in the planning of equipment and devices that meet cognitive, perceptual and motor demands of drivers, thus contributing to traffic safety.

Acknowledgments. The authors thank the National Council for Scientific and Technological Development – CNPq (Sérgio Tosi Rodrigues, grant # 458916/2014-5) and São Paulo Research Foundation - FAPESP (Gisele Chiozi Gotardi, grant # 2015/10851-3).

References

1. Caird, J.K., Willness, C.R., Steel, P., Scialfa, C.: A meta-analysis of the effects of cell phones on driver performance. Accid. Anal. Prev. **40**(4), 1282–1293 (2008). doi:10.1016/j.aap.2008.01.009
2. McEvoy, S.P., Stevenson, M.R., McCartt, A.T., Woodward, M., Haworth, C., Palamara, P., Cercarelli, R.: Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study. BMJ **331**, 428–433 (2005)
3. NHTSA, National Highway Traffic Safety Administration. State Laws (2013). http://www.distraction.gov/content/get-the-facts/state-laws.html
4. NHTSA, Distracted Driving 2014. Summary of Statistical Finding (2016). DOT HS 812 260. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812260
5. Alm, H., Nilsson, L.: Changes in driver behaviour as a function of hands free mobile phone. Accid. Anal. Prev. **26**, 441–451 (1994)
6. Caird, J.K., Scialfa, C., Ho, G., Smiley, A.: The effects of cellular telephones on driving behaviour and accident risk: results of a meta-analysis. Canadian Automobile Association (CAA), Foundation for Traffic Safety, Edmonton, Alberta (2004)
7. SÃO PAULO (Estado). Secretaria de Planejamento e Gestão. Departamento Estadual de Trânsito (2007). http://www.detran.sp.gov.br/
8. RIO DE JANEIRO (Estado). Secretaria de Estado da Casa Civil. Departamento Estadual de Trânsito (2008). http://www.detran.rj.gov.br/
9. Strayer, D.L.: Conversation disrupts visual scanning of traffic scenes. In: 9th Vision in Vehicles Conference, Brisbane, Australia (2001)
10. Pickrell, T.M., Ye, T.J.: Driver electronic device use in 2010. In: Traffic Safety Facts Research Notes. Report No. DOT HS 811 517. National Highway Traffic Safety Administration: Washington, DC (2011)
11. Sivak, M.: The information that drivers use: is it indeed 90% visual? Perception **25**(9), 1081–1090 (1996)
12. Land, M.F.: Eye movements and the control of actions in everyday life. Prog. Retin. Eye Res. **25**, 296–324 (2006)
13. Rodrigues, S.T.: O movimento dos olhos e a relação percepção-ação. In: Teixeira, L.A. (Org.) Avanços em comportamento motor. Movimento: Rio Claro, SP, 122–146 (2001)
14. Summala, H., Nieminen, T., Punto, M.: Maintaining lane position with peripheral vision during in-vehicle tasks. Hum. Factors **38**, 442–451 (1996)
15. Lehtonen, E., Lapp, O., Koirikivi, L., Summala, H.: Effect of driving experience on anticipatory look-ahead fixations in real curve driving. Accid. Anal. Prev. 49 (2014)
16. Carizio, B.G.: Efeito distrator do telefone celular sobre a condução simulada de automóveis: situações de uso manual e viva-voz.. Dissertação (Mestrado em Design) – Faculdade de Arquitetura, Artes e Comunicação, Universidade Estadual Paulista Júlio de Mesquita Filho, Bauru, 109 (2016)

17. Maxwell, S.E., Delaney, H.D.: Designing Experiments and Analyzing Data: A Model Comparison Approach. Wadsworth, Belmont, CA (1990)
18. Brasil. Ministério das Cidades. Departamento Nacional de Trânsito. Resolução nº 543, de 15 de julho de 2015. Altera a Resolução CONTRAN nº168, de 14 de dezembro de 2004, com a redação dada pela Resolução CONTRAN nº 493, de 05 de junho de 2014. Brasília, DF (2015)
19. NHTSA. National Highway Traffic Safety Administration. State Laws (2011). http://www.distraction.gov/content/get-the-facts/state-laws.html
20. Ishida, T., Matsuura, T.: The effect of cellular phone use on driving performance. IATSS Res **25**(2), 6–14 (2001)
21. Mccarley, J.S., et al.: Conversation disrupts change detection in complex traffic scenes. Hum. Factors **46**(3), 423–436 (2004)
22. Hooge, I.T.C., Erkelens, C.J.: Control of fixation duration during a simple search task. Percept. Psychophys. **58**, 969–976 (1996)

Risky Behavior and the Attitudes Towards the Police Safety Operations in the Czech Republic

Petr Zámečník(✉), Veronika Kurečková, Pavel Řezáč, and Vít Gabrhel

Transport Research Centre (CDV), Líšeňská 33a, Brno, Czech Republic
{Petr.Zamecnik,Veronika.Kureckova,Pavel.Rezac,
Vit.Gabrhel}@cdv.cz

Abstract. This study maps attitudes of drivers towards police actions and operations in the Czech Republic. Data was gathered through a representative survey, which involved 1482 persons aged 18 years and older. The survey covered three types of topics. (1) Orientation in the system of sanctions and perceiving of punishments for traffic offences. (2) Experiences with traffic police. (3) An efficiency of the police operations and actions. The result shows slight and time limited influence of police actions on drivers' behavior. Data also show insufficient deterrence and ignorance of system among drivers.

Keywords: Traffic safety · Risk behavior · Police action efficiency · Attitudes towards police

1 Introduction

In the Czech Republic, we have frequent and long-term risky behavior on the roads in comparison with other EU countries. This behavior is connected with high road accident rate and mortality rate. This is usually mentioned together with two reasons. (1) Insufficient system of graduate drive licensing and (2) minimum number of measures aimed at a small group of the most risky drivers. Less often we can find research which deal with insufficient deterrence as possible source of risk behavior.

It is well known that lack of feedback from system, along with a perceived improvement in driving skills during the first years of driving causes change in the perception of the consequences of risky behavior [1], overestimation of own driving skills [2], and a negative impact on risk perception [3]. The significance of these factors is modified by age, experience and education that have an impact on personality traits and attitudes to violations of traffic rules [4]. This results in higher participation of young drivers in traffic accidents, since they tend to behave more risky when driving, especially in the context of speeding [5], and more frequent abuse among men [6, 7].

Without appropriate deterrence, drivers tend to leave safe behavior learned in driving schools as the motivation for keeping rules falls with time. Part of the drivers would not follow the rules without threat of detection and penalties or would show much less respect towards the rules. That is why deterrence is important part of traffic safety and should be enhanced as much as possible. By deterrence is meant police surveillance, detection and prosecution of offenders.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_30

Main aim of the deterrence is motivation of road users to comply with the rules while it should also fulfill a preventive function. It is one of the essential components of traffic safety, since it has an immediate effect on the number of accidents and their consequences (depending on its intensity and quality). It is primarily focused on increasing the safety of traffic participants through the threat of detection and punishment (general level), and through penalties for those who were caught and convicted (specific level). Its effectiveness depends on the perception of sanctions in the driver population. Thus, the attitudes towards the police safety operations influence driver behavior considerably and have big potential impact on risky or safe behavior.

Whereby the most important seems to be certainty of punishment [8], while increasing the actual hardness of the punishment brings only a small reduction in the risk behavior of individuals [9], and little change in the behavior of the population [10, 11]. Severity of punishment is an effective factor in the case when the perceived certainty of being caught and punished is high [11, 12]. But for example in group of repeated offenders is perception of both very low (hence have harsh penalties for this group of little influence). In the case of the swiftness of punishment, there is no accordance among researchers but it can be regarded as a less important factor than the certainty of detection [13, 14]. Swiftness of punishment was considered as an important factor in the past, since the times of Criminology book "On Crimes and Punishments" by Cesare Beccaria from the 18th century to the late 20th century. In the context of contemporary research, however, it shows that this is characteristic rather for animals, not for humans, who are able to combine punishment with the offence even after a long time.

Overall, it can be stated, that if certainty of punishment is high, punishment comes quickly and is reasonably hard then there is a measurable reduction in risk behavior. Reduction is on both, specific (in particular man) and general (the whole population) levels. Discourse of harsh sanctions, which is regularly emerging within the social or political debate, has no support in traffic psychological or criminological literature. In addition, for hard sanctions such as imprisonment, we can follow a series of accompanying negative phenomena. For example, it has negative effect on future opportunities in terms of education and employability; on immediate vicinity of the offender; relationships and family background and reputation (stigma of conviction). Also, administrative burden and costs of these harsh penalties for traffic offences disproportionately burden the judicial system [15].

2 Aim and Methods of the Study

As the deterrence is key factor for reducing risky behavior and its efficiency depends on perceiving in general population, aim of this research was to map attitudes towards the police safety operations in the Czech Republic, which are a main tool of deterrence.

A survey was used as a method of data gathering. The aim of the survey was to conduct a statistical investigation mapping an attitudes and experience of the target population with the traffic police operations and actions. The survey was conducted by statistical agency, which interviewed respondents aged 18 years and older, in such a way as to ensure the representativeness of the sample in relation to the population of the

Czech population - in terms of socio-demographic profile of respondents. Basic selection of the sample was based on quota sampling, reflecting relevant socio-demographic characteristics of respondents older than 18 years. This means age distribution, proportional representation according to the region of residence, size of inhabited municipalities, education, and gender. The research involved a total of 1482 persons, who in the past six months did not participate in market research and public opinion research.

In the survey, apart the sociodemographic, drivers were asked on three clusters of questions. (1) How much are they oriented in the system of sanctions and perceiving of punishments for traffic offences; (2) What experiences do they have with traffic police; 3) How they evaluate efficiency of the police operations and actions.

3 Sample

While women accounted for 51.5% of the total respondents (N = 763) in the case of men, it was 48.5% (N = 719). From this perspective, the sample fairly accurately reflects the gender structure of the population - according to the current data of the Czech Statistical Office (CSO), women account for approximately 51% of the population [16].

Regarding the highest level of education, 16.5% (N = 244) of respondents reached basic (including unfinished) level of education, 37.1% (N = 550) completed secondary education, 34% (N = 504) high school education and finally, 12.4% (N = 184) of respondents achieved a university degree. If we compare these data with the Czech population [16], once again we can say relatively accurate reflection - at least within the categories of basic (population 17.6%) and higher (12.5% of the population) level of education. The share of respondents with completed secondary level of education is slightly different from the population - in the case of secondary education there were 33% and 27.1% of the population.

Regarding the age of the respondents, the average age in the sample corresponded to the age of 46.6 years (SD = 16.6). This is a value higher than the national average, i.e. 41.9 years (CSO, 2016), which is in line with the intention of research as it was oriented on adult population. The oldest person was 89 years, the youngest respondent 18 years.

Regarding the frequency of driving during the last year, almost a third (27.7%; N = 411) of the respondents drive every day and only slightly fewer respondents drive at least several times a month (26.5%; 392 persons). About a third of respondents in the last year or did not drive the car. There were 71 (4.8%) professional drivers.

4 Results

In terms of cluster 1 the vast majority of respondents admitted very small (if any) orientation in the current system of penalties for traffic offenses (including the number of demerit points and the amount of the fine). Only about 6% drivers has a good knowledge of the system of penalties (according to self-assessment), while 19% openly declared knowing nothing of the system and 42% of drivers declared unfamiliarity with the system.

This ignorance is reinforced by discrepancy between agreeing with the relatively severe penalties for offences and own behavior, which is characterized by committing those offences. Really interesting is, that there was no difference in assessing the severity (of punishment) between group who commit those specific offences and group who do not. And also that majority of respondents would welcome more police controls and actions.

As for experience with traffic police, 54% of drivers were stopped by police patrols during last 12 months. When considering overall experience with patrols and officers, only 9% of drivers had negative feelings from the encounter(s). Neither age nor gender were significant in this cluster. Answers were consistent and there were no differences in terms of any measured parameter of the sample.

From answers from cluster 3, it is clear, that police actions and operations have at least some impact on behavior. As much as 53% of drivers declared that they more follow the rules when there are any police actions in their region. Despite this change in behavior, only 3% of drivers use mobile applications, which are able to monitor police patrols and actions in order to avoid possible encounter.

The most important part of the third cluster was the evaluation of police actions and operations. When considering impact of police actions and operations on other drivers, respondents generally predicted influence on 50% of the driving population. Specific influence was rated in a way that for all surveyed categories was count share of sample that expects impact on at least half of driver population. As the most efficient were seen actions towards drink driving when 70% of respondents think, that these actions will positively influence at least 50% of drivers in short term (up to one week after noticing the action). On the opposite side there were actions focused on drug driving (43% of respondents assume positive impact on at least 50% of drivers) and prevention (40% of respondents assume positive impact on 50% of drivers).

When the sociodemographic variables are taken into account it can be seen that only education makes some differences. There were two significant positive relationships. Firstly, there is relationship between education and using hands-free instead of holding cell phone in hands (Kendall Tau-c (1079) = 0.15; $p < 0.001$). And secondly, between education and speeding up to 10 km per hour (Kendall Tau-c (1079) = 0.08; $p < 0.05$).

5 Discussion

As results shows, it is clear that police actions and operations have some impact on individuals, and up to the extent to which we can rely on assessing the others, also on driving population. Results also show, that people are not able to assess impact or influence precisely. For example, they are not able to distinguish among specific police actions and tend to evaluate them in same way. Based on little differences in scores for each type of action it could be said, that when evaluating actions, they evaluate deterrence in general. However, although the efficiency cannot be rated for each action, it is possible to track the perceived efficiency generally what brings light on deterrence effect in the Czech Republic. When taking this general efficiency in concrete numbers, the effect does not seem to be high. Half of the drivers admit behavior change for up to one

week after noticing action and suppose in average influence of action on half of the population. But when we take into account acceptance of severe punishments even for those offence which they commit and very poor knowledge of sanction system it can be stated that drivers do not fear of being caught or being punished.

This can be supported by fact that drivers perceive the system of punishments as a fair as there is no difference in evaluation between those who commit those offences and those who do not. Also direct work of the police is perceived as polite when more than 90% drivers did not notice any problems. With this numbers, it is not surprising that majority of drivers demands more police patrols on the roads. This leads again to assumption that drivers do not perceive the threat of detection and punishment as high or real.

In addition, there is one more indication for above-mentioned statement. Number of controlled drivers in the sample was 54%. This is extremely high within condition in the Czech Republic. According to police statistic [17], in 2015 there were 1664 police actions and operations in whole country and police checked altogether nearly 600 thousands drivers that year. In the Czech Republic, there is 10 million of inhabitants, 6.5 million have driving license and 5 million are active drivers. That leads to statistical probability of control approximately once per ten years. This means two things. Firstly, our sample does not match these statistical probabilities, and secondly even when controlled so often, they still do not fear of detection and punishment.

6 Conclusion and Recommendation

Based on our research we can evaluate the strength of deterrence in the Czech Republic as rather low. Even when drivers from the sample encounter police controls frequently they do not feel need to know the system and punishment for breaking traffic rules, and what is more, wish more police patrols on the roads. This assumption should be nevertheless confirmed by further research focused directly on perceiving fear or probability of punishment from the authorities.

The lack of knowledge of sanction system is probably caused by the fact that the Czech Republic does not have implemented any measure from graduated driver licensing. There is no accompanied driving, no prolonged test period for novice drivers, no driver improvement courses and all learner drivers obtain full license after few weeks in driving school. What is more, in driving schools there is possibility to have "individual plan" what means that learner drivers do only test rides and after 28 "school" hours (45 min) of test driving can undergo final exam. This is one of the shortest period of learning in EU countries and thus absolutely insufficient for safety behavior of novice drivers. Therefore, it is crucial to extend length of learning and cancel the possibility of individual plan as learner drivers in this type of learning only memorize the test questions and correct answers without insight into their meaning and importance. It is also necessary to include discussion about risk behavior and sanctions, which follow it, as awareness of consequences is important factor for preventing traffic offences. All these recommendations are periodically offered to Czech ministry of transport and highlighted in

various strategic documents (e.g. in National road safety strategy) but there is no political will for implementation so far.

Driving schools methods and driver licensing system are not the only ones that should be changed. This research has possible implication also for work of the traffic police. Results of our study show that drivers see no or little differences among police actions and thus have no clear idea of what is police doing (vast majority of respondents were not able to name any preventive police action). This strongly influences subjective probability of being caught for committing offences (in negative way) and consequently causes more risky behavior on the roads. Thus it is necessary for police to communicate better with public, explain what is being done for traffic safety and that there is plenty of various actions that are targeting unsafe behavior. Higher awareness of police work will enhance deterrence effect and affect drivers' subjective feeling of likelihood of detection of inappropriate behavior.

Acknowledgements. This article was produced with the financial support of the Ministry of Education, Youth and Sports within the National Sustainability Programme I, project of Transport R&D Centre (LO1610), on the research infrastructure acquired from the Operation Programme Research and Development for Innovations (CZ.1.05/2.1.00/03.0064).

References

1. Delhomme, P., Verlhiac, J.F., Martha, C.: Are drivers' comparative risk judgments about speeding realistic? J. Saf. Res. **40**(5), 333–339 (2009)
2. Whissell, R., Bigelow, B.: The speeding attitude scale and the role of sensation seeking in profiling young drivers at risk. Risk Anal. **23**(4), 811–820 (2003)
3. Rosenbloom, T.: Risk evaluation and risky behavior of high and low sensation seekers. Soc. Behav. Pers. **31**(4), 375–386 (2003)
4. Najeeb, M.A.: A study of the psychological factors that influence the rule violation behavior of drivers. In: 25th ICTCT Conference Road Safety in a Globalised and More Sustainable World – Current Issues and Future Challenges, ICTCT Press, Hasselt (2012). http://www.ictct.org/workshops/25-hasselt-belgium/
5. Smart, D., Vassallo, S.: In the Driver's Seat: Understanding Young Adults' Driving Behaviour (Research Report No. 12). Australian Institute of Family Studies, Melbourne (2005)
6. Cestac, J., Paran, F., Delhomme, P.: Young drivers' sensation seeking, subjective norms, and perceived behavioral control and their roles in predicting speeding intention: how risk-taking motivation evolve with gender and driving experience. Saf. Sci. **49**(3), 424–432 (2011)
7. Taubman-Ben-Ari, O., Mikulincer, M.: The effects of dispositional attachment orientations and contextual priming of attachment security on reckless driving. Transp. Res. Part F Traffic Psychol. Behav. **10**(2), 123–138 (2007)
8. Vito, G.F., Maahs, J.R., Holmes, R.M.: Criminology: Theory Research & (Criminal Justice Illuminated): Theory. Research and Policy. Jones & Bartlett Learning, Burlington (2006)
9. Mann, R.E., Vingilis, E.R., Gavin, D., Adlaf, E., Anglin, L.: Sentence severity and the drinking driver: relationships with traffic safety outcome. Accid. Anal. Prev. **23**(6), 483–491 (1991)
10. Elvik, R., Christensen, P.: The deterrent effect of increasing fixed penalties for traffic offences: the norwegian experience. J Saf. Res. **38**(6), 689–695 (2007)

11. Watson, B., Siskind, V., Fleiter, J.J., Watson, A., Soole, D.: Assessing specific deterrence effects of increased speeding penalties using four measures of recidivism. Accid. Anal. Prev. **84**, 27–37 (2015)
12. Briscoe, S.: Raising the bar: can increased statutory penalties deter drink-drivers? Accid. Anal. Prev. **36**(5), 919–929 (2004)
13. Fildes, B.N., Lee, S.: The Speed Review: Road Environment, Behaviour, Speed Limits, Enforcement and Crashes. Monash University Accident Research Centre, Melbourne (1993)
14. Nagin, D.S., Pogarsky, G.: Integrating celerity, impulsivity, and extralegal sanction threats into a model of general deterrence: theory and evidence. Criminology. **39**(4), 865–891 (2001)
15. Dant, T.: The driver-car. Theory Cult Soc. **21**(4–5), 61–79 (2004)
16. CSO: Report of Czech statistical bureau on sociodemographic statistics (2016). https://www.czso.cz/csu/gender/1-gender_vzdelani
17. PCR: Report of Czech traffic police on traffic accidents and police controls (2016). http://www.policie.cz/clanek/statistika-nehodovosti-900835.aspx

Does the Familiarity of Road Regulation Contribute to Driving Violation? A Simulated Study on Familiar and Unfamiliar Road Intersections among Young Chinese Drivers

Siyuan Huang[1,2], Daniele Ruscio[2], Dedy Ariansyah[2], Jun Yi[1(✉)], and Monica Bordegoni[2]

[1] School of Design, Hunan University, Changsha, China
oliviahuangshuishui@gmail.com, yijun7809@163.com
[2] Department of Mechanical Engineering, Politecnico di Milano, Milan, Italy
{daniele.ruscio,monica.bordegoni}@polimi.it,
dedy.ariansyah1@gmail.com

Abstract. Road accidents represent the primary cause of death among young drivers. To understand the emerging issue of young drivers' involvement in traffic accidents abroad, risk-taking behaviors were investigated in familiar and unfamiliar driving situations. Twenty-two young Chinese drivers completed a road regulation test followed by a simulated test drive. The number of traffic violations and accidents in familiar and unfamiliar driving intersections were correlated to road regulation knowledge, risk perception scores and to self-assessment of driving skills. Significant number of mistakes was found in risk-taking situations, regardless of the familiarity of the situation, especially for drivers that presented high ratings of self-assessed driving skills. Results show that risk-taking behaviors while driving in unfamiliar conditions are mediated by psychological factors, like self-assessment of being a good driver, more than the actual knowledge of road regulation rules. Implications for international driving can be considered for future research development.

Keywords: Young drivers · Road violations and accidents · Road intersection · Risk perception · Driving skills · Knowledge of traffic laws · International driving

1 Introduction

Road traffic injuries represent one of the main causes of death among young drivers worldwide, especially for the 15–29 years old category [1]. Studies have estimated that young drivers are more exposed to road accidents and present at the same time a higher propensity to risk-taking behaviors while driving [2]. Research on human factors has identified significant factors that correlate young's risk-taking behaviours with personality trait, spatial and cognitive ability development, as well as the susceptibility to peer pressure and to social norms [3, 4]. Considering the growing number of foreign people that are migrating into western countries like moving abroad for international study

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_31

programs, young drivers' involvement in road accident abroad, is emerging as a potential novel social issue. For instance, foreigner drivers in Italy are reported to be involved in the 23.7% of the "driving in the wrong way" accidents in 2014. According to a statistical sample of 1289 drivers involved in severe road accidents, 12% of the victims were indeed foreigners [5]. Even if this problem is increasingly impacting on modern society, less attention is given to concrete training programs designed specifically for young foreign drivers [6]. At the same time, young people have easy access to motor vehicle worldwide like renting a car from car-rental companies or use car-sharing mobile systems in several European countries. Thus, it is important to conduct scientific research on the way individual and cultural factors interact with the specific condition of driving abroad, in order to provide a comprehensive insight into young drivers' vulnerability and support specific training for reducing traffic injuries.

1.1 Individual Factors in Risk-Taking Behaviors

Several factors are able to explain road accident involvement of young drivers.

Risk Perception. Numerous studies have correlated risk perception abilities with traffic accident involvement [7]. Risk perception is defined as a subjective judgment that categorize an upcoming event as potentially dangerous [8]. It is a crucial ability in order to anticipate, evaluate and judge uncertainty to avoid negative consequences [9]. Young people aged from 16 to 25 [2] seem to present a systematic tendency to underestimate the threat level of driving situations and seem less able to visually anticipate danger, compared to experienced drivers [10]. At the same time, present higher scores in sensation seeking and willingness to take risks [11], making risk perception as one of the main factors that could explain the traffic accident involvement of young drivers.

Self-assessed Driving Skill and Road Regulation Knowledge. Driving abilities are usually considered as two distinct elements: the theoretical knowledge of the road regulation rules and the practical driving skill required to operate the vehicle [12]. According to Spolander [13], the driving skill can be considered composed by two further aspects: (1) the motor-coordination required for vehicle operation and (2) the defensive ability of being able to anticipate the interaction with other road users. Defensive ability includes the cognitive functions related to visual perception and working memory required to read the situation, but also the subjective assessment of self-driving skills required to safely manage a given driving intersection [14]. Drivers usually tend to overestimate their driving skills [15–17] and as a consequence they could underestimate the potential risks to perform some risk-taking behaviors, like overspending or overtaking [18, 19]. The lack of driving experience of young drivers could also affect their self-assessment of driving abilities, leading to inaccurate evaluations of road situations and a high possibility of risk-taking maneuvers. This difficulty could be more evident in concern to the issue of young drivers being exposed to unfamiliar road situations, where culturally different environment could make more complex the evaluation of road situations.

1.2 The Effect of Road Familiarity on Risk-Taking Behavior

Risk perception and self-assessment of driving abilities can be influenced by social and cultural norms [20]. Studies comparing the risk perception related to specific traffic hazards have found differences in attitudes and driving behaviors in drivers coming from different countries [21]. Cultural attitudes toward safety and risk can generate specific consolidated habits that will eventually shape the formal and informal definitions of what a "safe driver" could represent [22, 23]. That is why driving in an unfamiliar environment can increase the risk of traffic violations within different cultures [24, 25]. The likelihood of serious conflict in intersection was highlighted for instance within informal driving rules that existed in Sweden [26], with different traffic rule generating conflicts among users that were strictly following the formal driving rules. Drivers who obtain different traffic knowledge from the country in which he or she is driving might be involved in violations and accidents if they are not well-informed of the local road regulation rules. This is especially distinct when considering different traffic regulation across countries where the design of road intersection and regulated signs may be the same or similar but turn out to be different in actual implementation. For instance, in Italy, a T-junction regulated by a traffic light prohibits the car to turn right when the traffic light is red, while it is inconsistent for the traffic regulation in China.

At the same time, familiar circumstances can also generate underestimation of risk perception. The observation of real-road driving in Israel showed that drivers tended to perform more severe road traffic violations, as well as more dangerous behaviors when they drove in more familiar locations [27]. The impact of road familiarity was also observed on visual attention to traffic signs. The effect of familiarity decreased gaze fixation frequency and duration as drivers drove the same road several times [28], impairing drivers' ability to detect emergency situations [29]. Furthermore, research also found that a higher propensity of risk taking behavior was mediated by drivers having stronger motives, like overconfidence and a desire "to get there quickly" [2] in habitual routes. These motivations beyond risk-taking behaviors were more likely to appear in a familiar road than unfamiliar road conditions, as people were more reluctant to adhere to the rules or norms when they drove in familiar settings [27].

1.3 The Present Study

To better understand the involvement of young drivers in traffic violations and road accidents abroad, a driving simulation was tested to assess the way road regulation rules, risk perception and self-assessment of driving skills can affect the practical decision-making process during critical risk-taking situation. An experimental task was designed to induce some pressure and time-constrains, in familiar and unfamiliar road intersections. A group of young Chinese drivers studying in Milan, Italy were considered for the experiment. Knowledge of Chinese and Italian road regulation rules were tested in a questionnaire and in the practical driving situations. Participants' scores of risk perception and driving skills were also recorded. The objectives were: (1) determine if the knowledge of road regulation of young drivers correlated directly with their driving

performance; and (2) examine if mediating factors like risk perception and self-assessment of driving skills were able to explain their driving performance under time-pressure condition. We hypothesize that: (H1) a relationship between road regulation knowledge and driving performance could be found; (H2) risk perception and self-driving assessment could present a mediating effect on the driving performance; (H3) Familiar intersections would result in correct theoretical responses, but at the practical driving behavior more violations could occur under time-constrain situation, compared to the unfamiliar situations.

2 Method

2.1 Familiarity of Road Intersections

Driving scenarios were tested using a fixed-based driving simulator with a three-screens display that provided a horizontal and a vertical field of view of 165° and 36° respectively. The driving scenes reproduced a European driving environment that included familiar and unfamiliar road situation for Chinese drivers. Familiar situations were a roundabout and a left turn at a T-junction, both regulated by traffic light (Fig. 1A, C), according to the same traffic rules that apply to China (i.e. give way when the traffic light is red). Unfamiliar situations were built according to the Italian code of law that states that it is mandatory to give way to the drivers already inside a roundabout (Fig. 1B) and it is prohibited to take a right turn at a T-junction with the red light on (Fig. 1D). The familiarities of the road situations were examined in a preliminary pilot test among

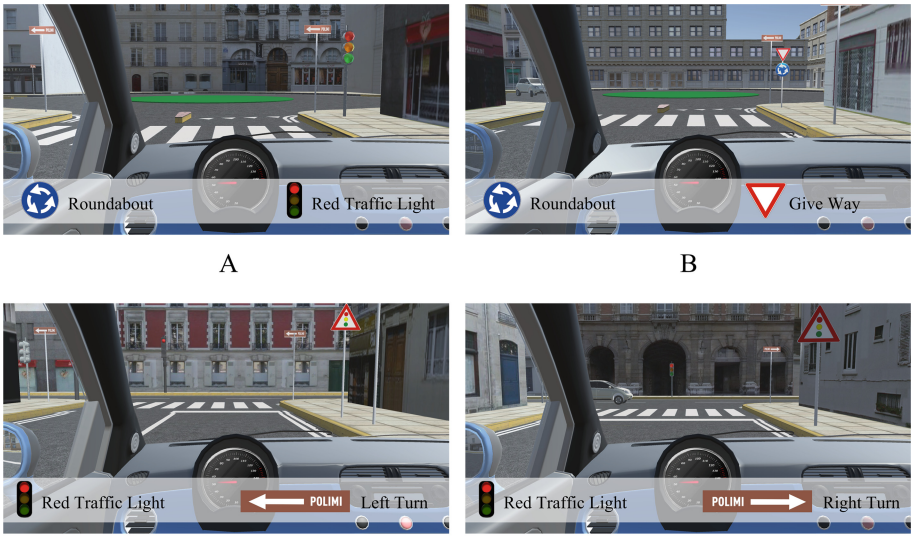


Fig. 1. Familiar road regulations for Chinese drivers (intersections A and C), and unfamiliar road regulations rules (intersections B and D) while driving abroad

one Italian and three Chinese young drivers that they confirmed the differences in road regulation between the two countries. In order to create a sense of urgency, participants were informed that they were late for class and they were required to reach the destination (10 km within 10 min). In the middle and toward the end of the test, auditory warning reminded the drivers that there were 4 km in 4 min left, and 2 km in 2 min left. Participants were also reminded to respect all the traffic laws and to reach the destination on time.

2.2 Measurement

Risk Perception Questionnaire. The revised form of Risk Perception Scale [31] validated for the Chinese sample [30] was used in the current test. Participants were required to evaluate the perceived risks of 23 driving situations, using a 10-point Likert Scale. The scale was composed of 17 items on general driving situations and 6 items on driving abroad situations. The total score for the scale ranged from a minimum of 23 points in the case of the driver perceived all the driving situations as extremely safe, to a maximum of 230 points in the case of the driver perceived all the driving situations as extremely dangerous.

Self-assessed Driving Skill Questionnaire. The Chinese adaptation [31] of the Driver Skill Inventory (DSI) [13] was used to measure self-assessment of driving skills. The scale was composed of two factors: (1) technical driving skills, and (2) defensive driving skills, measured using a 10-point Likert Scale, (1 = "I don't have this driving ability", 10 = "I completely master this driving ability"). The total score for the scale ranged from a minimum of 17 points in the case of the drivers' self-assessed driving skills are very poor, to a maximum of 170 points.

Road Regulation Knowledge. The specific knowledge of Chinese and Italian road regulation rules were measured by a questionnaire based on Chinese and Italian traffic regulation tests. The ad-hoc questionnaire resulted composed of: (1) questions on road signs meanings; and (2) a series graphical images of Chinese and Italian driving situations that required the participant to describe the correct driving behavior in those situations. The number of right and wrong responses was recorded, as well as the admission of not knowing the answer. 0-points were scored in case of right answers; 0.1-points were scored for admitting to not know the answer; 0.2-points were scored in case of wrong answers. The total score for the road regulation knowledge ranged from a minimum of 0 points for drivers who knew the correct behavior for all the situations, to a 0.4 score for drivers who admitted not knowing the correct response, to a maximum of 0.8 points in the case of wrong responses in all driving situations.

Driving Performance. The frequency of mistakes, including traffic violations and road accidents were scored from the video recording of the driving performance. The Manchester Driver Behavior Questionnaire (DBQ) factors were used as reference for the video scoring of the mistakes, in order to keep track of: violations, dangerous errors, and relatively harmless lapses [32]. In the case of correct driving performance and

relatively harmless lapses, 0 point were assigned to the driver; in the case of violations of the traffic law, 0.1 point were assigned; in the case of dangerous errors and accidents, 0.2 points were scored. The total score for traffic violations and accidents was calculated as "mistake score", which ranged from a minimum of 0 points in case of absence of mistakes, to a maximum of 0.4 when the driver violated the traffic law in all tested situations, or to a maximum of 0.8 points in case the driver violated the laws and also made dangerous errors that caused road accidents.

2.3 Sample

Twenty-two Chinese young drivers voluntarily participated in the driving test. Panel sampling strategy was used, in order to carefully select the Chinese students that reported, in a previous step of the research, to being in Milan since few months and with the intention to drive in Italy in the future. The sample consisted of 10 male and 12 female drivers, with a mean age of 24.73 (SD = 1.42) and an average driving mileage of 2,760 km per year (SD = 4,692).

2.4 Procedure

After the consent form was signed, a demographic questionnaire was filled along with the risk perception and the self-assessment of driving skills scales. After the questionnaire, an adaptation scenario was presented to gradually introduce the participant to the virtual simulation and avoid potential simulation sickness effects. In the experimental session, the driving task was introduced to participants, with the instruction to follow the road direction signs to reach the destination. Emphasis on the "late for class" motivation was provided, in order to investigate the decision-making process in risk-taking situation, still reminding that the respect of all the road regulation rules and safe driving would have been the first priority. The urban scenario included a random combination of familiar and unfamiliar road situations for the Chinese drivers abroad. After the virtual driving task, a final questionnaire tested the theoretical knowledge of the same road situations faced in the practical driving. A debriefing session ended the experiment, providing to the drivers explanations about the road intersections that they eventually failed to cross safely.

2.5 Calculation

Pearson correlation coefficients were calculated between the scores of the theoretical knowledge of road regulation (questionnaire) and the eventual mistakes in the virtual driving performance in the same intersections (driving performance). A 2×2 repeated-measures multivariate analysis of variance (MANCOVA) for experimental condition (familiar situations, unfamiliar situations) $\times$ type of response (questionnaire scores, simulated driving performance) was also performed, for the risk-taking behaviors in the roundabouts and for the T-junctions. Scores of self-assessed driving skills and risk perception were used as covariates in the model.

3 Results

3.1 Descriptive Statistics

Risk Perception and Self-assessment of Driving Skills. The standardized distribution of risk perception and self-assessment of driving skills are reported in Fig. 2. A linear tendency shows the variances on the scores of self-assessment of driving skills and perception of risks. Drivers that report lower scores of driving skills also perceive the driving situations as less dangerous. At the increase of values on self-reported driving skills, there is an increase in the level of perceived risks of driving situations.

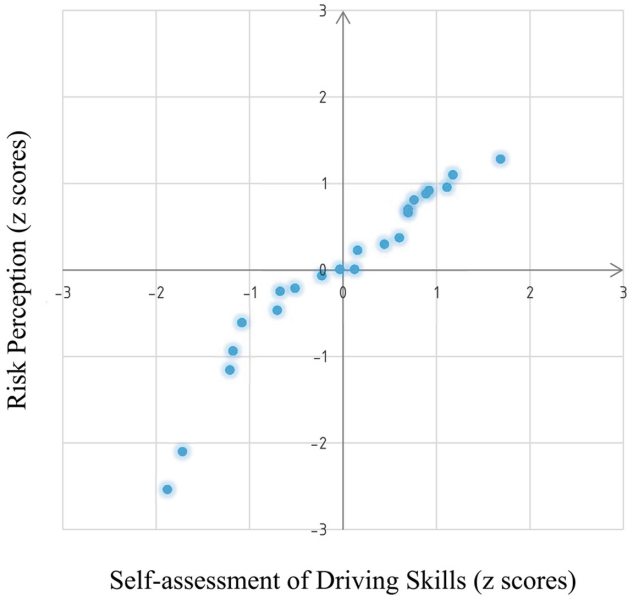


Fig. 2. Standardized scores of risk perception scale and self-assessed driving skills

Mistakes in Questionnaire and Virtual Driving. In familiar situations (a roundabout regulated by traffic light and a prohibited left turn at T-junction with red traffic light), 0% of the participants made mistakes in describing the correct behavior in road regulation questionnaire. However, in virtual driving task, 34.1% participants made a mistake in the same familiar intersections.

On the other hand, in unfamiliar situations (roundabout regulated by give way sign and a right turn at T-junction prohibited by red traffic light) 47.7% of participants provided wrong descriptions of the road regulation rule, and in the virtual driving performance 54.5% drivers made a mistake in the unfamiliar intersections (Fig. 3).

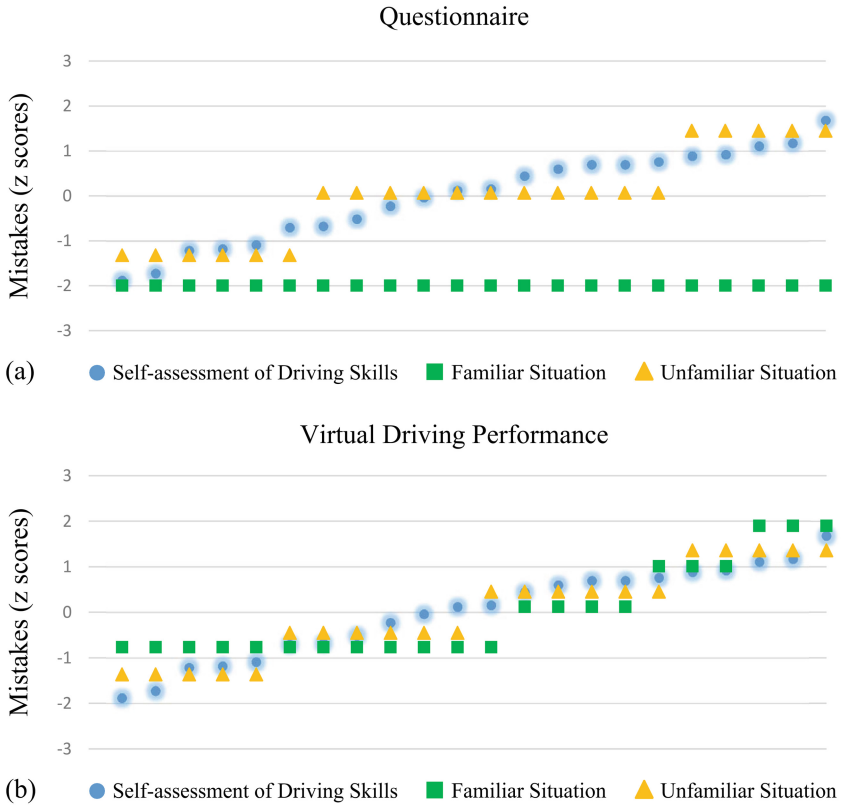


Fig. 3. Standardized scores for the mistakes in the familiar/unfamiliar situations from the (a) questionnaire and the (b) virtual driving performance. Standardized scores of self-assessed driving skills were included in the chart, as it resulted to be a significant covariate in the MANOVA

3.2 Explaining Mistakes in Questionnaire and Virtual Driving

Pearson correlation shows a significant and direct correlation only for the mistakes in the familiar and unfamiliar driving situations: $r(20) = .595$, $p = 003$. However, no direct correlation was found between the knowledge of road regulation and driving performance (H1). Results of 2 × 2 repeated-measures multivariate analysis of variance (MANCOVA) for experimental condition (familiar situation, unfamiliar situation) × type of response (answers in road regulation questionnaire, virtual driving performance), show different main effect for the type of response and an interaction with self-assessed driving skills (Fig. 3). Results of the ANOVA show that the scale of self-assessed driving skills is able to explain the different variability of the road regulation questionnaire responses versus virtual driving performance scores. Specifically, participants that presented higher scores in the self-assessment of driving skills made more mistakes in the familiar driving situations. In the roundabout, the main effect of the type of response was significant: $F(1,22) = 9.736$, $p = .005$, $\eta_p^2 = .327$,regardless of the familiarity of the

driving situation. Moreover, the interaction effect with the covariate "self-assessment of driving skills " was also significant: $F(1,22) = 5.807$, $p = .026$, $\eta_p^2 = .225$. Participants that were perceived to have higher driving skills were involved in more violations and accidents in the virtual driving task. As to the T-junction regulated by traffic light, the type of response was significant: $F(1,22) = 6.429$, $p = .020$, $\eta_p^2 = .243$, regardless of the familiarity of the driving situation, but no interaction with the covariates like self-assessment of driving skills or risk perception (Fig. 4).

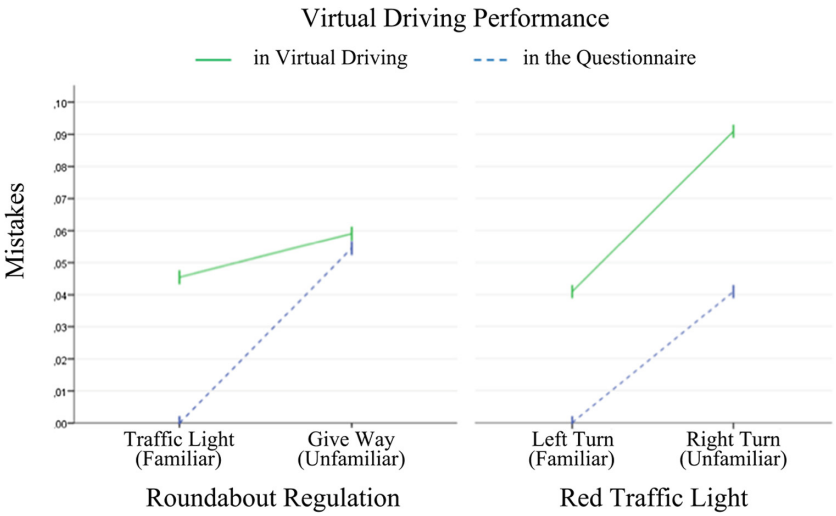


Fig. 4. Plot for the repeated-measures and multivariate analysis of variance for experimental conditions (familiar driving situation, unfamiliar driving situation) × type of responses (road regulation questionnaire, virtual driving performance), for the roundabout and T-junction.

4 Discussion

Accident involvement of young drivers revealed complex connections among risk-taking behaviors, road regulation knowledge, self-assessment of driving skills and road intersection design. In order to provide a more complex explanation of young drivers involvement in road accident while driving abroad, a systematic analysis on potential risk-taking behaviors was considered in familiar and unfamiliar situations. The main objectives of the study were: (1) understand whether road regulation knowledge correlated directly with driving mistakes in unfamiliar driving situations; and (2) whether mediating factors like risk perception and self-assessment of driving skills were able to explain the driving performance.

A simple relationship between road regulation knowledge and virtual driving performance (H1) was not found. A mediating effect of the perceived level of driving skills produced a significant interaction with driving mistakes and violations, in both familiar and an unfamiliar environment. The self-assessment of driving skills resulted

more important to explain violations and mistake compared to the risk perception scale and the theoretical knowledge of the road regulation. Drivers who rated themselves as better-than-average drivers performed higher number of violations and mistakes, in both familiar and unfamiliar road situations (H2). This result suggests that the frequent violations and mistakes that could occur under particular time-constrain circumstances while driving abroad, can be related not only to lack of road regulation knowledge, but also to some significant connections with the self-attribution of being a good driver.

The fact that all the young Chinese drivers tested, knew the correct behavior in familiar driving situations did not automatically imply that they would respect the rules, in the practical driving situations (H3). The perception of being a good or bad driver played a crucial role during the specific risk-taking decisions that occurred in the situation of "being late". Participants that rated higher scores in perceived driving skills were more prone to take risks and thus committed more traffic violations and accidents compared to participants that perceived to have less driving skills, which were more watchful. In addition, it was possible to see a difference in the type of road interaction on the driving violations. The roundabout produced a similar amount of driving mistakes in familiar and unfamiliar situations: highlighting the importance of subjective self-assessment of driving skills, in determine the regulation behavior. On the other hand, in the T-junction, a wider spread between the numbers of mistakes made in the familiar vs. unfamiliar situation was present, indicating the importance of road regulation knowledge for this specific intersection.

These indications could be used to improve the few existing training courses available for foreign drivers, as it could be possible to better focus the test on the reasons that leads to accidents. The difference between the theoretical knowledge and the practical driving performance is reporting a specific issue in young drivers abroad. Drivers should consider re-defining themselves no longer as "very good drivers" when they drive for the first time in a new driving environment. Some concrete training course could be helpful to provide more information on the road regulation rules, but also improve a new awareness of the implications of driving abroad.

These results should be contextualize considering that the present study had some limitations in the type and number of road regulation situation selected, and for the particular task related to "being late for class" tested. Future development could use a group of Italian sample as a control group, to improve the validity of the virtual simulation used and consider cross-cultural differences. Additional workload indexes related to physiological measures and eye movements could be also considered to better understand the implications of accidents involvements while driving abroad. However it was still possible to see the interplay between different variables in determine some risk-taking behaviors. Traffic violations while driving abroad can be prevented not only by providing theoretical knowledge of road regulation rules, but also considering some second levels abilities related to defensive driving [12]. Empowering the ability to anticipate and evaluate driving situation is essential for all road users and especially for young and novice drivers. Train judgment skills, perception skills and achieve a more complex awareness on driving abilities, could become crucial to improve the adoption of a safer behavior also unfamiliar driving situations.

Acknowledgments. The authors are grateful to the participants for investing their efforts in this study, also thanks to the support of I.DRIVE Lab (Interaction Between Driver, Road Infrastructure, Vehicle, And Environment – Politecnico di Milano - www.idrive.polimi.it) and Chinese Scholarship Council.

References

1. World Health Organization: Violence, Injury Prevention, World Health Organization. Global status report on road safety 2013: supporting a decade of action. World Health Organization (2013)
2. Jonah, B.A.: Accident risk and risk-taking behaviour among young drivers. Accid. Anal. Prev. **18**(4), 255–271 (1986)
3. Machin, M.A., Sankey, K.S.: Relationships between young drivers' personality characteristics, risk perceptions, and driving behaviour. Accid. Anal. Prev. **40**(2), 541–547 (2008)
4. Oltedal, S., Rundmo, T.: The effects of personality and gender on risky driving behaviour and accident involvement. Saf. Sci. **44**(7), 621–628 (2006)
5. Associazione Sostenitori Amici Polizia Stradale. it il portale della sicurezza stradale. http://www.asaps.it/
6. Singer, E., Endreny, P.M.: Reporting on Risk: How the Mass Media Portray Accidents, Diseases, Other Hazards. Russell Sage Foundation, New York (1993)
7. Sagberg, F., Bjørnskau, T.: Hazard perception and driving experience among novice drivers. Accid. Anal. Prev. **38**(2), 407–414 (2006)
8. Knuth, D., Kehl, D., Hulse, L., Knuth, D., Kehl, D., Hulse, L., Schmidt, S.: Risk perception, experience, and objective risk: a cross-national study with European emergency survivors. Risk Anal. **34**(7), 1286–1298 (2014)
9. Lund, I.O., Rundmo, T.: Cross-cultural comparisons of traffic safety, risk perception, attitudes and behaviour. Saf. Sci. **47**(4), 547–553 (2009)
10. Ciceri, M.R., Ruscio, D.: Does driving experience in video games count? Hazard anticipation and visual exploration of male gamers as function of driving experience. Transp. Res. Part F Traffic Psychol. Behav. **22**, 76–85 (2014)
11. Zuckerman, M., Neeb, M.: Demographic influences in sensation seeking and expressions of sensation seeking in religion, smoking and driving habits. Personality Individ. Differ. **1**, 197–206 (1980)
12. Lajunen, T., Summala, H.: Driving experience, personality, and skill and safety-motive dimensions in drivers' self-assessments. Personality Individ. Differ. **19**(3), 307–318 (1995)
13. Spolander, K.: Drivers' assessment of their own driving abilitiy (1983)
14. Sundström, A.: Self-assessment of driving skill—a review from a measurement perspective. Transp. Res. Part F Traffic Psychol. Behav. **11**(1), 1–9 (2008)
15. Delhomme, P.: Comparing one's driving with others': assessment of abilities and frequency of offences. Evidence for a superior conformity of self-bias? Accid. Anal. Prev. **23**(6), 493–508 (1991)
16. Freund, B., Colgrove, L.A., Burke, B.L., McLeod, R.: Self-rated driving performance among elderly drivers referred for driving evaluation. Accid. Anal. Prev. **37**(4), 613–618 (2005)
17. Svenson, O.: Are we all less risky and more skillful than our fellow drivers? Acta Psychologica **47**(2), 143–148 (1981)
18. Finn, P., Bragg, B.W.: Perception of the risk of an accident by young and older drivers. Accid. Anal. Prev. **18**(4), 289–298 (1986)

19. Matthews, M.L., Moran, A.R.: Age differences in male drivers' perception of accident risk: the role of perceived driving ability. Accid. Anal. Prev. **18**(4), 299–313 (1986)
20. Kouabenan, D.R.: Beliefs and the perception of risks and accidents. Risk Anal. **18**(3), 243–252 (1998)
21. Şimşekoğlu, Ö., Nordfjærn, T., Rundmo, T.: Traffic risk perception, road safety attitudes, and behaviors among road users: a comparison of Turkey and Norway. J. Risk Res. **15**(7), 787–800 (2012)
22. Warner, H.W., Özkan, T., Lajunen, T., Tzamalouka, G.: Cross-cultural comparison of drivers' tendency to commit different aberrant driving behaviours. Transp. Res. Part F **14**(5), 390–399 (2011)
23. Özkan, T., Lajunen, T., Chliaoutakis, J.E., Parker, D., Summala, H.: Cross-cultural differences in driving behaviours: A comparison of six countries. Transp. Res. Part F **9**(3), 227–242 (2006)
24. Lajunen, T., Corry, A., Summala, H., Hartley, L.: Cross-cultural differences in drivers' self-assessments of their perceptual-motor and safety skills: Australians and Finns. Personality Individ. Differ. **24**(4), 539–550 (1998)
25. Sivak, M., Soler, J., Tränkle, U.: Cross-cultural differences in driver self-assessment. Accid. Anal. Prev. **21**(4), 371–375 (1989)
26. Björklund, G.M., Åberg, L.: Driver behaviour in intersections: formal and informal traffic rules. Transp. Res. Part F **8**(3), 239–253 (2005)
27. Rosenbloom, T., Perlman, A., Shahar, A.: Women drivers' behavior in well-known versus less familiar locations. J. Saf. Res. **38**(3), 283–288 (2007)
28. Martens, M.H., Fox, M.R.: Do familiarity and expectations change perception? Drivers' glances and response to changes. Transp. Res. Part F **10**(6), 476–492 (2007)
29. Yanko, M.R., Spalek, T.M.: Route familiarity breeds inattention: a driving simulator study. Accid. Anal. Prev. **57**, 80–86 (2013)
30. Zhang, Q., Jiang, Z., Zheng, D., Wang, Y., Man, D.: An application of the driver behavior questionnaire to Chinese careless young drivers. Traffic Injury Prevention **14**(8), 867–873 (2013)
31. Tsai, W.: A Study on the Factors Influencing Driving Skills and Behaviors. Doctoral dissertation, Taiwan, China (2008)
32. Reason, J., Manstead, A., Stradling, S., Baxter, J., Campbell, K.: Errors and violations on the roads: a real distinction? Ergonomics **33**(10–11), 1315–1332 (1990)

Research on Blinking-Luminescence Travel Support for Visually Impaired Persons

Tomohiro Okada[1], Takashi Miyamoto[1], Shin-ichi Ito[1], Katsuya Sato[1], Norihiro Ikeda[2], Osamu Sueda[1], and Shoichiro Fujisawa[1(✉)]

[1] Institute of Technology and Science, Tokushima University, 2-1 Minamijosan-jima, Tokushima, Tokushima 770-8506, Japan
{c501542002,c501302087,s.ito,katsuyas, sueda.osamu,sfujisawa}@tokushima-u.ac.jp
[2] Kictec Inc., 150 Umegaoka, Agui, Chita, Aichi 470-2295, Japan
ikeda@kictec.co.jp

Abstract. Of the five human senses, vision is said to account for 80% of the information acquired. The presentation of light is an effective way to convey visual information. A blinking presentation especially attracts attention. A low-vision person acquires peripheral information through residual vision. In this case, light effectively supports the travel of low-vision persons. A blinking light stimulates more than a continuous light. However, a blinking light is not sufficiently compatible with the surrounding environment. This research investigates an environmentally friendly blinking pattern to arouse attention.

Keywords: Low-vision · Visual information · LED · Travel support · Sensory test · Brain waves

1 Introduction

Crossing crosswalks is one of the most dangerous situations for visually impaired persons. Crosswalk entrances located on the boundary with the roadway are among the most dangerous areas for visually impaired persons. Visually handicapped persons include blind and low-vision (LV) persons. A light presentation is an effective aid for LV persons, who account for about 80% of the visually handicapped. The guidance of visually impaired person has the result of a lot of researches [1–4]. We developed an LED block to strengthen support at crosswalk entrances. LV persons can detect the developed LED blocks with their residual vision [5, 6]. The LED block utilizes solar power and is always on. However, it is difficult to find such a LED block in the road environment because of car headlights and peripheral lights. In such situations, it is possible to facilitate discovery by blinking the light. However, the stimulation of the blinking light may be too strong. In general, there is a correlation between the stimulation of the senses and brain waves. However, we proposed a pattern of blinking in which the stimulation of the senses was low and the stimulation of the brain was high. The proposed luminescence pattern does not blink on-off but rather smoothly turns on and off. This luminescence pattern can reduce power consumption, does not interfere with peripheral

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_32

residents, and is reassuring to the movement support of LV persons. This paper verifies the difference of stimulation by various blinking luminescence patterns and proposes an effective luminescence pattern.

2 Experiment

2.1 Blinking Pattern

Fourteen blinking patterns are prepared in this experiment. Table 1 shows the combination of each blinking pattern. Blinking has three parameters (the blinking cycle, the turning off time, and the blinking method). A is an ONOFF blinking pattern and is called a simple blink. B is a blinking pattern that lightens and darkens gradually then disappears completely. It is called fade-in fade-out blinking. Blinking has three parameters (the blinking cycle, the turning off time, and the blinking method). Figure 1 presents a chart of the blinking method. Figure 2 shows the blinking device.

Table 1. Combination of blinking patterns.

	Blinking cycle [s]	Turning off time [s]	Blinking method
P1	4	1	A
P2			B
P3		1.5	A
P4			B
P5		2	A
P6			B
P7		2.5	A
P8			B
P9	5	1.5	A
P10			B
P11		2.5	A
P12			B
P13		3	A
P14			B

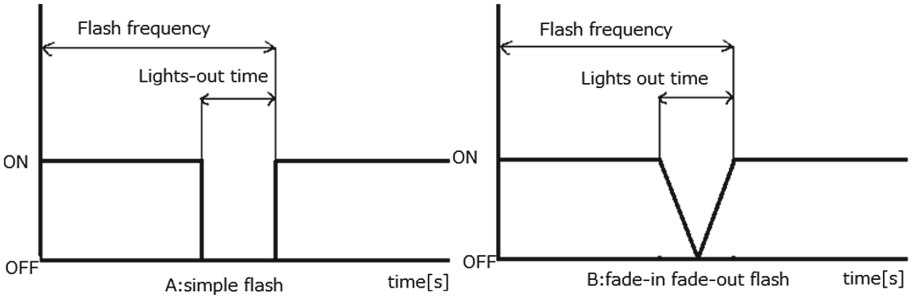
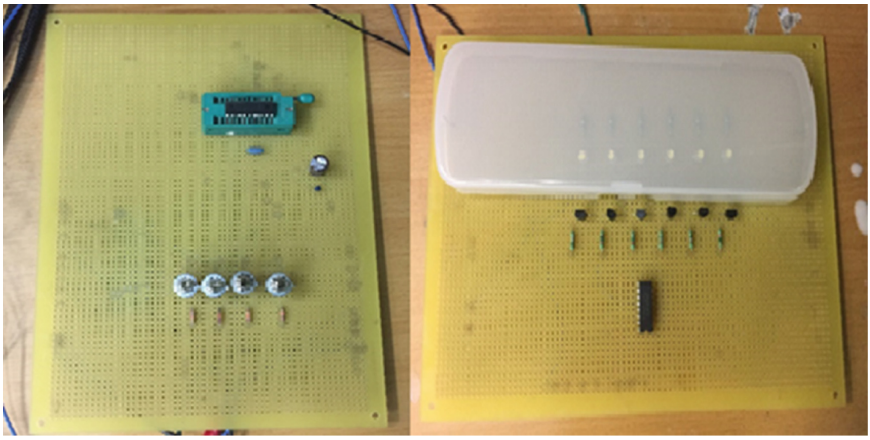


Fig. 1. Blinking method.



(a) (b)

Fig. 2. Blinking device. (a) Switch of blinking pattern; (b) luminescence device of LED.

2.2 Experiment Outline

The blinking light is presented to the subject, the brain wave is measured, and the sensory stimulation is tested.

- The subjects were 20 healthy 22.25 year olds.
- The brightness of the experiment environment was about 50 lx. (Brightness under a streetlight at night is assumed.)
- Number of blinking patterns: 14.
- Blinking is presented to the subject at random.

2.3 Experiment Procedure

Figure 3 depicts the experiment. The experiment procedure is as follows.

1. The electrode of the electroencephalograph is attached to the subject.
2. The subject adapts to the 50 lux brightness of the laboratory.
3. In the beginning, a continuous light is presented for comparison with the blinking stimulation. Afterwards, the blinking light is presented for 30 s. The brain wave at that time is measured.
4. After the blinking light is presented for 30 s, the subject evaluates the stimulation of the blinking light compared with a continuous light in seven stages (−3 to +3).
5. This experiment with 14 blinking light patterns is performed twice. The subject rests after each of the seven measurements.

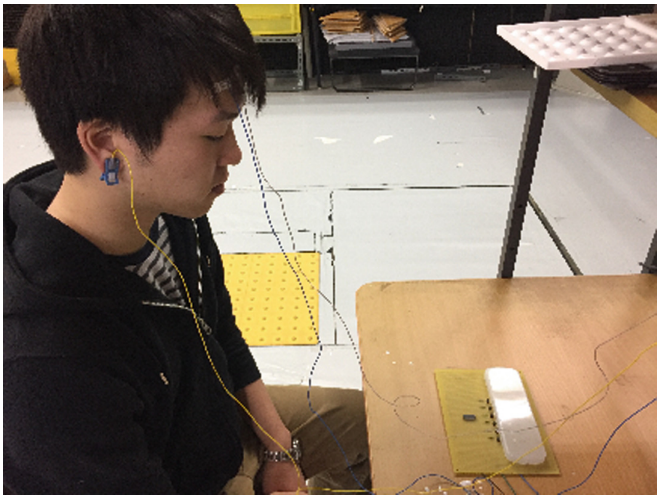


Fig. 3. Experiment representation.

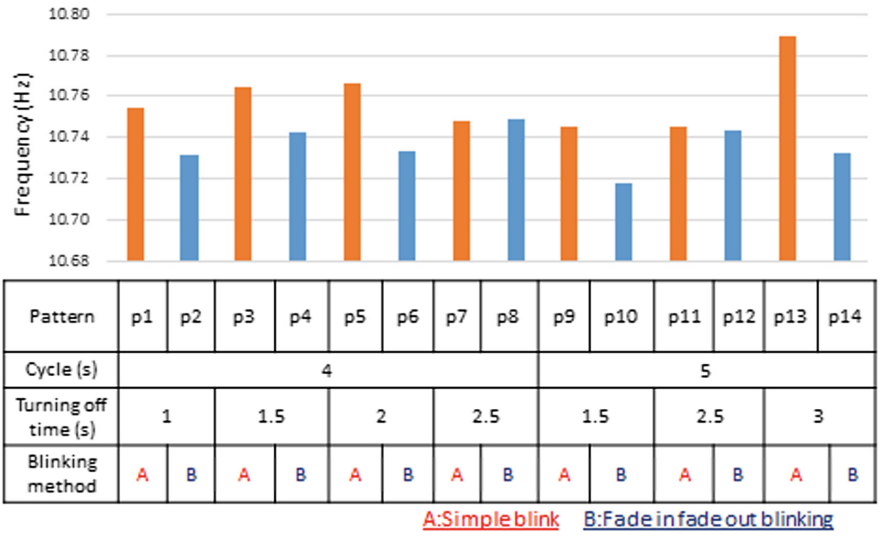


Fig. 4. Result of brain waves

2.4 Evaluation Method

This experiment adopts awake the level of the brain and person's sensory test method of evaluating two. An awake brain reacts to stimulation. Awake the level is a degree of the reaction. In the actual experiment, awake the level pays attention to the α wave of the brain wave. When the frequency of the obtained brain wave is high, awake the level is high. When the average frequency is low, awake the level is low. When awake the

level is high, the brain reacts to stimulation well and recognizes the stimulation well [3, 4, 7]. Next, the sensory test is described. The subject evaluates the strength of the stimulation of each blinking pattern based on a continuous light. The subject rated the psychological strength in seven stages.

3 Experiment Result

3.1 Result of Brain Wave Measurement

Each of the 14 blinking light patterns was measured by 20 subjects twice (28 times in total), and the mean value of the brain wave of each blinking set was calculated. Figure 2 shows the brain wave result. In the spindle, frequency (Hz) and the horizontal axis are the blinking patterns.

3.2 Result of Sensory Test

Twenty subjects measured each of the 14 blinking light patterns twice, and the mean value of each blinking pattern was calculated. Figure 5 shows the result of the sensory test. The significance level was 5%, and there were significant differences in all pairs of A and B.

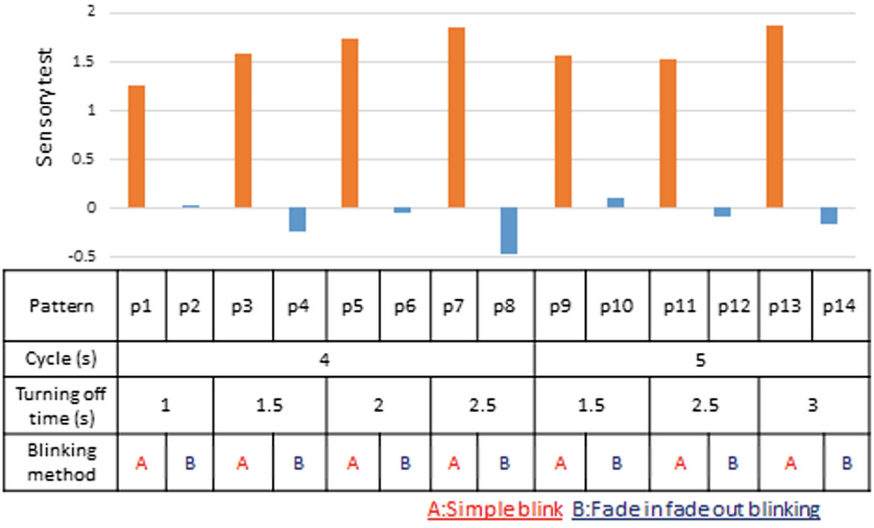


Fig. 5. Result of the sensory test.

4 Consideration

In general, awaking of the brain wave understands a simple blinking is higher than the turnoff blinking (see Fig. 4). However, there was no difference at awake the level in pairs P7–P8 and P11–P12. The cycle of these patterns was 4 s and 5 s, respectively. However, both turn-off times was 2.5 s. Moreover, simple blinking is stronger than the turnoff blinking stimulation by all patterns in the sensory test. Stimulation of the turn on and off pattern of 2.5 s (P8 and P12) seems easier than a simple blinking in the sense. However, it feels the same strength as simple blinking for awakening the brain.

5 Conclusion

This paper verified the difference of stimulation by various blinking luminescence patterns and proposed an effective luminescence pattern. In general, the stimulation of turn off and on blinking is easier than simple blinking. In addition, awaking of the brain has received stimulation equal with a simple blinking at the turn off and on time of 2.5 s. It is easy under the surrounding environment. Moreover, a blinking pattern for which awakening is high could be found. It can be expected that light will be applied to support visually impaired persons in the future.

References

1. Loomis, J.M., Philbeck, J.M., Zahorik, P.: Dissociation between location and shape in visual space. J. Exp. Psychol. **28**(5), 1202–1212 (2002)
2. Loomis, J.M., Kbapp, J.M.: Visual perception of egocentric distance in real and virtual environments. In: Hettinger, L.J., Haas, M.W. (eds.) Virtual and Adaptive Environments, pp. 21–46. Erlbaum, Mahwah (2003)
3. Stoelzel, C.R., Huff, J.M., Bereshpolova, Y., Zhuang, J., Hei, X., Alonso, J., Swadlow, H.A.: Hour-long adaptation in the awake early visual system. J. Neurophysiol. **114**(2), 1172–1182 (2015)
4. Abolafia, J.M., Vergara, R., Arnold, M.M., Reig, R., Sanchez-Vives, M.V.: Cortical auditory adaptation in the awake rat and the role of potassium currents. Cereb. Cortex **21**(5), 977–990 (2011)
5. Nagahama, H., Inagaki, T., Ikeda, N., Takahashi, K., Takeuchi, K., Ogino, H., Sato, K., Ito, S., Fujisawa, S.: Proof experiment of LED block equipped with projections to locate travel direction for blind and vision impaired persons. In: Proceedings of the AHFE 2016 International Conference on Human Factors in Transportation, pp. 159–169 (2016)
6. Ikeda, N., Takahashi, K., Inagaki, T., Sato, K., Ito, S., Seiyama, M., Fujisawa, S.: Emitting LED block at crosswalk entrance for visually impaired persons. Procedia Manuf. **3**, 3774–3778 (2015)
7. Keller, A.J., Houlton, R., Kampa, B.M., Lesica, N.A., Mrsic-Flogel, T.D., Keller, G.B., Helmchen, F.: Stimulus relevance modulates contrast adaptation in visual cortex. eLife **6**, 1–12, Article e21589 (2017)

An Emerging Framework to Inform Effective Design of Human-Machine Interfaces for Older Adults Using Connected Autonomous Vehicles

Phillip L. Morgan[1,2(✉)], Alexandra Voinescu[1,2], Craig Williams[1,3], Praminda Caleb-Solly[1,4], Chris Alford[1,2], Ian Shergold[1,5], Graham Parkhurst[1,5], and Anthony Pipe[1,4]

[1] University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol S16 1QY, UK
{phil.morgan,alexandra.voinescu,craig6.williams, praminda.caleb-solly,chis.alford,ian.shergold, graham.parkhurst,tony.pipe}@uwe.ac.uk

[2] Psychological Sciences Research Group, University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK

[3] Centre for Health and Clinical Research, University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK

[4] Bristol Robotics Laboratory, University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK

[5] Centre for Transport and Society, University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK

Abstract. Connected autonomous vehicles (CAVs) represent an exciting opportunity for wider access to mobility; especially for individuals unable to drive manual vehicles. Interaction with CAVs will be through human-machine interfaces (HMIs) providing journey-related and other information with some interactivity. These should be designed with potential users as part of a co-design process to maximize acceptance, engagement, and trust. This paper presents an emerging framework to inform the design of in-vehicle CAV HMIs with a focus on older adults (70-years+). These could be amongst early adopters of CAVs and tend to have the highest level of cognitive, sensory, and physical impairments. Whilst there are numerous principles on HMI design for older adults there are fewer on HMIs for AVs, and a need for research on CAV HMI design principles for older adults. Our emerging framework is novel and important for designers of CAV HMIs for older adults and other potential users.

Keywords: Connected autonomous vehicle · Human machine interface · Older adults · Design

1 Introduction

Connected autonomous vehicle (CAV) technology is developing rapidly. These are referred to as Level 5 AVs in frameworks of vehicle autonomy defined by SAE

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_33

International [1] as *'the full-time performance by an automated driving system of all aspects of the dynamic driving task under all roadway and environmental conditions that can be managed by a human driver'*. There is no fallback of system controls on the person(s) being driven who can have 'mind' as well as 'eyes' off the road during CAV journeys. As such, fully autonomous vehicles may not require manual input devices such as steering wheels which require a requisite level of dexterity. Instead, interaction is likely to be through screen and speech-based human-machine interfaces (HMIs) providing vehicle-related (e.g., speed, time to destination, information about local area) and other information (e.g., in-car entertainment) with perhaps some interactivity capabilities to specify and personalize journeys.

The current paper showcases work from an early phase of an Innovate UK funded project – Flourish – *Empowerment through Trusted Secure Mobility (2016–2019)*. The paper is related to one of the key project aims: exploring user interaction by focusing on the needs and experience when using CAV technology. It represents early stages of a larger framework being developed to inform the effective design of HMIs for a population sector likely to benefit from early adoption of CAVs: older adults. Within the emerging framework, older adults are defined as individuals' 70-years or above with normal aging related cognitive, sensory, and/or physical impairments. The key focus is on factors likely to determine user acceptance, engagement, trust, and likely continued usage of CAV HMIs.

2 Review Method and Research Questions

Search terms were used to guide the review (with multiple combinations) including: human machine interface, display design, autonomous vehicle, connected, driverless car, automation, elderly and older adults. Key disciplines of interest were: psychology, human factors, ergonomics, aging, gerontology, transport studies, engineering, robotics, and computer science. The search had three stages: (i) *University of the West of England-Bristol* online library database, (ii) general scholarly search engine (*Google Scholar*), and, (iii) general search engine (*Google*). The procedure resulted in multiple references deemed relevant from titles, keywords, and abstracts/summaries that were reviewed thoroughly and scored by members of the research team in terms of relevance to in-vehicle CAV HMI design questions outlined below. This involved a scale ranging from 0 (not relevant) to 5 (very relevant). Those rated 5 were regarded as key references, 3–4 very relevant, and 1–2 lower status (e.g., to support general points). References rated zero were discarded. Our CAV HMI design questions were as follows:

1. What are the leading generic HMI design principles that could inform the development of a CAV in-vehicle HMI for use by older adults?
2. What aging-related impairments should be considered in regards to the accessibility and usability of a CAV in-vehicle HMI?
3. What aging-related impairments should be considered in regards to the functionality and adaptability of a CAV in-vehicle HMI?

3 Emerging Principles

3.1 Leading Generic HMI Design Principles that Could Inform the Development of a CAV in-Vehicle HMI for Use by Older Adults

Numerous early HMIs seemed to violate many contemporary interface design principles such as providing too many features to perform a similar function(s) (e.g., find, search, locate). However, highly influential frameworks emerged such as those by Shneiderman [2], Nielsen [3], and Wickens, Lee, Liu, and Becker [4] to better guide the design of HMIs and improve human-computer interaction (HCI) by placing user experience, needs and capabilities at the center of the design process. Example principles from Shneiderman's [2] Eight Golden Rule framework include:

- Strive for consistency (e.g., include consistent terminology for menus and prompts);
- Offer informative feedback (e.g., to verify the status of a request);
- Design dialogues to yield closure (e.g., provide feedback when task completed);
- Offer simple error handling (e.g., require the user to go through checking steps that do not simply involve clicking the same button multiple times);
- Reduce short-term memory load (e.g., avoid multiple menus and/or pages).

Nielsen [3] took into consideration factors such as the importance of the match between the system and the real world, establishing and adhering to conventions and standards, maintaining visibility of system status, and simplicity and aesthetic integrity. Together with other interface design principles and frameworks (e.g., [5–7), Nielsen [3] and Shneiderman and Plaisant [8] continue to lead the way in terms of informing effective generic interface design principles for the majority of knowledgeable to expert frequent HMI users. Wickens et al. [4] developed principles with perhaps even more focus on human cognitive capabilities and limitations. These include:

- Perception (e.g., make displays legible and/or audible);
- Attention (e.g., minimize the cost of accessing important information);
- Memory (e.g., ensure key to-be-remembered information is easily accessible);
- Mental models (e.g., ensure pictorial realism between icons and what they represent);
- Situation awareness (e.g., features to predict/remind of key future and past events).

Many leading generic HMI design principles derived over the past three decades or so can be tentatively applied to the design of in-vehicle HMIs. Some can be directly applied to areas such as: automated system design; spoken input and dialog guidelines; guidelines for visual and auditory displays, and; traffic information guidelines (e.g., [9]). These and other principles will be discussed in more detail below. In terms of specific principles, work by Weir [10] noted that in-vehicle interfaces should be designed to meet requirements related to key factors such as usability, driver comfort, and acceptable levels of attentional demands in dual task conditions. It is also important to aim for optimal functionality and usability in terms of comfort (e.g., climate controls), entertainment (e.g., news links, music), telematics (e.g., network communications), and driver support (e.g., driving condition updates).

Work on HMI design for AVs has begun to increase, particularly during the latter half of the past two decades. Much early work focused on vehicles with one or few automated functions (e.g., Level 1–2). Research involving HMI designs for higher levels of vehicle autonomy (Levels 3–4) has emerged over the past 5–10-years (e.g., [11, 12]). To date, our review has uncovered very little on the design of HMIs for Level 5 fully autonomous vehicles, with a gap on those designed for older adult users.

Over a decade ago, Cuevas [13] suggested that the four most important factors to consider within the design of HMIs for vehicles with high levels of autonomy include:

- Ethnographic/anthromorphic qualities: e.g., attribute human qualities to system displays and functions;
- Cognitive factors: e.g., perception, attention, memory, and mental models;
- Predictive modelling: e.g., step-by-step modelling of interactions with new interfaces informed by established models and architectures;
- Empirical testing: i.e., test with target end-user groups.

More recent work picks-up on the HMI design issues of information overload (e.g., [14]) and situation awareness (e.g., [15]). Essentially, '…design needs to focus on communicating the information needs of drivers in order to give them the best chance of behaving appropriately for the situation' ([9], p. 143) as well as communicating the system limits in a dynamic ongoing (not static) manner [16].

Examples of recent activity on HMI design for AVs has involved 'automation displays' for functions such as adaptive cruise control (ACC) (e.g., [9]). For example, Stanton, Dunoyer, and Leatherland [17] compared three methods of displaying 'stop and go' (S&G-ACC) information to drivers under different driving conditions. Methods involved: a static icon (vehicle ahead, distance lines/arrows), a flashing icon (like static but changed color when close to vehicle ahead), and a radar display (distance ahead, road position). Simple icon-based interfaces did not seem to support driver situation awareness, whereas the more complex radar display helped but increased workload.

Overall, there is much less research on generic HMI design principles for AVs and CAVs (especially Level 5) in comparison to the wealth of literature concerning general HMI design principles that *could* inform the development of in-vehicle CAV HMIs. Many general HMI principles should nevertheless help to inform the design of early generation HMIs for CAVs, including those designed for use by older adults.

3.2 Aging-Related Impairments that Should Be Considered in Regards to the Usability and Accessibility of a CAV in-Vehicle HMI

Older age is often associated with cognitive, sensory, and mobility impairments (e.g., [18, 19]). CAV HMIs designed for use by older adults with different needs and abilities should match accessibility and usability requirements of this specific population [20, 21]. CAV HMI accessibility for older adults is related to the physical and ergonomic aspects of the in-vehicle environment whereas usability relates to aspects of the HMI including: learnability; efficiency; memorability: error handling; and satisfaction (linked with likelihood of continued use).

Cognitive aging can span multiple areas, but here we focus on attention and memory. Attention is prone to age-associated impairments such as maintaining focus on a stimuli and/or dividing focus between stimuli. General design recommendations have been proposed to increase older adults' usability of a system whilst aiming to alleviate some of the implications of age related decline in attention (e.g., [22, 23]). These include: use of simple displays; differentiating important and relevant information; minimizing screen clutter; and limiting distractor stimuli. Working memory involves trying to retain and manipulate information over short time periods usually not greater that 20–30-seconds [24]. To support age-associated decline in working memory function, recommendations for HMI design have been proposed (e.g., [23, 25]) These include: avoidance of complex long instructions to minimize memory overload; use of graphical aids to support complex tasks; labelling items clearly; and, adopting familiar conceptual models and/or metaphors (e.g., red for 'stop'). Long-term memory (LTM) is also affected by aging (e.g., [21, 23, 26, 27]). LTM is often thought of as capable of holding memories more permanently than working memory for future activation and retrieval [28]. In considering LTM issues related to aging, design should include: using simple, minimal and intuitive steps to perform tasks; offering extra practice to learn procedures; and replacing time-based instructions (e.g., perform x in 60-seconds) with event based instructions involving context specific memory cues to reduce the demands on aspects such as prospective memory.

Visual and auditory sensory impairments are more likely to occur amongst older adults (e.g., [29–31]). There is a vast amount of previous work on HMI design recommendations taking into account sensory impairments with a key focus on vision (e.g., [23, 32–35]). Example visual impairment related principles include: using large screens with large buttons (15–20-mm minimum); making information clearly visible using size, enhancing color and contrast features; spacing buttons (3.17-mm minimum); and avoiding dynamic text presentation formats. There has also been a lot of work on HMI design recommendations for auditory impairments (e.g., [21, 36–39]). Example principles include: ensuring that sound signals are at least 60 dB; enhancing discriminability of sounds with a frequency range of at least 500–1000 Hz; avoiding synthetic speech; and ensuring that key information is communicate in an auditory as well as a visual format.

Physical usability requirements for older adults often arise due to age-related changes in motor control which can affect fine motor movement and coordination [20]. To improve HMI accessibility and usability amongst older adults, a range of recommendations have been made (e.g., [23, 40–44]). These include: using haptic displays with touchscreen and voice-command capabilities; avoid simple features that require multiple click actions; avoid pointing and dragging tasks; offering a wide range of different access points within the interface; limiting difficult or long successive actions; and, using light pens for touchscreen interactions. It is also important to provide grip balustrades for support and balance when using HMIs in dynamic situations such as being driven in a car.

Overall, there appears to be a vast range of HMI principles on accessibility and usability that might inform effective design of CAV HMIs for use amongst older adults. Some are quite intuitive yet evidence informed and others are based on tested minimal requirements to ensure that HMIs are accessible and usable.

3.3 Aging-Related Impairments Should Be Considered in Regards to the Functionality and Adaptability of a CAV in-Vehicle HMI

Availability and usefulness of CAV HMI functions might increase the usability and acceptance of CAVs in general. Having the option to adapt some of these functions to better fit specific user requirements is also likely to be very important.

For AV HMIs, useful functions can increase trust and safety because they provide the user with valid and reliable driving information. Many functions that can maintain and enhance safe in-vehicle mobility have been tested and developed including: navigation aids (e.g., route guidance), visual aids (e.g., night vision enhancement), attentional and cognitive aids (e.g., distraction-management system), and crash avoidant aids (e.g., collision warnings). Common recommendations (see e.g., [28, 45–48]) include: vehicle speed and journey time information; step-by-step route guidance within navigational systems; and, minimizing complex interface features with multiple functions.

Other in-vehicle HMI functions that might provide a positive experience and keep users engaged with the system have also been recommended by a number of researchers (e.g., [49–53]). These include: night vision enhancement (NVE); forward collision warnings (FCW); health monitoring capabilities; and infotainment systems.

The usefulness and usability of any function is likely to be dependent on individual user needs and requirements and these may vary considerably amongst older adults; especially those who may have distinct cognitive and/or sensory and/or physical impairments. Also individual aging-related impairments can change, so the system should lend itself to being easily adapted to ensure continued usability.

Some in-vehicle HMIs are now becoming more adaptable (e.g., [54–56]). According to Lavie and Meyer [57], effective adaptation will involve intermediate levels of adaptation based on task type (e.g., routine vs non-routine), task difficulty, and user requirements (e.g., younger vs older adults). General adaptability recommendations have been proposed by many researchers (e.g., [57–63]). Examples include: provide text-to-speech/speech-to-text options; provide the option to zoom-in and out; and, use of persona-based user modelling. To support adaptability and personalization, the system needs to have information regarding the user's abilities and limitations which can be captured in a user model and/or through use of a reliable and valid set of tests and measures of factors (see [61, 64]). These include: cognitive (e.g., mental processing speed, attention, working memory), sensory (e.g., vision, hearing), and physical abilities.

Our preliminary review has also revealed that HMI adaptability amongst older adults does not always need to be based on cognitive, sensory and physical impairments. We should also consider driving habits, culture, and preferences. For example, and as posited by Heimgärtner [65], having the ability to control and change things like: preferred route type; preferred speed; default tours; and, preferred HMI interaction styles.

Overall, effective in-vehicle HMI functions for AVs and CAVs are likely to include important in-vehicle information such as vehicle speed and journey time as well as outer-vehicle information such as distance between own and other vehicles and traffic and news updates. Physiological monitoring (particularly for individuals with age related health conditions) is also likely to be important. Adaptability (and ease of adaptability) is also important; particularly for individuals with specific age-related cognitive and/or

sensory and/or physical impairments. This likely means that a 'one-size-fits-all' CAV HMI will not be as effective, especially amongst older users of AVs and CAVs. However, it is noted that there is little research on adaptive in-vehicle HMIs for use amongst older adults in highly and/or fully autonomous CAVs.

4 Conclusions and Future Directions

The current paper has focused on a synthesis of principles from an initial literature review conducted to inform the design of effective CAV HMIs for use by older adults above the age of 70-years. This emerging framework represents an important part of our current Innovate UK funded Flourish project that will involve experimental trials with older adults in simulated and road based CAVs in order to test and develop effective in-vehicle HMIs for use amongst this population. There are numerous useful generic HMI and in-vehicle HMI design principles that relate to four general areas: accessibility, usability, functionality, and, adaptability. The review has highlighted that experience and training in the resultant systems, as well as the ability to create bespoke individual user-focused solutions, are likely to play a key role in overcoming many of the possible barriers to adoption, use and continued use.

The review is also very important, because it has identified a major gap in the literature on AV and CAV HMI design principles that consider the needs and requirements of older adults who may be more likely to have cognitive and/or sensory and/or physical mobility impairments. There seems to be little research in this area and thus our short term aim is to further develop the current work into a comprehensive framework to inform effective design of HMIs for older adults using CAVs. This should be informed by requirements and expectations gathered from our target test population (i.e., potential users as part of a co-design process to maximize acceptance, engagement, and trust).

There is a crucial need for more concentrated research on the design and testing of in-vehicle HMIs for CAVs amongst older adults as well as other population sectors. There is a dearth of direct research in this area, and our Flourish project represents a major effort to inform this gap in knowledge and to offer effective design recommendations and solutions.

Acknowledgments. The reported research forms part of an Innovate UK research project - FLOURISH: Empowerment through Trusted Secure Mobility (2016–2019). See http://flourishmobility.com/. We thank Tracey Poole (Atkins UK) for reading an earlier draft of this paper.

References

1. SAE International.: U.S. Department of transportation's new policy on automated vehicles adopts SAE International's levels of automation for defining driving automation in on-road motor vehicles, (2016) https://www.sae.org/news/3544
2. Shneiderman, B.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley, New York (1987)
3. Nielsen, J.: Usability Engineering. Elsevier, Amsterdam (1994)

4. Wickens, C.D., Lee, J., Liu, Y., Becker, S.G.: An Introduction to Human Factors Engineering. Pearson, Hoboken, NJ (2004)
5. Gould, J.D., Lewis, C.: Designing for usability: key principles and what designers think. Commun. ACM **28**(3), 300–311 (1985)
6. Heckel, P.: The Elements of Friendly Software Design. Warner Books, Anaheim, CA (1984)
7. Mayhew, D.J.: Principles and Guidelines in Software User Interface Design. Prentice-Hall, London (1991)
8. Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction, 5th edn. Addison-Wesley Publishers, Boston (2010)
9. Walker, G., Stanton, N., Salmon, P.: Human Factors in Automotive Engineering and Design. Human Factors in Transport Series. Ashgate Publishing Ltd, Farnham (2015)
10. Weir, D.H.: Application of a driving simulator to the development of in-vehicle human–machine-interfaces. IATSS Res. **34**(1), 16–21 (2010)
11. Häuslschmid, R., Bengler, K., Olaverri-Monreal, C.: Graphic toolkit for adaptive layouts in in-vehicle user interfaces. In: Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp, 292–298 (2013)
12. Lorenz, L., Kerschbaum, P., Schumann, J.: Designing take over scenarios for automated driving: how does augmented reality support the driver to get back into the loop? In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 1681–1685. Sage Publications, Los Angeles (2014)
13. Cuevas, H.M.: An illustrative example of four HCI design approaches for evaluating an automated system interface. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 892–896. Sage Publications (2004)
14. Jämsä, J., Kaartinen, H.: Adaptive user interface for assisting the drivers' decision making. In: 6th IEEE International Conference on Cognitive Info communications (CogInfoCom), pp. 17–18 (2015)
15. Baxter, G., Besnard, D., Riley, D.: Cognitive mismatches in the cockpit: will they ever be a thing of the past? Appl. Ergon. **38**(4), 417–423 (2007)
16. Seppelt, B.D., Lee, J.D.: Making adaptive cruise control (ACC) limits visible. Int. J. Hum. Comput. Stud. **65**(3), 192–205 (2007)
17. Stanton, N.A., Dunoyer, A., Leatherland, A.: Detection of new in-path targets by drivers using stop & go adaptive cruise control. Appl. Ergon. **42**(4), 592–601 (2011)
18. Deary, I.J., Corley, J., Gow, A.J., Harris, S.E., Houlihan, L.M., Marioni, R.E., Penke, L., Rafnsson, S.B., Starr, J.M.: Age-associated cognitive decline. Br. Med. Bull. **92**(1), 135–152 (2009)
19. Freedman, V.A., Martin, L.G., Schoeni, R.F.: Recent trends in disability and functioning among older adults in the united states: a systematic review. JAMA **288**(24), 3137–3146 (2002)
20. Charness, N., Boot, W.R.: Aging and information technology use potential and barriers. Curr. Dir. Psychol. Sci. **18**(5), 253–258 (2009)
21. Fisk, A.D., Rogers, W.A., Charness, N., Czaja, S.J., Sharit, J.: Designing for Older Adults: Principles and Creative Human Factors Approaches. CRC Press, Boca Raton (2009)
22. Czaja, S.J., Lee, C.C.: Information Technology and Older Adults. CRC Press, Boca Raton (2009)
23. Farage, M.A., Miller, K.W., Ajayi, F., Hutchins, D.: Design principles to accommodate older adults. Glob. J. Health Sci. **4**(2), 2 (2012)
24. Baddeley, A.: Working memory: theories, models, and controversies. Annu. Rev. Psychol. **63**, 1–29 (2012)

25. Mulvenna, M., Carswell, W., McCullagh, P., Augusto, J.C., Zheng, H., Jeffers, P., Wang, H., Martin, S.: Visualization of data for ambient assisted living services. IEEE Commun. Mag. **49**(1), 110–117 (2011)
26. Glisky, E.L.: Changes in cognitive function in human aging. In: Riddle, D.R. (ed.) Brain Aging: Models, Methods, and Mechanisms, pp. 3–20. CRC Press, Boca Raton (2007)
27. Spencer, W.D., Raz, N.: Differential effects of aging on memory for content and context: a meta-analysis. Psychol. Aging **10**(4), 527 (1995)
28. Baddeley, A., Eysenck, M.W., Anderson, M.C.: Memory, 2nd edn. Psychology Press, Sussex (2015)
29. Dickinson, A., Arnott, J., Prior, S.: Methods for human–computer interaction research with older people. Behav. Inf. Technol. **26**(4), 343–352 (2007)
30. Lee, H.K., Scudds, R.J.: Comparison of balance in older people with and without visual impairment. Age Ageing **32**(6), 643–649 (2003)
31. West, S.K., Munoz, B., Rubin, G.S., Schein, O.D., Bandeen-Roche, K., Zeger, S., German, S., Fried, L.P.: Function and visual impairment in a population-based study of older adults. The SEE project. Invest. Ophthalmol. Vis. Sci. **38**(1), 72–82 (1997)
32. Echt, K. V., Morrell, R.: Designing web-based health information for older adults: visual considerations and design directives. In: Older adults, health information, and the World Wide Web. pp. 61–87 (2002)
33. Morrell, R. W., Echt, K. V.: Designing written instructions for older adults: learning to use computers. In: Handbook of Human Factors and the Older Adult, pp. 335–361 (1997)
34. Rubin, G.S., Roche, K.B., Prasada-Rao, P., Fried, L.P.: Visual impairment and disability in older adults. Optom. Vis. Sci. **71**(12), 750–760 (1994)
35. Watson, D.G., Maylor, E.A.: Aging and visual marking: selective deficits for moving stimuli. Psychol. Aging **17**(2), 321 (2002)
36. Gordon-Salant, S.: Hearing loss and aging: New research findings and clinical implications. J. Rehabil. Res. Dev. **42**(4), 9 (2005)
37. Pak, R., McLaughlin, A.: Designing Displays for Older Adults. CRC Press, Boca Raton (2010)
38. Strawbridge, W.J., Wallhagen, M.I., Shema, S.J., Kaplan, G.A.: Negative consequences of hearing impairment in old age a longitudinal analysis. Gerontologist **40**(3), 320–326 (2000)
39. Zajicek, M., Morrissey, W.: Speech output for older visually impaired adults. In: People and Computers XV—Interaction Without Frontiers, pp. 503–513. Springer, London (2001)
40. Aslan, U.B., Cavlak, U., Yagci, N., Akdag, B.: Balance performance, aging and falling: A comparative study based on a Turkish sample. Arch. Gerontol. Geriatr. **46**(3), 283–292 (2008)
41. Darejeh, A., Singh, D.: A review on user interface design principles to increase software usability for users with less computer literacy. J. Comput. Sci. **9**(11), 1443 (2013)
42. Dieudonné, V., Mahieu, P., Machgeels, C.: INPH, a navigation interface for motor-disabled persons. In: Proceedings of the 15th Conference on l'Interaction Homme-Machine, pp. 202–205 (2003)
43. Page, T.: Touchscreen mobile devices and older adults: a usability study. Int. J. Hum. Factors Ergon. **3**(1), 65–85 (2014)
44. Fink, J., Kobsa, A., Nill, A.: Adaptable and adaptive information provision for all users, including disabled and elderly people. New Rev. Hypermedia Multimed. **4**(1), 163–188 (1998)
45. Eisses, S.: ITS Action Plan. RappTrans. European commission D4 Final Report (2011)
46. Emmerson, C., Guo, W., Blythe, P., Namdeo, A., Edwards, S.: Fork in the road: in-vehicle navigation systems and older drivers. Transp. Res. Part F **21**, 173–180 (2013)
47. May, A., Ross, T., Osman, Z.: The design of next generation in-vehicle navigation systems for the older driver. Interact. Comput. **17**(6), 643–659 (2005)

48. Pauzie, A.: Development of ergonomic mock-ups for usability testing of in-vehicle communicating systems. Hum.-Comput. Interact. **2**, 228 (2003)
49. Arnaout, G.M., Bowling, S.: A progressive deployment strategy for cooperative adaptive cruise control to improve traffic dynamics. Int. J. Autom. Comput. **11**(1), 10–18 (2014)
50. Bekiaris, E., Panou, M., Touliou, K.: HMI for elderly and disabled drivers to get safely real-time warnings and information while driving. In: The 13th World Conference on Transport Research (2013)
51. Iulian, D.A., Leonte, M.G.: Driver warning assistant for monitoring heart rate and SpO2 using mobile phones. Appl. Mech. Mater. **656**, 395–402 (2014)
52. McMellon, C.A., Schiffman, L.G.: Cybersenior empowerment: how some older individuals are taking control of their lives. J. Appl. Gerontol. **21**(2), 157–175 (2002)
53. Reimer, B.: Driver assistance systems and the transition to automated vehicles: a path to increase older adult safety and mobility? Public Policy Aging Rep. **24**(1), 27–31 (2014)
54. Alvarez-Cortes, V., Zayas-Perez, B. E., Zarate-Silva, V. H., Uresti, J. A. R.: Current trends in adaptive user interfaces: challenges and applications. In: Electronics, Robotics and Automotive Mechanics Conference (CERMA), pp. 312–317 (2007)
55. Gonzalez-Rodriguez, M., Manrubia, J., Vidau, A., Gonzalez-Gallego, M.: Improving accessibility with user-tailored interfaces. Appl. Intell. **30**(1), 65–71 (2009)
56. Kurschl, W., Augstein, M., Burger, T., Pointner, C.: User modeling for people with special needs. Int. J. Pervasive Comput. Commun. **10**(3), 313–336 (2014)
57. Lavie, T., Meyer, J.: Benefits and costs of adaptive user interfaces. Int. J. Hum Comput Stud. **68**(8), 508–524 (2010)
58. Chung, M.K., Lee, D., Jeong, C.: The effects of zoomable user interfaces and user age in searching for a target with a mouse on a two-dimensional information space. Int. J. Ind. Ergon. **41**(2), 191–199 (2011)
59. Ferreira, F., Almeida, N., Rosa, A.F., Oliveira, A., Casimiro, J., Silva, S., Teixeira, A.: Elderly centered design for interaction–the case of the S4S medication assistant. Proced. Comput. Sci. **27**, 398–408 (2014)
60. Hanson, V. L.: The user experience: designs and adaptations. In: Proceedings of the 2004 International Cross-Disciplinary Workshop on Web Accessibility (W4A), pp. 1–11. ACM 7 (2004)
61. Mejía, A., Juárez-Ramírez, R., Inzunza, S., Valenzuela, R.: Implementing adaptive interfaces: a user model for the development of usability in interactive systems. In: Proceedings of the CUBE International Information Technology Conference, pp. 598–604 (2012)
62. Rice, M., Alm, N.: Designing new interfaces for digital interactive television usable by older adults. Comput. Entertain. (CIE). **6**(1), 6 (2008)
63. Verwey, W.B.: On-line driver workload estimation: Effects of road situation and age on secondary task measures. Ergonomics **43**(2), 187–209 (2000)
64. Browne, D., Norman, M., Riches, D.: Why build adaptive systems? In: Brown, D., Totterdell, P., Norman, M. (eds.) Adaptive User Interfaces, pp. 15–58. Elsevier, London (2016)
65. Heimgärtner, R.: Cultural differences in human computer interaction: results from two online surveys. In: ISI, pp. 145–157 (2007)

Road and Rail: Human-Machine Interaction

Identifying Difference Thresholds of Haptics and Acoustics of Control Devices

Tobias Glohr[1(✉)], Andreas Zimmermann[2], and Thomas Maier[1]

[1] Institute for Engineering Design and Industrial Design, Research and Teaching Department Industrial Design Engineering, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany
{tobias.glohr, thomas.maier}@iktd.uni-stuttgart.de
[2] MAN Truck & Bus AG, Engineering Research Electronic Systems, Dachauer Strasse 667, 80995 Munich, Germany
andreas.zimmermann@man.eu

Abstract. In the context of the holistic design and development of control devices in cars, haptics and acoustics, besides providing a clear feedback, are becoming more important as quality aspects. Previous research has therefore yielded haptic and acoustic design recommendations. However, these design recommendations include defined single parameters, but no range in which high quality is still achieved. In order to specify an adequate tolerance zone of perceived high quality for controls, a four-part study was conducted to identify the difference thresholds. The first part investigated the acoustic pitch of control devices. The second, third and fourth parts of the study investigated selected haptic parameters of rotatory and translation controls.

Keywords: Difference thresholds · Psychophysics · Control devices · HMI

1 Introduction

The challenge for today's automobile industry is to transform automobiles from individual transportation systems into a living environment. The human-machine interface is therefore becoming more important. The aim is to satisfy the customers' demand for comfort and to provide a means of differentiation between competitors. One fundamental aspect of the human-machine interface is the development of new controls to make for a high-quality interior. Beside the appearance of controls, haptic and acoustic design parameters too are growing in importance because they provide feedback for the driver. Glohr et al. [1] emphasize the necessity of a perfect match of haptics and acoustics in the course of the development of control devices in addition to the importance of an unambiguous feedback. Previous research has largely focused on the specification of haptic and acoustic parameters for a perceived high quality. Hampel [2], Reisinger [3], Anguelov [4] and Kosaka and Watanabe [5] analyzed haptic parameters and derived design recommendations for rotatory (e.g. knobs) and translation (e.g. push buttons) controls. This research usually defines parameters but no range in which high quality is still achieved.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_34

In order to specify an adequate tolerance zone for high perceived quality for controls, it is necessary to identify when differences in stimuli are also perceived as differences by humans. To quantify this difference, thresholds can be determined. The difference threshold is the intensity which is necessary to differentiate one stimulus from another [6–8], i.e. the just noticeable difference (JND) between two stimuli [6]. With regard to the formulation of design recommendations of controls, including a tolerance zone, the objective of this study is to identify the difference thresholds of essential haptic and acoustic parameters.

In previous research, Kuehner [9] has examined the difference thresholds of inertia, damping and friction of rotatory control devices. Furthermore, Doerrer [10] investigated the difference thresholds of the haptic force surge of translation controls. According to his research, sudden force fluctuations of 0.1 N cannot be detected for reaction forces between 0.5 and 2.5 N. However, the investigated range of force values does not match the parameters' values for a high perceived quality as stated in the results of Anguelov [4] and Reisinger [3]. Because the difference thresholds depend on the initial stimulus intensity [6], the results obtained for the force surge parameter cannot be brought into line with the values which are regarded as showing a perceived high quality. In addition, there are no experimental studies covering difference thresholds of other parameters of controls such as distance, torque, angle or the acoustic pitch. Hence, a four-part study was conducted to examine the difference thresholds of haptic and acoustic parameters relevant to quality perception.

2 Methods

Based on research in psychophysics, Fechner [11] developed three classic methods for identifying thresholds [12]: method of constant stimuli, method of limits and method of adjustment. However, Kuehner [13] stresses that according to Falmagne [14] these methods yield additional information and data outside the main focus on the thresholds. Furthermore, when the method of limits and processing methods are used, subjects easily detect the experimental design, and collection of data is thus not exact [13]. In the field of psychophysics there are additional adaptive methods which can be seen as a variation of the method of limits [13]. In contrast to the classic methods, adaptive methods adjust the intensity of the presented stimulus based on the capability of individual subjects and their feedback in previous trials [15]. One adaptive method is the Transformed Up/Down Method, which is a modified version of the Up/Down Method. In addition to using adaptive methods, it is also essential to choose a corresponding answer paradigm [9].

2.1 Answer Options

In the context of the classic methods for identifying thresholds mentioned above, a yes/no or greater/less/equal answer option is usually used [9]. Because of a lack of verification by the examiner, these methods suffer from a possible subjective influence of the participants. A further objective option is the Forced-Choice Method [16].

Here the participant has to choose from a number of alternatives which stimulus differs from others regardless of whether the participant actually recognizes a difference in the stimuli presented [16].

2.2 Transformed Up/Down Method

The adaptive Transformed Up/Down Method by Levitt [17] is a modified version of Dixon & Mood's [18] Up/Down Method [13]. In the original Up/Down Method the difference in intensity between the stimuli presented is increased as long as participants do not recognize any difference between the correct stimulus and the others. If the participants detect the difference in stimuli correctly, the stimulus difference in the next trial will be decreased [13, 15]. Using the Up/Down Method the probability of correct answers is 50% [13, 15]. The modified Transformed Up/Down Method enhances the level of convergence and complicates the transparency in experimental design [13]. For example, using a Transformed Up/Down Method including a 1 up/2 down rule increases the probability of correct answers to 71% [13, 15]. In this case, the difference in stimulus intensity increases after an incorrect response. In contrast, the stimulus intensity decreases after two consecutive correctly detected stimuli [15]. Figure 1 shows a possible procedure for both methods. To speed up detection of a threshold for each individual subject, the stimulus intensity can be presented in larger steps until the first turnaround point T* is reached [13]. Thus, the stimulus intensity at the first turnaround point is not taken into account [13]. The individual threshold results from the average of the stimulus intensities at the turnaround points T1–T6 [13].

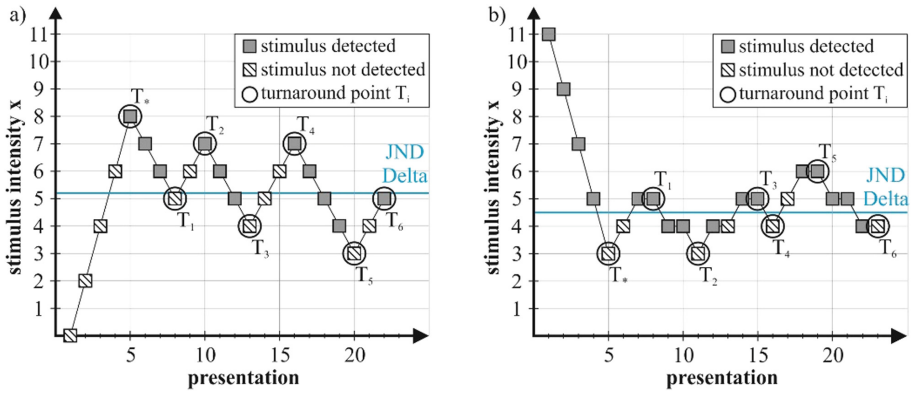


Fig. 1. Up/Down-Method (a) and Transformed Up/Down Method including 1 up/2 down rule (b) referring to [13]

A disadvantage of the 1 up/2 down rule compared to the Up/Down method is an impending increase in test duration. Therefore, in the present study the Transformed Up/Down Method has been modified (see Fig. 2). After two consecutive correct responses the stimulus intensity decreases after one correct response in the following

trail. In addition, the threshold is obtained from the average of stimulus intensities at five turnaround points (T_1–T_5). The modified Transformed Up/Down Method is used in combination with a three-alternative forced-choice method. For example, the initial stimulus intensity 3.0 N (x = 0) is presented twice and the highest stimulus intensity 5.7 N (x = 11) once. The subjects have to choose the stimulus, which differs from the two others.

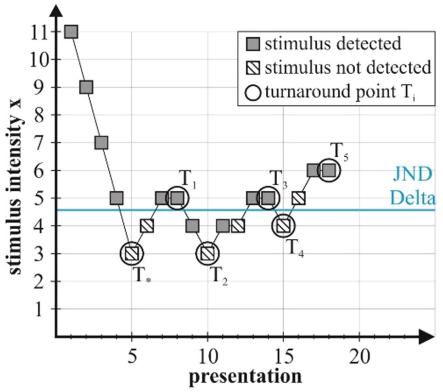


Fig. 2. Modified Transformed Up/Down Method

2.3 Identification of Haptic and Acoustic Parameters

Previous research has identified the essential parameters for high-quality perception of haptic and acoustic controls. The results of Anguelov [4], Reisinger [3] and Kosaka and Watanabe [5] show that force surge and reaction force have a remarkable influence on the haptic perception of translation controls. Anguelov [4] further emphasizes the influence of the distance of the force surge. During his examinations Reisinger [3] also in-vestigated the position/distance of the reaction force. With respect to the parametrization of rotatory controls, Anguelov [4] and Reisinger [5] state the influence of torque, number of increments and profile of the torque/angle characteristics. With regard to the acoustics Anguelov's [4] results indicate that acoustic pitch has an influence on the qualitative perception of controls. In addition, the results of Glohr et al. [1] also mention the interactions of haptics and the acoustic pitch. The complex structure of click sounds makes it hard to parameterize them [20]. For this reason those click sounds for rotatory and translation controls were used which were identified as high-quality click sounds in the results of Glohr et al. [20].

Aiming at pleasant haptic and acoustic feedback, the initial characteristics of the push-button and the knob are shown in Table 1. The chosen parameter values are based on design recommendations from the results of previous research. With regard to the rotatory controls, the knob therefore had a constant diameter of 50 mm [4], 20 detents and a constant saw tooth profile [3].

Table 1. Initial characteristics for translation and rotatory control

Parameter	Design recommendations	Chosen value
Reaction force (F1)	3.1 N–3.3 N [4], 3.0 N [3]	3.0 N
Force surge	40–55% [4], 33% [3]	47%
Distance (Force surge)	0.5–0.6 mm [4]	0.5 mm
Distance reaction force (S1)	0.5 mm [3]	0.5 mm
Pitch translation[a]	900–1000 Hz [4]. 6200 Hz [20]	6200 Hz
Torque	30–40 mNm [4], 25 mNm [3]	25 mNm
Detent	25–35 steps [4], 20 steps [3]	20 steps
Pitch rotatory[a]	900–1000 Hz [4]. 2400 Hz [20]	2400 Hz

[a]Alternative parameter: peak frequency [20]

2.4 Experimental Design

Due to potential mental fatigue of the participants during long testing, the study consists of four different parts with some changes in the subjects in each case.

In the first part the acoustic pitch of controls is tested. The second part investigates the haptic difference thresholds of rotatory controls (torque/detent). In the third and fourth parts the haptics parameters of translation controls are varied (reaction force, force surge, force surge distance, reaction force distance).

All haptic parameters are varied in 22 steps/stimulus intensity levels. These intensity levels are divided into an up and a down direction. The up direction (▲) describes an increase in the single parameter values (11 steps), while the down direction (▼) indicates a decrease in the parameter values in a further 11 stimulus intensity levels. The step range of the haptic parameters is based on a standard number row R40 with an increment of $\phi = 1.06$ in accordance with DIN 323-1 [21]. Thus, the initial parameters are varied for each stimulus intensity by the following factors: 1.06, 1.12, 1.18, 1.25, 1.32, 1.40, 1.50, 1.60, 1.70, 1.80, 1.90. For example, the variation of the reaction force (initial: F = 3.0 N) leads to a starting stimulus intensity level at x = 11 of F = 5.7 N (3.0 N × 1.90 = 5.7 N). As a result, the delta in variation concerning stimulus intensity level x = 11 is 2.7 N (5.7 N–3.0 N).

Because humans have a higher perceptual ability for acoustics than for haptics (s. [19]), the pitch is manipulated in 40 steps. The step range in this case belongs to no standard number row.

To test the difference thresholds of the selected parameters, a rotatory and a translation simulator are used, like the ones in the studies of Glohr et al. [1] and [20]. The positions of the different stimuli were varied in a randomized manner. Tables 2, 3 and 4 show the variation of the haptic and acoustic parameters for each stimulus intensity.

All four parts of the study are conducted in a pre-study experimental design, which means a large number of samples (N = 30) can be dispensed with. Instead, a minimum sample size of 12 subjects is envisaged (see [22]).

Table 2. Variation - acoustics

Parameter	Direction	Stimulus intensity	Delta variation	Stimulus intensity	Delta variation
Pitch	▼/▲	20	±400 Cents	10	±150 Cents
		19	±375 Cents	9	±125 Cents
		18	±350 Cents	8	±100 Cents
		17	±325 Cents	7	±75 Cents
		16	±300 Cents	6	±50 Cents
		15	±275 Cents	5	±40 Cents
		14	±250 Cents	4	±30 Cents
		13	±225 Cents	3	±20 Cents
		12	±200 Cents	2	±10 Cents
		11	±175 Cents	1	±5 Cents

Table 3. Variation – haptics translation

Parameter	Direction	Stimulus intensity	Delta variation	Stimulus intensity	Delta variation
Force (F1)	▼/▲	11	±2.7 N	5	±0.96 N
		10	±2.4 N	4	±0.75 N
		9	±2.1 N	3	± 0.54 N
		8	±1.8 N	2	±0.36 N
		7	±1.5 N	1	±0.18 N
		6	±1.2 N		
Force surge	▼/▲	11	±42.0%	5	±15.0%
		10	±38.0%	4	±12.0%
		9	±33.0%	3	±8.0%
		8	±28.0%	2	±6.0%
		7	±24.0%	1	±3.0%
		6	±19.0%		
Distance (S1)	▼/▲	11	±0.45 mm	5	±0.16 mm
		10	±0.40 mm	4	±0.13 mm
		9	±0.35 mm	3	±0.09 mm
		8	±0.30 mm	2	±0.06 mm
		7	±0.25 mm	1	±0.03 mm
		6	±0.20 mm		
Distance (Force surge)	▼/▲	11	±0.45 mm	5	±0.16 mm
		10	±0.40 mm	4	±0.13 mm
		9	±0.35 mm	3	±0.09 mm
		8	±0.30 mm	2	±0.06 mm
		7	±0.25 mm	1	±0.03 mm
		6	±0.20 mm		

Table 4. Variation – haptics rotatory

Parameter	Direction	Stimulus intensity	Delta variation	Stimulus intensity	Delta variation
Torque	▼/▲	11	±23.0 mNm	5	±8.0 mNm
		10	±20.0 mNm	4	±6.0 mNm
		9	±18.0 mNm	3	±5.0 mNm
		8	±15.0 mNm	2	±3.0 mNm
		7	±13.0 mNm	1	±2.0 mNm
		6	±10.0 mNm		
Detent	▼/▲	11	±18 steps	5	±6 steps
		10	±16 steps	4	±5 steps
		9	±14 steps	3	±4 steps
		8	±12 steps	2	±2 steps
		7	±10 steps	1	±1 steps
		6	±8 steps		

2.5 Analysis

As mentioned, the individual difference threshold of one subject results from the average of stimulus intensities at the five turnaround points (T1–T5) (see Sect. 2.2). The final difference threshold for each parameter results from the average of all test persons (see Sect. 3.1).

Since the overall goal of the study is the identification of adequate tolerance zones for the haptics and acoustics of controls, this analysis method has to be modified. Clarification is provided in the following example: The average identified difference threshold concerning Force F1 is 3.7 (rounded 4). In this case, however, 30% of the participants are unable to perceive a difference in stimuli of this size.

Thus, an analysis using the average would presumably exclude a high percentage of the population. An additional analysis is therefore included using the 95th percentile as cut-off. This procedure ensures that statistical outliers (correction level = 5%) are not taken into account and can be used to determine the limits of the tolerance zone (see Sect. 3.2).

If the participant's perception exceeds the limits of the presented stimulus intensities, the analysis takes into account x = 12 (haptics)/x = 21 (acoustics) concerning the upper limit and x = 0 (at the respective presentation point). The investigation for the single parameter ends at the presentation point, where the limit is reached. Further possible open turnaround points will also be counted as x = 12 (haptics)/x = 21 (acoustics) concerning the upper limit and x = 0 (haptics/acoustics) concerning the lower limit.

3 Results

All parts of the study were conducted with N = 13. The first part (acoustics) of the study included participants which were on average 32.2 years old (SD = 10.8, 5 females and 8 males). In the second part concerning the thresholds of rotatory parameter, the subjects were on average 34.5 years old (SD = 12.4, 7 females and 6 males). In the third and fourth parts of the study (translation) the subjects were on average 32.7 years old (SD = 10.9, 6 females and 7 males).

3.1 Difference Thresholds

To sum up, except for reaction force distance (S1), difference thresholds could be detected. The difference threshold for the reaction force distance in the down direction is outside the tested stimulus intensity (x > 11). Furthermore, the results for distance S1 for the up direction show a just barely achieved stimulus intensity (x = 11). The results moreover indicate that the stimulus intensities for the up and down directions are nearly the same for torque, translation force and force surge. The most significant difference between up and down direction is that for force surge distance.

Two participants were able to identify the smallest (down direction) difference stimuli presented for the acoustic pitch (rotatory). Analysis of reaction force distance (S1) data shows that several participants were unable to differentiate even the largest presented difference in stimuli: 11 (up direction), 5 (down direction). In addition, 5 participants were unable to differentiate even the largest (up direction) presented difference stimuli for force surge distance.

Tables 5, 6 and 7 show the detailed results of the difference thresholds for all investigated parameters.

Table 5. Difference thresholds – acoustics (first part)

Parameter	Direction	Stimulus intensity[a] (identified)	Stimulus intensity (rounded)	JND delta	Absolute value[b]
Pitch (translation)	▼	12.6	13	225 Cents	6810 Hz
Pitch (translation)	▲	10.5	11	175 Cents	5420 Hz
Pitch (rotatory)	▼	6.1	7	75 Cents	2460 Hz
Pitch (rotatory)	▲	4.0	4	30 Cents	2300 Hz

[a]20 steps of stimulus intensities
[b]Alternative parameter: peak frequency [20]

Table 6. Difference thresholds - haptics rotatory (second part)

Parameter	Direction	Stimulus intensity (identified)	Stimulus intensity (rounded)	JND delta	Absolute value
Torque	▼	5.9	6	10 mNm	35 mNm
Torque	▲	5.6	6	10 mNm	15 mNm
Detent	▼	5.7	6	8 steps	28 steps
Detent	▲	4.6	5	6 steps	14 steps

Table 7. Difference thresholds - haptics translation (third & fourth part)

Parameter	Direction	Stimulus intensity (identified)	Stimulus intensity (rounded)	JND delta	Absolute value
Force (F1)	▼	3.7	4	0.75 N	3.75 N
Force (F1)	▲	3.7	4	0.75 N	2.25 N
Force surge	▼	4.7	5	15.0%	62.0%
Force surge	▲	4.9	5	15.0%	32.0%
Distance (S1)	▼	11.3	12	N/D	N/D
Distance (S1)	▲	10.4	11	0.45 mm	0.05 mm
Distance (Force surge)	▼	9.8	10	0.40 mm	0.90 mm
Distance (Force surge)	▲	5.3	6	0.30 mm	0.20 mm

3.2 Recommendations Concerning Tolerance Zones

Based on the correction level of 5%, data of 95% of subjects are taken into account (N = 12). The results of the 95% solution show that with regard to the reaction force distance (S1) both thresholds (up/down directions) are outside the tested stimulus intensity ($x > 11$). Like the analysis using averages, this analysis (95th-percentile) also identifies the most significant difference between up and down direction in force surge distance. In addition, the stimulus intensity for up and down direction is the same for torque and translation force. Taken together the two analyses (average vs 95th-percentile) differ mainly in the absolute limits of the tolerance zone (see Tables 8, 9 and 10).

Table 8. Recommendations concerning tolerance zones – acoustics (95%)

Parameter	Direction	Stimulus intensity[a] (identified)	Stimulus intensity* (rounded)	Delta	Absolute value[b]
Pitch (translation)	▼	17.6	18	350 Cents	7610 Hz
Pitch (translation)	▲	16.2	17	325 Cents	5310 Hz
Pitch (rotatory)	▼	8.2	9	125 Cents	2520 Hz
Pitch (rotatory)	▲	6.8	7	75 Cents	2280 Hz

[a]20 steps of stimulus intensities
[b]Alternative parameter: peak frequency [20]

Table 9. Recommendations concerning tolerance zones - haptics rotatory (95%)

Parameter	Direction	Stimulus intensity (identified)	Stimulus intensity (rounded)	Delta	Absolute value
Torque	▼	8.8	9	18 mNm	43 mNm
Torque	▲	8.8	9	18 mNm	7 mNm
Detent	▼	6.8	7	10 steps	30 steps
Detent	▲	5.8	6	8 steps	12 steps

Table 10. Recommendations concerning tolerance zones - haptics translation (95%)

Parameter	Direction	Stimulus intensity (identified)	Stimulus intensity (rounded)	Delta	Absolute value
Force (F1)	▼	4.8	5	0.96 N	3.96 N
Force (F1)	▲	4.6	5	0.96 N	2.04 N
Force surge	▼	7.8	8	28.0%	75.0%
Force surge	▲	6.4	7	23.0%	24.0%
Distance (S1)	▼	12	12	N/D	N/D
Distance (S1)	▲	12	12	N/D	N/D
Distance (Force surge)	▼	11	11	0.45 mm	0.95 mm
Distance (Force surge)	▲	8	8	0.30 mm	0.20 mm

4 Discussion

In view of the pre-study characteristics of the current experimental design, DIN ISO 8587 [22] states, that the number of subjects between 12 and 15 is sufficient.

The specification of thresholds must be regarded as critical. It is essential to gain information about the experimental design and the way the test has been conducted. For example, the thresholds for vibration can depend on contact surface or force [9].

Overall, the results for the difference thresholds provide valuable information about possible tolerance zones within the perceived high quality of control devices such as knobs and push-buttons in vehicles. This simplifies the specification and the compliance of haptics and acoustics with it. In addition, the results for recommendations give a useful initial position concerning haptic and acoustic limit values for further variation studies. Furthermore, both results show that the distance of reaction force has no significant influence on the qualitative haptic perception. Because of this lack of differentiation in perception, the distance of reaction does not have to be considered in further investigations concerning variation.

References

1. Glohr, T., Schynoll, S., Zimmermann, A., Maier, T.: Interaction of Haptics, Acoustics and Optics in the Perception of Rotatory Control Devices. In: Soares, M., Falcão, C., Ahram, T. Z. (eds.), Advances in Ergonomics Modeling, Usability & Special Populations. In: Proceedings of the AHFE 2016 International Conference on Ergonomics Modeling, Usability & Special Populations, pp. 599–611, Springer, Orlando (2016)
2. Hampel, T.: Untersuchungen und Gestaltungshinweise für adaptive multifunktionale Stellteile mit aktiver haptischer Rückmeldung. dissertation, Universität Stuttgart (2011)
3. Reisinger, J.: Parametrisierung der Haptik von handbetätigten Stellteilen. dissertation, Technische Universität München (2009)
4. Anguelov, N.: Haptische und akustische Kenngrößen zur Objektivierung und Optimierung der Wertanmutung von Schaltern und Bedienfeldern für den Kfz-Innenraum, dissertation, Technische Universität Dresden (2009)
5. Kosaka, H., Watanabe, K.: Reaction Forces of Switches and Push Feeling, SAE Technical Paper Series 960945, Publication No. SP-1155, Warrendale (1996)
6. Handwerker, H.O.: Allgemeine Sinnersphysiologie. In: Schmidt, R.F., Lang, F., Thews, G. (eds.) Physiologie des Menschen mit Pathophysiologie, pp. 273–294. Springer, Heidelberg (2005)
7. Ansorge, U., Leder, H.: Wahrnehmung und Aufmerksamkeit. VS Verlag, Wiesbaden (2011)
8. Meilgaard, M.C., Civille, G.V., Caar, B.T.: Sensory Evaluation Techniques. CRC Press, Boca Raton (2007)
9. Kuehner, M.: Haptische Unterscheidbarkeit mechanischer Parameter bei rotatorischen Bedienelementen. dissertation, Technische Universität München (2014)
10. Doerrer, C.: Entwurf eines elektromechanischen Systems für flexibel konfigurierbare Eingabefelder mit haptischer Rückmeldung, dissertation, Technische Universität Darmstadt (2003)
11. Fechner, G.T.: Elemente der Psychophysik. Breitkopf & Härtel, Leipzig (1989)

12. Gescheider, G.A.: Psychophysics – The Fundamentals. Lawrence Erlbaum Associates, Mahwah (1997)
13. Kuehner, M., Bubb, H., Bengler, K., Wild, J.: Adaptive Verfahren in der Psychophysik: Effiziente Bestimmung von Absolut- und Unterschiedsschwellen. Ergon. aktuell **13**, 26–30 (2012)
14. Falmagne, J.-C.: Psychophysical measurement and theory. In: Boff, K.R., Kaufman, L., Thomas, J.P. (eds.) Handbook of Perception and Human Performance, Volume 1 - Sensory Processes and Perception, pp. 1–1–1-66. Wiley-Interscience, New York (1986)
15. Kingdom, F.A.A., Prins, N.: Psychophysics: A Practical Introduction. Academic Press, Amsterdam (2010)
16. Ehrenstein, W.H., Ehrenstein, A.: Psychophysical Methods. In: Windhorst, U., Johansson, H. (eds.) Modern Techniques in Neuroscience Research, pp. 1211–1241. Springer, Berlin (1999)
17. Levitt, H.: Transformed Up-down Methods in Psychoacoustics. J. Acoust. Soc. Am. **49**, 467–477 (1971)
18. Dixon, W.J., Mood, A.M.: A method for obtaining and analyzing sensitivity data. J. Am. Stat. Assoc. **43**(241), 109–126 (1948)
19. Kunsch, K., Kunsch, S.: Der Mensch in Zahlen. Eine Datensammlung in Tabellen mit über 20.000 Einzelwerten, 3. ed, Elsevier, München 2007
20. Glohr, T., Hasenberger, E., Zimmermann, A., Maier, T., Krump, G.: Untersuchungen zur Betätigungsakustik als Aspekt der Wertigkeit von Bedienelementen im Fahrzeug. 63. Frühjahrskongress der Gesellschaft für Arbeitswissenschaften, Brugg-Windisch, GFA Press, Dortmund (2017)
21. DIN 323-1: Normzahlen und Normzahlreihe. Deutsches Institut für Normung, Beuth, Berlin (1974)
22. DIN ISO 8587: Sensorische Analyse – Prüfverfahren – Rangordnungsprüfung. Deutsches Institut für Normung, Beuth, Berlin (2010)

Digital Touchpoints in Campus Slow Traffic Service System

Jintian Shi[1(✉)] and Ke Ma[2]

[1] College of Architecture and Urban Planning, Tongji University,
No. 1239 Siping Road, Shanghai 200092, China
shijintian1017@126.com
[2] College of Design and Innovation, Tongji University,
No. 281 Fuxin Road, Shanghai 200092, China
suta@tongji.edu.cn

Abstract. Recently in China, the scale of university campus has increasingly expanded, both campus buildings and road networks become complex and compound. Besides this phenomenon and allowing for the uniqueness of campus, how to effectively address the issue of campus transportation deserves our deep thinking. This paper took Jiading Campus of Tongji University as a practical case, after some design research on the campus, it turned out that slow traffic service system was a proper campus transportation mode for Jiading campus. And this campus slow traffic service system was designed to construct on a digital touchpoints architecture, based on service-oriented landmarks (SOLs). This paper aims at exploring how to design these digital touchpoints and organize their association to provide people better campus slow traffic services.

Keywords: Service system · Campus transportation · Digital touchpoints · IoT

1 Introduction

When we enjoy the convenience vehicle-dominant modern traffic mode brought, we also must face the problems caused by the automotive, such as environmental pollution, high-frequency traffic safety accidents, worse conflict between cars and people. University campus, which used to be quiet and orderly, also can't avoid these problems. Campuses are very distinctive communities [1]. The environment, stakeholders, and activities are different with the conditions of other areas in cities. And an increasing number of universities, especially in China, have established their new big campuses within more multiple services, which poses a challenge to campus transportation system. After quantities of research based on features of campuses, we found that although most campuses do not totally exclude the automobile, slow traffic mode is the expected way to get around even though other ways of transportation may also be possible [1]. While in campus slow traffic system, how to accommodate human needs for transportation? How to develop users' needs to effective services? How to evidence services into digital touchpoints? And how to organize all the digital touchpoints systematically for a better service quality?

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_35

This paper would give answers to the above questions by elaborating the slow traffic service system for Jiading Campus of Tongji University, from the perspective of service system design thinking.

2 Needs for Campus Slow Traffic

As to campus transportation problem, it seems to have various solutions. But for most of the current university campuses, slow traffic mode is supposed to be a primary option. College campuses encouraged a modal shift from cars to other modes, in particular to bicycle and walking; its main purpose is to reflect on the opportunity to create sustainable campuses from a bicycle and pedestrian planning viewpoint [1, 2]. There are some urgent needs of campuses for slow transportation service system.

2.1 Environmental Requirements

Slow traffic is also called as non-motorized transportation, and it mainly refers to walking and cycling, whose speed is less than 15 km per hour [3]. In some urban mobility systems, buses are also incorporated into urban slow traffic system as a complementary part, to connect with walking traffic and cycling network. Slow traffic mode encourages people to adopt the transport of "walking and bus" or "bicycle and bus" to alleviate traffic jam and call for environment protection.

Campus as an important educational institute in the city, has also in demand of a sustainable traffic mode for a healthier campus micro-climate.

2.2 Object-Oriented Needs

Due to the uniqueness of campus, slow traffic model seems to have more advantages over other modes. Universities in modern cities have the opportunity to improve important societal measures of innovativeness, equity, sustainability, and resiliency and become the leaders of regional economic development [4]. Research and projects on non-motorized transportation planning is still meager if compared with research on other modes [5]. Campus, as experimental platform for social innovation, should discover and optimize the prospective mobility pattern in the future. Meanwhile, comparing with urban environment, the stakeholders and using contexts in campus are relatively simple, so that most transportation needs could be meet by slow traffic.

Some research on urban slow mobility system could not be directly applied into campus, due to differences on road networks, flow of people, vehicle size and so on. Slow traffic solutions specific to campus have lots of possibilities to explore.

2.3 Social Demands

"Maybe we don't need service design, but service is actually everywhere [6] ". In the current service-economy time, we must integrate the existing resources to innovate for meeting common interests of all related stakeholders, standing on the point of service

system logic. The current society is a complex world full of wicked problems and uncertainties, and the interrelations among different aspects become stronger yet harder to figure out. Just as suggested in the above paragraphs, university campus demands slow traffic mode, but how to construct and conduct this campus transportation system? Traffic system is not just a single vehicle mode, but an integrated network containing interactive and collaborative transportation means, stakeholders, activities and various touchpoints. The service system thinking could help campus slow traffic offerings to be conducted more effectively.

3 Service System Approach in Campus Slow Traffic

And how to use service system approach to design campus slow traffic? Service design follows the working principle of being user-centered, co-creative, evidencing, sequencing and holistic [6]. And as for service system design, it focuses more on the associations among divergent services and the interrelationships between sub-systems [7]. Service system design has many methods supporting design process. The design approaches of service could be classified into four clusters according to design activities, representations, recipients and contents [8]. Different approaches could help service system design procedure to develop in a consistent progress. Service system design consists of the service front stage which focuses on the user experience of customers, and service back stage dealing with the operation of service providers and management of the whole system. The core of innovativeness of service is not just on the user experience on front stage, but the systematical and personnel organizational innovation at back stage.

In slow traffic network, there are many complicated interrelationships lying among various aspects, including divergent transportation means, multiple needs, different stakeholders and kinds of scenarios. The association and connection among these various aspects in the whole system are the key factors influence the traffic service quality. And touchpoints acting as bridges, are the inevitable media of this connection.

Touchpoints in Service System. Touch-points are the points of contact between a service provider and customers. A customer might utilize many different touch-points as part of a use scenario (often called a customer journey) [9]. As one of the five features of service is "evidencing", touch points are the vital media to make services materialized and explicit. Touchpoints in service system could be classified into three kinds: physical touchpoints, emotional touchpoints [10] and digital touchpoints. Physical touchpoints mainly refer to some tangible contacting media connecting users with service system; and emotional touchpoints, more like a kind of emotional contact, always exist between service customers with providers.

As to digital touchpoints, Morrissey deemed that in online marketing, there were four kinds of normal digital Touchpoints-Web sites, e-mails, display advertising and paid search-were evaluated on their power to boost the effectiveness of consumer engagement [11].

While in service system, physical touchpoints, digital touchpoints and emotional ones, they integrate mutually and coexist between service receivers with providers. The

digital touchpoints in this paper mainly not only refer to virtual digital interfaces, but also include some physical digital devices could be interacted with people.

How to establish an effective slow traffic service system for campus? And how digital touchpoints perform in this system to create a better service? In the following we would take the slow traffic service design for Jiading Campus of Tongji University as an example, to explore how a digital touchpoint network which is based on service-oriented landmarks, to capture service-defined ambient data to generate and optimize traffic services for people.

4 Slow Traffic Service System of Jiading Campus

Jiading campus is in Jiading District, located in northwestern of Shanghai, and takes on the first place of largest campus with 167 hectares, among five campuses of Tongji University. There are scattered functional zones and multiple road network in this large campus, which requires an effective transportation system to meet daily travel demands (Fig. 1).

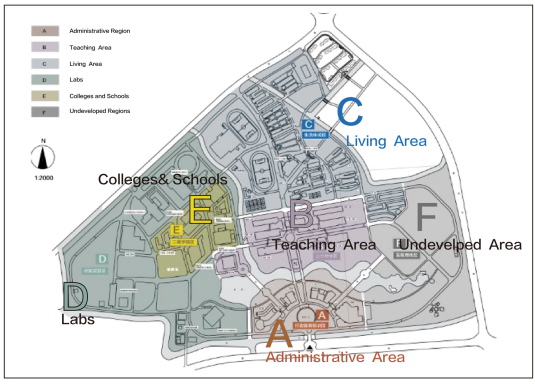


Fig. 1. Functional zones of Jiading Campus

4.1 Existing Problems in Campus Traffic

After a series of field study of the current situation of slow traffic, it was found that the current campus transportation mode was pedestrian-and-vehicle coexistence system, and some problems (Fig. 2) exited in this system: the walking paths were in lack of the mutual consistence and adequate way-finding signage; quantities of random-parking bicycles occupied much space; no public buses call for the long-distance commute resulting from the big size of campus; the travel experience in campus remained to be optimized. Then we conducted some research on the campus to explore what kind of transportation mode could meet users' needs.

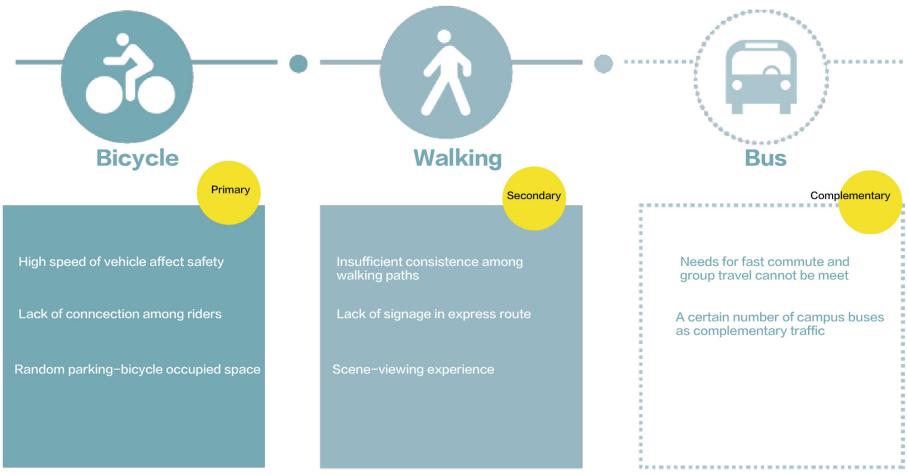


Fig. 2. Existing problems in the traffic system of Jiading Campus

4.2 User Research

No matter which kind of campus traffic system was adopted, its aim was always to offer services to people to satisfy their commute demands, therefore a comprehensive understanding of our users is inevitable to get a proper solution. Based on the current problems mentioned above, we used some quantitative and qualitative user research methods to discover opportunities.

Observation. Pretending as normal pedestrians, we looked for suitable "the observed", whose reaction was explicit to capture and interpret. Eventually some observed passerby contributed much to the results were selected to main objectives.

Field Statistics. The statistic work was deployed at five points with large people traffic flow in the campus. Altogether four statistics at each point in every 15-minute shift, and each time last 15 min. The statistics was taken on weekdays and weekends. This method was helpful to evaluate average people and vehicle traffic flow, which would provide decision basis for the distribution of digital touchpoints (Fig. 3).

Questionnaires. We designed a questionnaire, consisting of twenty-three fixed questions and another two open questions, to investigate the current situation of slow traffic in Jiading Campus. The questionnaire was filled out anonymously and issued both through online and offline.

Interview. After observation and questionnaires, some original data was collected. Based on these data, we could establish a general command of the traffic situation. And one-by-one interview at this stage could give us more chances to know the deeper motivations and reasons behind this traffic situation.

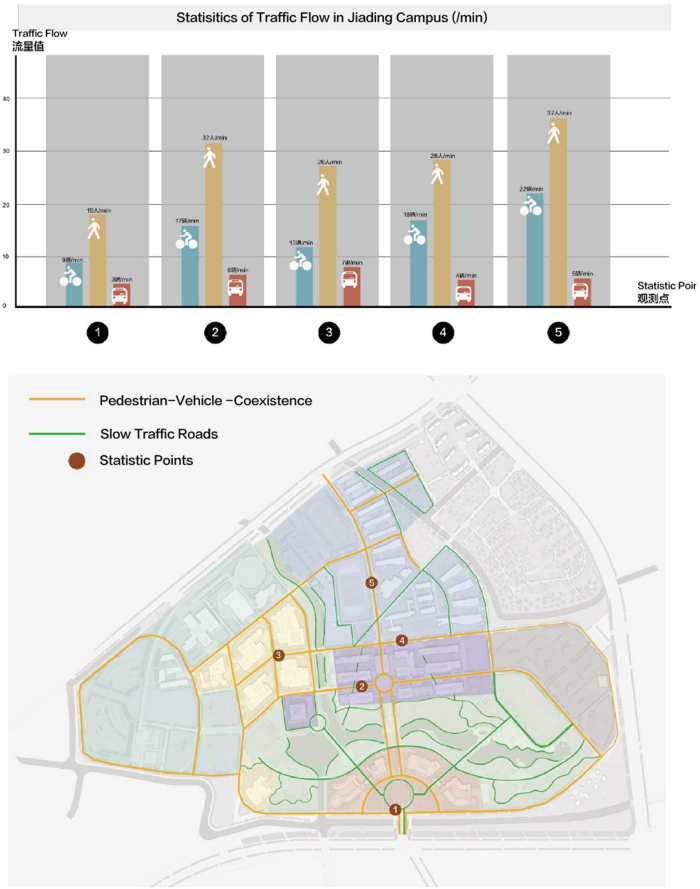
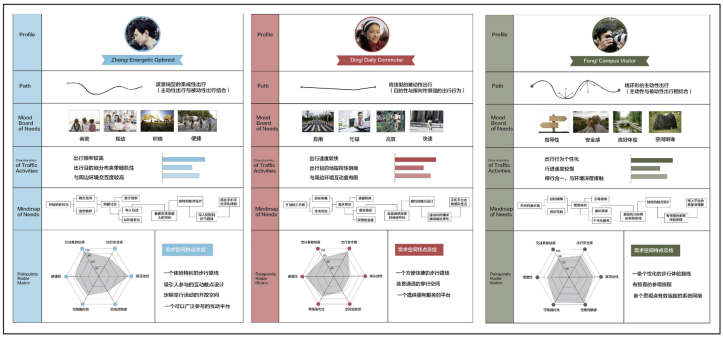


Fig. 3. Statistics of traffic flow (pedestrians, riders, vehicles) in Jiading Campus

Personas. Alan Cooper proposed personas as a tool to model user experience in interactive software design [12]. Parker and Heapy argued that in the context of service design "Creating for 'real' people can really animate discussions and enable service providers and users to engage emotionally as well as rationally [13]. The personas are then used to better understand the experiences and needs of staff."

In this project, our final users were students, teachers and administrative staff, and visitors altogether theses three clusters. Based on their characteristics, we set up three personas to get a more holistic understanding of final users. A multiple layer of traffic needs from different personas were done (Fig. 4).

After these user research, it was found that a slow traffic service system was required by Jiading Campus. This system was supposed to provide the students and staffs, even also visitors, with transportation services for their conscious or unintentional travel activities. This system could create a better walking and cycling experience: it could make daily commute more convenient and faster, give more customized guidance of



Profile	Zhang/ Energetic Optimist	Ding/ Daily Commuter	Feng/ Campus Visitor
Needs for Campus Traffic Service	Spcial Exprience Fancy Interactive Touchpoints Open Space with Regular Activities Strong-Social-Tie Campus	A Convenient Traffic Path An Easy Access to Service Points A Bicycle System support booking Swift Feedback from Services	A Customized bicycle Path Reasonable Distribution of Service Points Devices Easy to Use Service System Inclusive to visitors

Fig. 4. Personas and different needs for campus traffic service

campus touring, strengthen the social tie among users through services, and get a deeper semantic understanding of Jiading Campus.

4.3 The Slow Traffic Service System Design

Just as Fig. 5 the slow traffic service system for Jiading Campus consists of three sub-systems: walking, bicycling, and bus sub-system as complementary. In this integral system, the walking sub-system and bicycle sub-system take on the core place, and mutual enrichment occurred among these three sub-systems.

The overall service system has been built up, and as the key factors to perform services, the operation of digital touchpoints in the system deserves explorations. For creating better services to meet users' needs, various digital touchpoints in three different sub-systems were designed to interact and connect mutually, and they were arranged in a systematic network. And the following fifth part would be elaborated on how these digital touchpoints function in the whole service system.

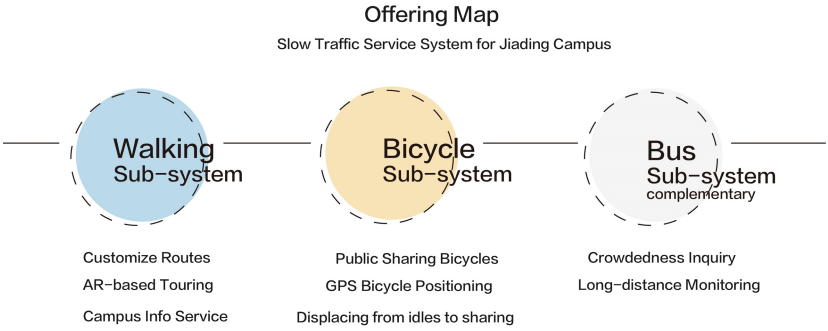


Fig. 5. Offerings of slow traffic service system for Jiading Campus

5 Digital Touchpoints in the Service System

5.1 Touchpoints Architecture

We designed a digital touch points architecture to facilitate the slow traffic service system of Jiading campus, which includes key touchpoints and complementary touchpoints. The key touchpoints, serving as a cross-system media, provides fundamental service in all three sub-systems, while the complementary touchpoints are designed to offer customized service for each sub-system. Together with these two components, the defined sub-systems could generate and optimize effective slow traffic service for people. Figure 6 illustrates the overall touchpoints architecture.

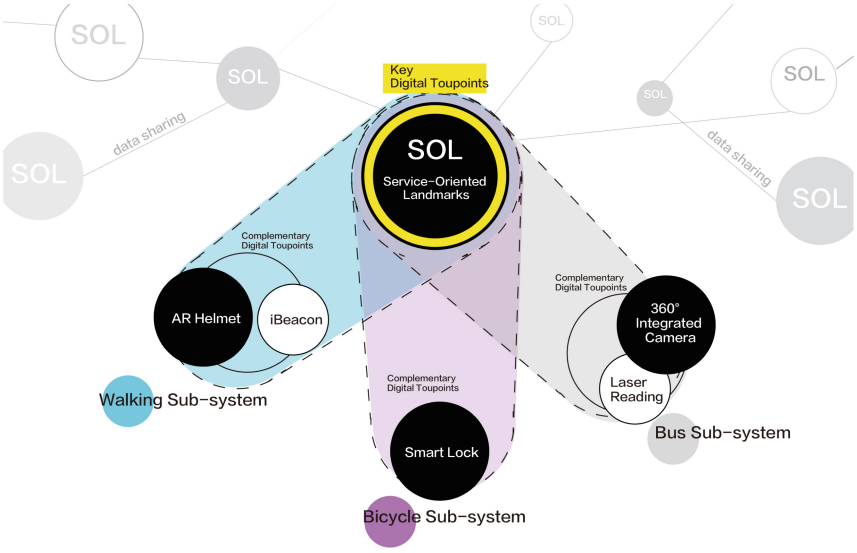


Fig. 6. Digital touchpoints architecture

5.2 Service-Oriented Landmarks

Service-oriented landmark (SOL), an interactive digital device, is identified as the key touchpoint in our service system. First, situated in primary walking path, SOLs possess recognizable appearance as the hallmark of service providers. People would intend to seek for service wherever they encounter a SOL in the campus. Figure 7 illustrates the overall SOL distribution map in Jiading campus. Second, SOL serves as an IoT-based device to collect service-defined data from the surrounding environment and simultaneously share to the cloud system for computational synthesis. Through the data sharing mechanism, our service system can get to know the real-time human activity and environmental factors around all the SOLs to provide sensible service. Third, we set out to design a tangible interface for SOL where through both screen-based and physical interaction, people can acquire and exchange information related to the slow traffic. Therefore, we identified the SOLs as the fundamental service providers in our slow traffic service system.

Fig. 7. Digital touchpoints distribution map in Jiading Campus

5.3 Digital Touchpoints in the Three Sub-systems

For each sub-system, we customized different types of digital touchpoints which are mainly digital or smart devices. The functionality, interface, interaction for each touchpoint would be introduced.

Touchpoints in Walking Sub-system

SOLs. SOLs in walking sub-system are used for customizing applicable walking routes such as the express route, campus touring route and personalized route. The interface would automatically present the navigation route according to people's selection. In

addition, through the interface showing all the SOL symbols and the environmental information nearby each SOL on the campus map, people can link the interested symbols together to generate the personalized route.

AR helmet. We set out to deploy the AR helmet touchpoints at the attractions and featured architectures for touring service, allowing visitors to obtain introduction about the history and story with an augmented reality experience.

iBeacon base station. We adhere the iBeacon base station to the surface of SOLs near the supermarkets and cafeteria to provide useful shopping and food information like discounts and the specials. The mobile app would alert an instant message when people are walking and approaching the iBeacon. This proactive campus information service establishes the intimate connection with the walking sub-system.

Touchpoints in Bicycle Sub-system

Smart Lock. The distinctive phenomenon in Jiading Campus was that there were many unclaimed bicycles parking in the streets, which occupied much space. So how to replace the existing idle resource into sharable ones, to create more values. The smart lock is the vital media displace idle bicycles into public sharing bikes, which function in GPS positioning, identity recognition and motor-control lock actuation. With the smart lock in the bicycle, people can use the school ID card to turn on the lock and afterwards lock it to finish the service. Meanwhile, the location information would transmit to the cloud system when it is not in use (Table 8).

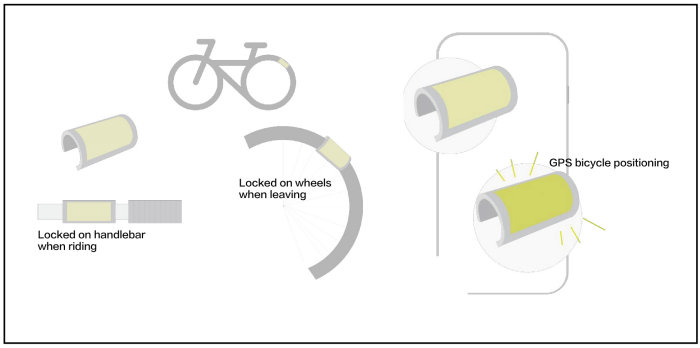


Fig. 8. Smart lock

SOLs. SOLs in bicycle sub-system serves as digital signage for navigating and gathering bicycles. People can manipulate the interface of any SOL to find out the location of free bicycles, which better improves the efficiency of finding bicycles.

Touchpoints in Bus Sub-system. There is just a small number of buses running in Jiading campus now, which hardly meet the needs of users. Therefore, when taking bus,

people always have to queue up to wait for some time, which results in a waste of time of passengers.

360° 4DOF Camera. To address the problem, we designed 360° 4DOF Camera integrated into the SOLs at the bus station. The camera can capture the real-time video stream for people to oversee the current human crowdedness. Additionally, the camera supports 4DOF rotation and motion to capture a panoramic view.

Laser Reading. Laser Reading is another proactive method to track whether there are passengers queuing up. If the bus station turns out to be crowded, the situation will be reflected on the SOLs Map.

SOLs. Through the interface, people can check the crowdedness of the bus stations to decide whether it is advisable to take the bus. Given that they can use multi-touch gestures like swiping left/right, zooming in/out to manipulate remote monitoring of the video view.

5.4 Technology Overview

Data Capture and Sharing. We proposed the cloud-based technology architecture of data capture and sharing. Figure 9 illustrates the overview. Distinctive data capture from IoT sensors, GPS sensors, cameras and laser readings are wirelessly transmitted to the cloud system. After the algorithmic computation, the semantic information is requested back to the SOLs for people to navigate and oversee.

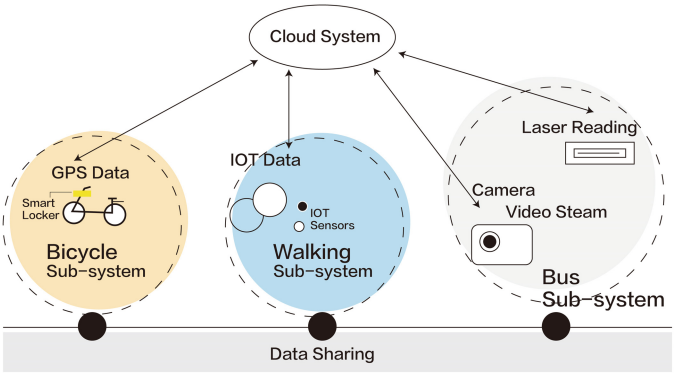


Fig. 9. Technology architecture

Technical Details

SOLs. An SOL touchpoint consists of touch-screen interface, android-based OS, IoT sensors, Wi-Fi module and microcontroller board (shown in Fig. 10.).

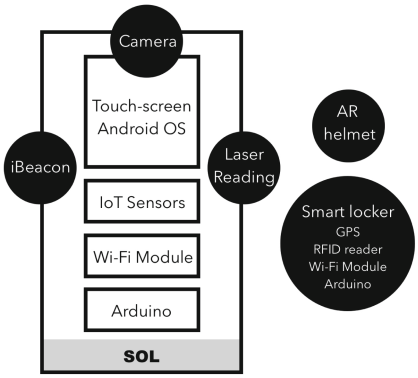


Fig. 10. Hardware overview

Smart Lock. The smart lock includes a GPS positioning sensor to track the position, a RFID reader for the identity recognition, a servo motor for the actuation and a built-in Wi-Fi microcontroller board for the computation and transmission.

6 Conclusion

Although regularly discussed and described in slow traffic, there is insufficient research from the perspective of service system design or into the nature of campus transportation. The emerging advanced technics give us more technical supports to provide better services to meet people's needs, while how to integrate these technical tools and innovate into digital touchpoints network remain to have many potentials to explore. There is therefore a need to support campus traffic mode innovation through service system approach to organize digital touchpoints.

This research, as a rough attempt, takes a tiny step towards improving campus slow traffic service using digital touchpoints as main methods. In this service project, different digital touchpoints were structured in a network based on the common service-oriented landmarks (SOL). The ambient data required by services was captured by SOL and shared among different digital touchpoints in a cloud-based technology architecture. How this kind of digital touchpoint network assist the whole service system to perform in other contexts, such as tourist attractions, shopping malls, deserve further research in the next step of work.

References

1. Balsas, C.J.L.: Sustainable transportation planning on college campuses. Transp. Policy **10**(1), 35–49 (2003)
2. Engwicht, D.: Reclaiming Our Cities & Towns[J]. Sources Toronto Canada (2009)
3. Slow traffic system. http://baike.baidu.com/link?url=Ik-7RMv7JqjS54plpPUJ3PVrWbUqC3c9SS8yiGca2bNviPFoGmHsy9eTetwoq-xljaFovyUZuDsp6cCyquF1_5C0b-khZm6qm8TOCB9X7Q1le55L55MowtsnLeTsvUVjHpdQcIwf6IsVTJKM6M4-bq

4. Spohrer, J., Giuiusa, A., Demirkan, H.: Service science: reframing progress with universities. Syst. Res. Behav Sci. **30**(5), 561–569 (2013)
5. Forester, J.: Bicycle Transportation, 2nd Edition. Environmental Politics & Policy, (1994)
6. Stickdorn M. This is Service Design Thinking. 2011
7. Sun X, Shi J. Environment-Specific Smart Service System Design// Cross-Cultural Design. Springer International Publishing, Cham (2016)
8. Service Design Tools. http://www.servicedesigntools.org/taxonomy/term/24
9. Clatworthy, S.: Service innovation through touch-points: development of an innovation toolkit for the first stages of new service development. J Exerc Sci Physiother **5**(2), 15–28 (2011)
10. Guseynova, N.: Emotions in design process. How to find an emotional touchpoint with the user. Servdes Org (2012)
11. Morrissey, B.: optimizing digital touchpoints. Success. Mark. **1**, 74–75 (2008)
12. Idoughi, D., Seffah, A., Kolski C.: Adding user experience into the interactive service design loop: a persona-based approach. Behav. Inf. Technol. **31**(3), 287–303 (2012)
13. Heapy, J., Parker, S.: The Journey to the Interface. London Demos (2006)

Cognitive Vehicle Design Guided by Human Factors and Supported by Bayesian Artificial Intelligence

Ata Khan(✉)

Carleton University, Ottawa, ON K1S5B6, Canada
ata_khan@carleton.ca

Abstract. Researchers and automotive industry experts are in favour of accelerating the development of "cognitive vehicle" features, which will integrate intelligent technology and human factors for providing non-distracting interface for safety, efficiency and environmental sustainability in driving. In addition, infotainment capability is a desirable feature of the vehicle of the future, provided that it does not add to driver distraction. Further, these features are expected to be a stepping-stone to fully autonomous driving. This paper describes advances in driver-vehicle interface and presents highlights of research in design. Specifically, the design features of the cognitive vehicle are presented and measures needed to address major issues are noted. The application of Bayesian Artificial Intelligence is described in the design of driving assistance system, and design considerations are advanced in order to overcome issues in in-vehicle telematics systems. Finally, conclusions are advanced based on coverage of research material in the paper.

Keywords: Autonomous vehicle · Human-factors · Cognitive vehicle · Advanced driving assistance system · Automation in driving

1 Design Features of the Cognitive Vehicle

Since the emergence of the *intelligent vehicle and highway system* (IVHS) field a few decades ago, intelligent vehicles have taken several giant leaps and the car of the future is expected to be "smarter" and "autonomous" with safeguards regarding shared authority with human driver. Researchers and automotive industry experts are leaning towards a vehicle design that exhibits "cognitive" features. This can become a reality only if intelligent technology and human factors are integrated for providing non-distracting interface for safety, efficiency and environmental sustainability in driving. In addition, infotainment capability is a desirable feature of the vehicle of the future, if it does not add to driver distraction.

Early R&D in the cognitive car was described by Heide and Henning [1], Stiller et al. [2], and Hoch et al. [3]. The National Research Council of Canada (NRC) took an interest in the subject by developing a Cognitive Vehicle Technology Platform that was intended to avail opportunities in the rapidly growing field of automotive information and communication technologies (ICT). The NRC's R&D in this area was a response to

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_36

both proliferation of on-board sensors and electronic controls for enhancing safety and performance, and a rising demand for integrated and wirelessly connected communications and infotainment devices [4].

Cognitive connected and autonomous vehicle research aims at integrating intelligent technology and human factors. Although R&D efforts underway in many countries are timely, there is a need to go a step beyond the existing programs in a number of key high priority areas. Figure 1 presents R&D themes, classified according to function. These themes cover capabilities with recognized market value and are fully integrated. For various themes, the components are application-specific (e.g. open architecture for multi-functional advanced driving assistance system design).

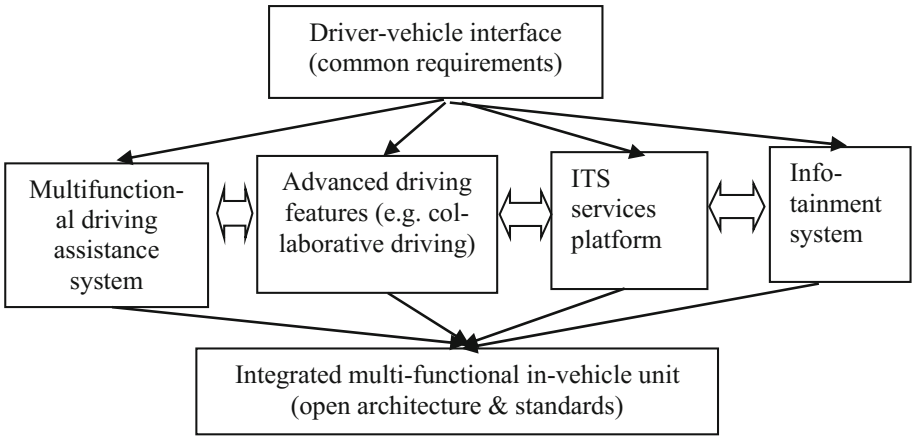


Fig. 1. Cognitive vehicle R&D themes

A key area of research is the driver-vehicle interface that comes into play in all functions of the vehicle of the future. In this case, common requirements are to be defined in support of the following components:

- Driving assistance system
- Advanced driving features such as collaborative driving
- Intelligent transportation system (ITS) services platform
- Infotainment system

In addition to the above functional components of the advanced technology vehicle, an in-vehicle unit is required with the capability to integrate multi-functions noted above. As in the case of the design of the individual components, this in-vehicle unit should exhibit open architecture and standards.

The rationale for developing cognitive vehicle features is provided below.

- The full potential of smarter transportation valued by drivers, especially younger generations of drivers, cannot be achieved without the cognitive vehicle.
- In the era of policy trends encouraging the convergence of transportation systems and services, the cognitive vehicle is central to future intelligent transportation.
- Sensors and communication devices account for substantial investments in transportation systems. Without the cognitive vehicle, the instrumentation of transportation will not be effective.

Based on industry analyses and research studies, cognitive vehicle features of high market potential are defined this paper. As a step in this direction, the following observations on improving vehicle design are noteworthy.

- For improving safety, mistakes made by distracted drivers should be reduced.
- Selected active safety features are likely to gain favour with the users, provided that their designs are improved substantially [5].
- Advanced driver assistance (driver support) systems that take into account "driver intent" are necessary.
- Ability to support non-distracted and non-aggressive driving in longitudinal and transverse guidance of the vehicle – if activated by the driver for reasons of comfort, convenience and safety in driving, will be a highly valuable design contribution.
- Natural non-distracting driver-vehicle interface that reduces driver stress and workload is essential.
- Ability to connect with other vehicles, infrastructure and devices is essential for future generation of vehicles [6].

2 Issues in Highly Automated Driving and Design Solutions

There are major issues in driving that a cognitive vehicle design should address. These are driver distraction, driver workload, driver stress, and inconsistent and confusing design features. In addition to common general distractions (e.g. passengers, food), there are concerns that in-vehicle technologies and applications may add to common sources of distractions. In-vehicle access to intelligent transportation system services and infotainment may add to distraction unless their designs are non-distracting. Even a driving assistance system and cooperative driving technologies may have a distracting side effect in the form of too many warnings and false alarms.

Although technologies play an important role in developing a cognitive vehicle, there is a need to go beyond strictly technical features by including well-researched means for vehicle-human interaction. Table 1 defines features that should be built into the design of the cognitive vehicle in order to enable challenging applications.

Table 1. Cognitive vehicle features and applications

Features	Applications
Ability to provide non-distracting user interface	For all driving, ITS service, and infotainment tasks, a non-distracting interface will be required
Awareness of position and surroundings	Cooperative collision avoidance, collaborative driving, ITS services (e.g. navigation)
Ability to gather and send out data	Safety, traffic control and management, ITS services
Ability to process data	Implementation of algorithms for safety, efficiency, and eco-driving
Ability to cooperate/collaborate	Cooperative collision avoidance, collaborative driving
Communication for active safety	Informing driver and neighbouring vehicles about potential collision hazards (e.g., informing left-turning vehicles about pedestrian on a crosswalk)
In the event of a crash, capability to send and receive information	Emergency/safety applications
Informs the driver about situations (warnings, advice); defines how to deal with driver distraction and aggressiveness	Safety applications (e.g., informing driver about object or stalled vehicle and aiding the driver in making an early lane change, collision warning)
Diagnostic capability	Enhancement of efficiency in the use of vehicle and convenience of the driver
Infotainment capability (while managing driver distraction)	Driver and passenger convenience

Taking each component of the cognitive vehicle design, illustrated in Fig. 1, the following figures highlight the problems as well as measures required to solve these through vehicle design. Figure 2 notes the requirements of driver-vehicle interface applicable to all components of vehicle design and ultimately operation in the real world.

In developing driver-vehicle interface designs for various applications implied by R&D themes illustrated in Fig. 1, details of the following common guidelines will be needed:

- Placing minimal demand on driver workload
- Reducing the frequency of driver multitasking
- Reducing the complexity of distracting tasks thus lowering demand on driver attention
- Reducing driver workload by managing available options
- Accounting for driver distraction with adaptive algorithms
- Monitoring driver attention status using in-vehicle technologies for alerting drivers.

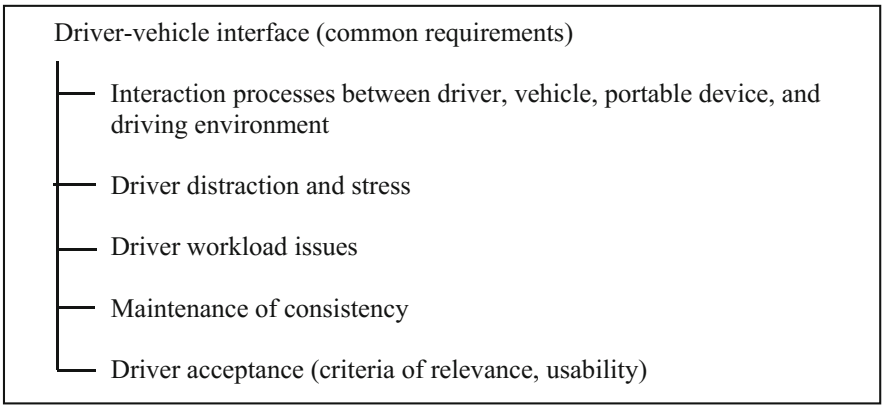


Fig. 2. Driver-vehicle interface

A major challenge in developing the cognitive connected vehicle is to add an advanced driving assistance system in order to meet the objectives of safety, convenience, economic efficiency, and environmental sustainability in driving. The automotive industry is very keen on innovating in this part of the cognitive vehicle design. See Fig. 3 for design features.

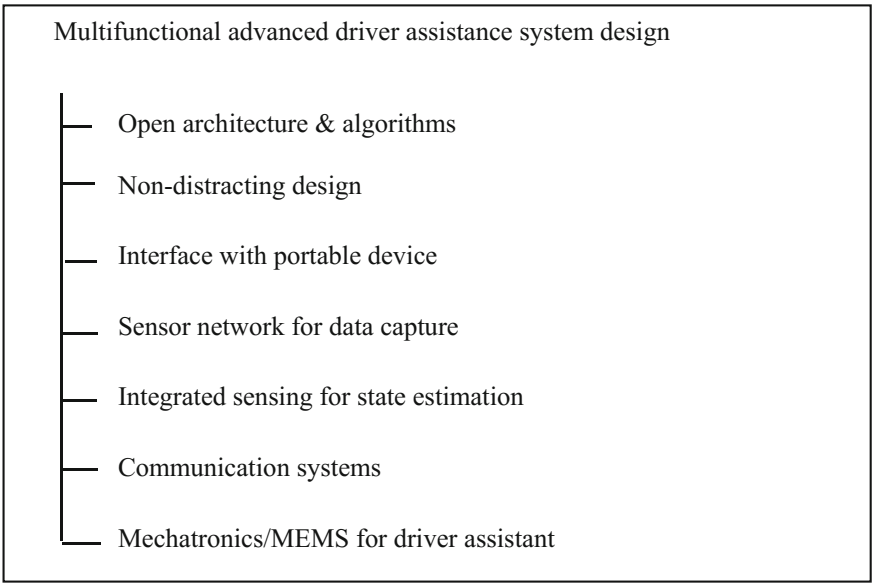


Fig. 3. Driver assistance system design

In order to put technology to work for supporting new and advanced driving features, research is underway in collaborative driving and other applications of vehicle-to-vehicle and vehicle-infrastructure communications. In addition to maximizing the throughput from highways, environmental benefits are achievable. In order to safeguard safety and enhance user acceptance, technology and human factors have to be integrated in designs (Fig. 4).

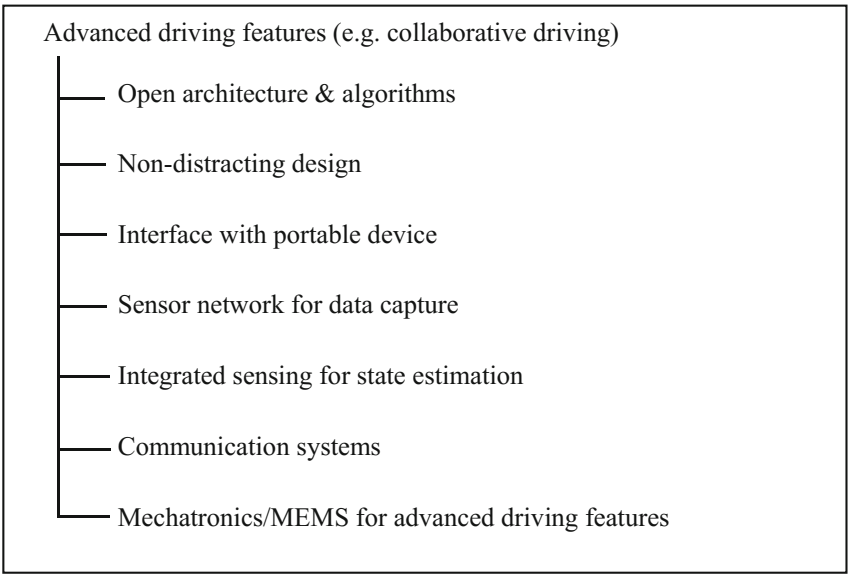


Fig. 4. Design for advanced driving features

For enhancing the value of a future vehicle to the owner, in addition to other features, it is necessary to formally incorporate the special requirements of intelligent transportation system (ITS) and services in vehicle design. The market potential for cognitive vehicles with integrated non-distracting and intelligent transportation features is very large. Examples of ITS services include parking information, route guidance and other traveller information, and insurance and road user charges that are expected to become on-line services (Fig. 5).

The infotainment system is intended to receive a wide variety of information and data of interest to the human driver and other occupants of the vehicle. In case of autonomous driving, such information can be used as well for non-driving functions. In addition to receiving information on-line, the internet capability will allow connectivity with out-of-vehicle sites. The buyers of advanced technology vehicles are likely to be

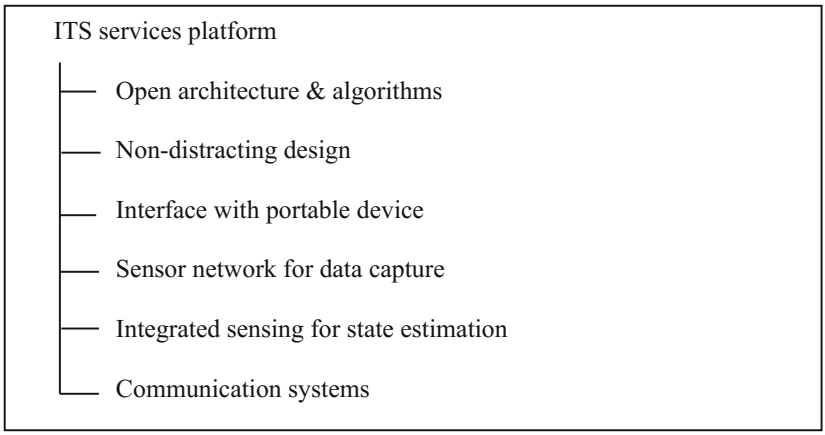


Fig. 5. Design for ITS services

keen on this convenience feature, which could also have the added entertainment value. On the negative side, due to the potential for prolonged communications such as requesting downloads or viewing received information, driver distraction becomes an issue. Therefore, the design of the infotainment system requires much attention regarding human factors (Fig. 6).

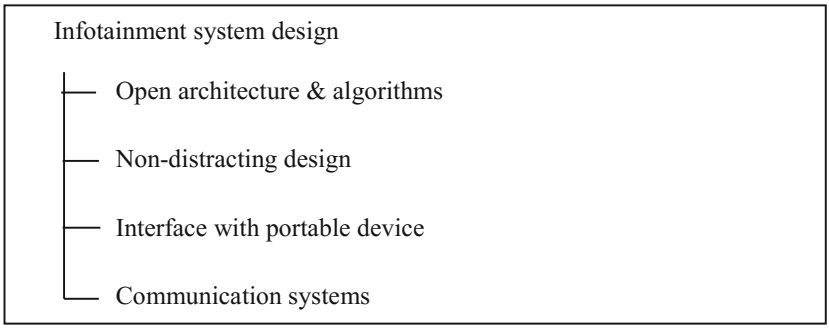


Fig. 6. Design for infotainment system

3 Application of Bayesian Artificial Intelligence

Korb and Nicholson have defined artificial intelligent (AI) as the "intelligence developed by humans, implemented as an artefact" [7]. The Bayesian methodology is best suited for system design and decision analysis when uncertain "states of nature" are encountered, but knowledge of uncertain factors such as driving states, driver distraction and driver intention can be refined with additional information acquisition.

Bayesian AI can be of much value in cognitive vehicle design, especially in the design of the advanced driving assistance system. For an introduction to the Bayesian AI application to driver assistance design, please refer to Khan [8, 9]. Applications in the design of driving assistance system can manage driver workload, warn drivers of risky situations, and monitor driver performance. Illustration of countermeasures directed to the driver shows the potential to overcome the distracted driving problem [10].

For design purposes, the Bayesian artificial intelligence integrates two cognitive features of the advanced driving assistance system. The first one models a desired human action (e.g., non-distracted non-aggressive driving). The second is to model the rational view of what is "optimal" and apply it as a decision criterion.

Khan has suggested a three-step approach. The use of algorithms for Bayesian analysis of driving missions is the first step. Next, is the computation of expected gains/utilities. Finally, the optimal course of action is identified on the basis of maximum gain/utility criterion [11].

Application of intelligent technology and human factors in the driving assistance system design will provide seamless transition between human control and automation. This advance in design should overcome driver dissatisfaction with false alarms, which arise due to lack of formal treatment of distracted driving and driver intent in existing driver assistant designs. The advanced design reported by Khan has the capability to avoid rear as well as lateral crashes. The high-level architecture of the advanced driver assistance system is shown in Fig. 7.

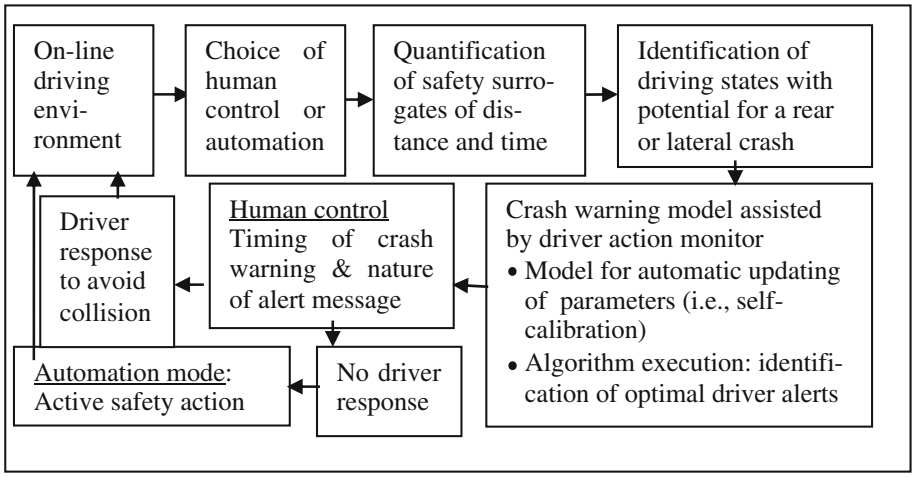


Fig. 7. High-level architecture of driving assistance system's safety function [11].

4 In-Vehicle Telematics Issues

In-vehicle telematics can distract human driver and encourage multi-tasking. To overcome such issues, the advanced driving assistance system provides a solution. To illustrate how driving can be affected by the use of in-vehicle telematics, driving simulator data is presented here.

Driving simulator experiments were carried out with the participation of young drivers who were given distracting tasks to do while driving. The simulator outputs contain useful information on the driving behaviour of participants. In non-distracted driving, the choice of speed, acceleration/deceleration rates, headway to the leading vehicle, and gap acceptance when changing lanes provide information on the lack of or presence of aggressiveness in driving.

During a driving mission, if the driver becomes distracted and is not alert enough to perceive a hazard, this results in insufficient distance for stopping or to make an evasive manoeuver. Excessive time could be required by a distracted driver to perceive the hazard and then excessive deceleration rate (close to 1 g) will be required to avoid a collision.

Such sudden emergency braking sends shock waves in the traffic stream and is known to be a safety hazard. Also, a distracted driver may experience lane migration and violate safety distance in the lateral direction.

Figures 8, 9, 10, and 11 illustrate the driving record of a young driver who was using a hands-free phone and encountered a hazard that required action. In this driving trajectory, rapid changes between deceleration and acceleration can be seen. In addition, the acceleration pattern is unusual. Although the perception-reaction time of 1.17 s was acceptable, other attributes of driving could be assisted with the cognitive vehicle's advanced driving assistance system.

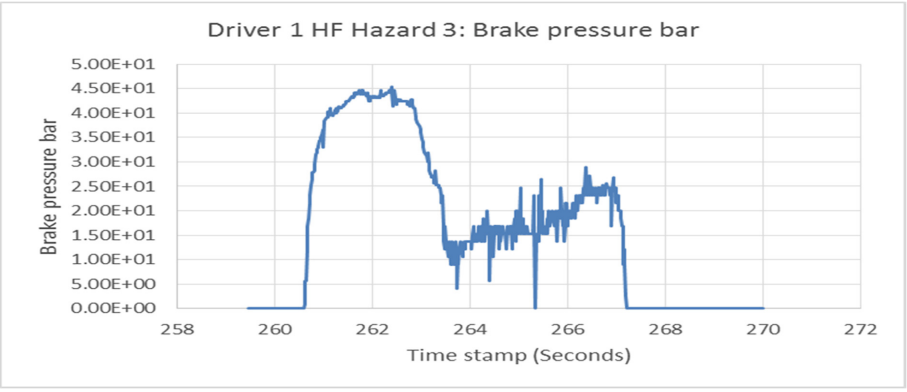


Fig. 8. Brake action while using hands-free communication device

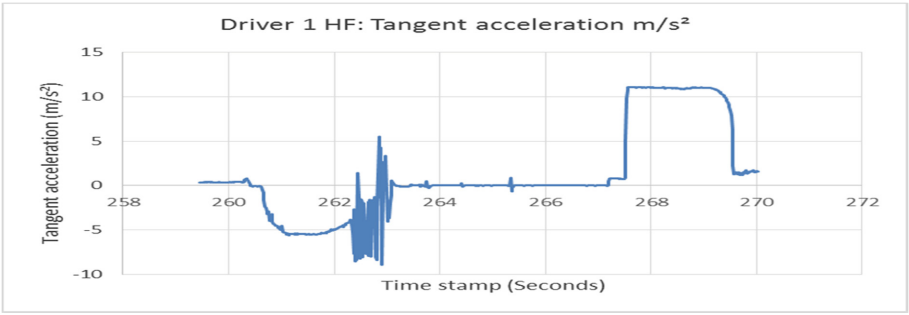


Fig. 9. High level deceleration and acceleration pattern

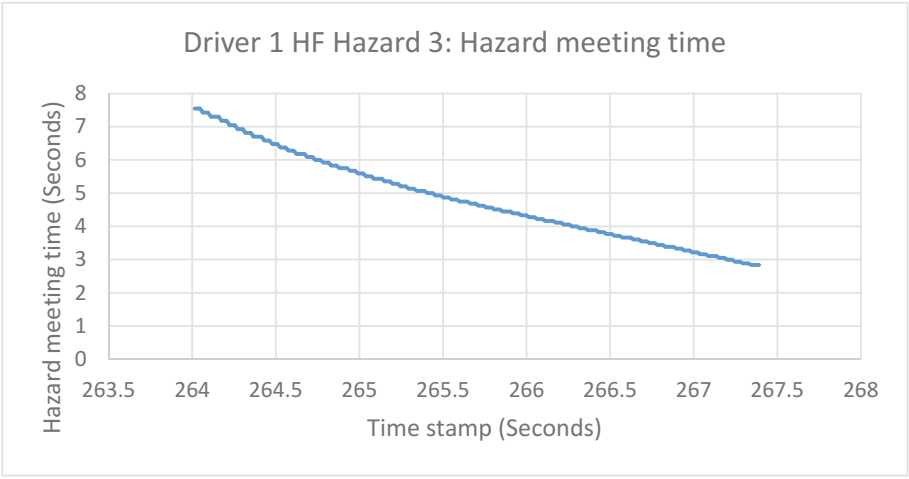


Fig. 10. Hazard meeting time

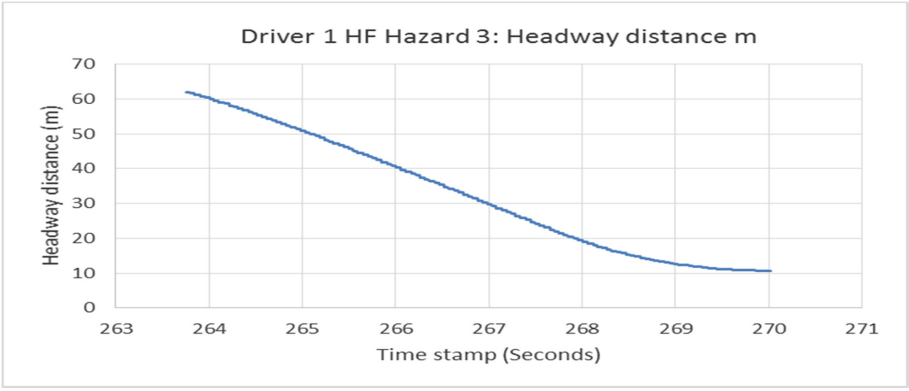


Fig. 11. Headway distance to leading vehicle

5 Conclusions

(1) The design of driver-vehicle interface should enable the provision of information in an effective manner and at the same time should eliminate or at least minimize distractions that contribute to elevated probability of crashes.
(2) Through new knowledge such as carefully designed application of Bayesian AI, in-vehicle systems will evolve that will be effective in serving their function, including necessary and optimized alerts.

Acknowledgements. This paper is based on research sponsored by the Natural Sciences and Engineering Research Counsel of Canada (NSERC) and the Ministry of Transportation of Ontario (MTO). The views expressed are those of the author.

References

1. Heide, A., Henning, K.: The "cognitive car": a roadmap for research issues in the automotive sector. Annu. Rev. Control **30**, 197–203 (2006)
2. Stiller, C. Farber, G., Kammel, S.: Cooperative cognitive automobiles. In: Proceedings of the 2007 IEEE Intelligent Vehicles Symposium, pp. 215–220, Istanbul, Turkey, 13–15 June 2007. WeC1.1
3. Hoch, S., Schweigert, M., Althoff, F.: The BMW SURF project: A contribution to the research on cognitive vehicles. In: Proceedings of the 2007 IEEE Intelligent Vehicle Symposium, pp. 692–697, Istanbul, Turkey, 13–15 June 2007. ThB1.26
4. National Research Council (NRC): Cognitive vehicles, key industrial sectors (2011). http://www.nrc-cnr.gc.ca/eng/sectors/automotive/cognitive-vehicles.html
5. Traffic Technology Today: News: autonomous vehicle software developer secures funding for 'human-like' driving, 5 February 2016
6. European Commission: EU accelerates connected vehicles. ITS International, 25 January 2016. And C-ITS Platform, Final Report, January 2016. Also see European Commission, 2015. Connected Cars. Business Innovations, Internet of Things, Case Study 43, February 2015
7. Korb, K.B., Nicholson, A.E.: Bayesian artificial intelligence. Chapman & Hall/CRC, London (2004)
8. Khan, A.M.: Cognitive connected vehicle information system design requirement for safety: role of Bayesian artificial intelligence. Syst. Cybern. Inform. **11**(2), 54–59 (2013)
9. Khan, A.M.: Bayesian-Monte Carlo Model for collision avoidance system design of cognitive vehicle. Int. J. Intell. Transp. Syst. Res. **11**(1), 23–33 (2013). Springer
10. Khan, A.M.: Modelling human factors for advanced driving assistance system design. In: AHFE International Conference, Orlando, FL (2016)
11. Khan, A.M.: Design of real-time transition from driving assistance to automation function: Bayesian artificial intelligence approach. In: Fischer-Wolfarth, J., Meyer, G. (eds.) Advanced Microsystems for Automotive Applications. Lecture Notes in Mobility, pp. 39–48. Springer, Berlin (2014)

Assessing In-Vehicle Secondary Tasks with the NHTSA Visual-Manual Guidelines Occlusion Method

Mikael Ljung Aust(✉), Robert Broström, and Sofia Lindvall

Volvo Car Corporation, Volvo Jacobs vag, 405 41 Gothenburg, Sweden
{Mikael.Ljung.Aust,Robert.Brostrom,Sofia.Lindvall}@volvocars.com

Abstract. The NHTSA Visual-Manual Distraction Guidelines propose two methods for evaluating which in-vehicle tasks should be allowed while driving. Of these, previous studies have shown that the Eye Glance measurement using Driving Simulator method (EGDS) has low reliability. This study therefore explored if the other proposed method, Occlusion Testing, provides a better alternative. Occlusion testing did indeed show more robust performance than EGDS. It also was more permissive, i.e., tasks that failed with EGDS passed with Occlusion. Furthermore, results indicate that for many tasks, Occlusion testing can be replaced with a simpler step count method. Two drawbacks were also identified. First, tasks requiring longer system search times are systematically allowed fewer interaction steps for completion. Second, test outcomes are highly dependent on participant instructions. Participants who perceive the evaluation as an individual test use more effort and develop optimization strategies, which means complex tasks that normally would fail will pass.

Keywords: Human factors · In-vehicle systems · Visual distraction · Secondary tasks · Performance assessment · Evaluation methods

1 Introduction

Naturalistic driving studies indicate that looking away from the forward roadway for more than a certain period of time is a key crash contributing factor [1–4]. This fact, combined with the societal transition towards ubiquitous use of information technology, has put driver distraction on top of the traffic safety agenda.

From a vehicle design point of view, many potential sources of distraction such as eating, drinking and interacting with passengers lie outside the vehicle manufacturers' influence. However, they do control the design of all in-vehicle systems that involve visual-manual interaction while driving.

To guide the design of these system, several design guidelines have been published. These include the European Statement of Principles [5], the JAMA guidelines [6] and the Alliance of Automobile Manufacturers guidelines [7]. These guidelines are often difficult to apply, because while they usually set performance thresholds that in-vehicle tasks need to meet to be allowed while driving, they rarely propose explicit methods for testing task compliance with said thresholds. For example, the AAM guidelines [7] suggest that each visual-manual task should require less than 2 s in mean glance duration

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_37

and maximum 20 s total eyes off-road time, but do not specify under which conditions this should be tested.

The latest addition to such guidelines, the Visual-Manual NHTSA Driver Distraction Guidelines for in-vehicle Electronic Devices [8], departs from the above tradition, because in addition to specifying performance criteria against which in-vehicle tasks should be evaluated, they also specify two test methods for assessing criteria compliance for in-vehicle tasks. These are Eye Glance measurement using Driving Simulator testing (EGDS) and Occlusion testing.

Of these two, previous studies have shown that EGDS as currently formulated has low reliability. EGDS test outcomes depend much more on individual glance behaviors among the randomly selected test participants than on how the in-vehicle task under evaluation is designed [9–14]. This makes EGDS very difficult to apply both in formative and summative evaluation. For formative evaluation, it lacks relevance because the result of one test can not be used to predict the outcome of the next, even if the tested task only changed marginally. In summative evaluation, EGDS outcomes will be near stochastic, since individual traits among the randomly selected participants have a dominating influence on test outcome [9–14].

These previous studies have lead to the insight that EGDS, while more attractive due to perceived ecological validity, probably should be discarded. The current study therefore focuses on the second proposed method, i.e., Occlusion testing [8]. Occlusion testing uses a visual occlusion technique where the test participant is wearing eye glasses with inbuilt shutters. The shutters open and close in 1.5 s segments (1.5 s open, 1.5 s closed, 1.5 s open, 1.5 s closed, and so on). The shutter closed times are meant to serve as a surrogate for the time drivers would look at the forward roadway when timesharing between driving and performing a secondary task while driving a real vehicle.

The pass/fail criteria for the Occlusion method are quite simple. If a driver can complete a *testable task* [15] within 8 shutter open segments (i.e., a cumulative shutter open time of 8×1.5 s = 12.0 s), it should be allowed while driving, and otherwise not.

A key difference between Occlusion and EGDS testing is that the duration of the off-road glances used to carry out the secondary task are self paced in EGDS but fixed in Occlusion testing. It follows that the large between-driver off-road glance duration variability found in EGDS testing [9, 13] would be expected to shrink, since that variability extended well beyond the Occlusion method's limit of maximum 1.5 s per glance. For Occlusion testing, it can thus be hypothesized that performance should be more similar between drivers, and outcomes less susceptible to between-driver variability, albeit perhaps at the price of forcing part of the test population to complete the secondary task in a non-familiar manner.

The aim of the current study was to test if test outcomes with Occlusion testing are robust (e.g., compared to using EGDS). In particular, given that Occlusion in the NHTSA configuration affords eight shutter open segments, it is not much of a stretch to view this as an elaborate step count method. If one could determine a performance threshold in terms of the likelihood that a task will comply with the pass criteria as a function of the number of shutter open segments, formative Occlusion testing could be replaced with a simple step count method up to that reliability threshold where compliance starts to deteriorate.

In addition to the main aim, two factors deemed to have large potential to influence test outcomes were studied. The first factor was character input method. To assess this, tasks where character input was required altered between keyboard and scribble inputs, to see if this would systematically alter test outcomes. The second factor relates to the fact that Occlusion based testing is not a dual-task paradigm like EGDS, but rather a single task paradigm. Under single task paradigms, test persons are more likely to view the task at hand as an individual performance test, especially if they become aware that task performance is being assessed towards a pass/fail criteria. This might trigger general alertness and/or optimizing behaviours which would lead to participants performing tasks faster/better than they world during normal, everyday driving. As can be imagined, this would have a drastic effect on test outcomes. The study therefore also tested whether optimizing behaviours could be induced, and if so, what the effects would be.

2 Methodology

Three Occlusion tests were performed, each covering a range of different in-vehicle tasks. The first test involved 17 participants, the second involved 26 participants and the third 24 participants, making for a total of 67 participants.

Participant sampling loosely conformed to the NHTSA guideline recommendations [8] regarding general criteria, participant impartiality, age mix and gender balance. However, strict adherence to the recommendations was not considered necessary, as the aim of the study was to explore strengths and weaknesses of the Occlusion method rather than certify any particular set of in-vehicle tasks.

2.1 Equipment

Test were conducted using the Plato System Occlusion glasses from Translucent Technologies [16]. A touch-screen tablet containing a prototype infotainment system, designed to replicate a typical in-vehicle center stack interface, was mounted on a panel at such height, position and angle that it was in the same position relative to the test participant as it would have been if the test participant was interacting with it when mounted in a real vehicle.

2.2 Tasks Performed

The first group tested 14 tasks, the second group tested 12 tasks and the third group 8 tasks. All tasks, while spread over typical infotainment domains such a navigation, media playing and radio, were selected to achieve a specific number of interaction steps for completion.

For the first group, the step range was fairly evenly distributed from 4 to 13 steps, in order to cast a net wide enough to capture any potential compliance deterioration threshold. As results for the first group indicated that such a compliance threshold might lie in the 7–9 step range, the tasks for the second group were mainly selected in this step count range.

Table 1. Descriptions and step counts for the tasks performed by each group in the study

Group	Number of steps	Task	Input method	
			Keyboard	Scribble
1	3>	Change FM radio frequency band		
	5	Select a favourite destination and start navigation		
	6	Search for a song and play it		x
	6	Search for a song and play it	x	
	7	Search for a contact and start call	x	
	7	Search for a song and play it		x
	8	Set street address and start navigation	x	
	8	Find POI along route and start navigation		
	8	Set street address		x
	9	Send text message		x
	9	Search for a contact and start call		x
	10	Send text message		x
	10	Search for a contact and start call	x	
	13	Call your own number	x	
2	4	Choose radio channel from a list of favourites		
	5	Set a favourite destination and start navigation		
	6	Search for a song and play it		x
	7	Set street address and start navigation		x
	7	Set street address and start navigation	x	
	7	Search for a contact and start call		x
	8	Set street address and start navigation	x	
	8	Set street address and start navigation		x
	8	Find POI along route and start navigation		
	9	Set street address and start navigation		x
	9	Set street address and start navigation	x	
	9	Find POI along route and start navigation		
3	2	Activate Bluetooth		
	2	Change temperature		
	2	Set seat heat		
	3	Call someone from list of recent calls		
	4	Select and play a song from USB memory		
	5	Set favourite destination and start navigation		
	13	Call your own number	x	
	14	Set new destination	x	

In the third group, where the aim was to study effects of biasing participants into a "competitive" mode, a set of 8 tasks were selected. 6 had a low step count and were hence expected to meet threshold limits. One was the phone dialing task, in order to

replicate the outcome of this task in group 1, and one was a 14 step navigation task, i.e., a task so difficult that is should be hard to complete with under the compliance thresholds (see Table 1 below). Character input methods was systematically varied between similar tasks.

2.3 Procedure

The test procedure, including test participant training, followed the Occlusion testing protocol [8]. Upon arriving to the test facilities, participants were given a brief description on the test procedure. They were then asked to sit down in front of the prototype HMI, and guided to adjust their position so they could reach the tablet in the same way as if they were seated in a real car.

For each task, the test participant was first instructed on how to carry out the task. Next, the participant was allowed to practice the task until performance was error free. After this, the Occlusion glasses were put on and task trial began. Every task started from a home screen view, and ended when it had been correctly executed.

In study one and two, no special instructions regarding how to perform the tasks were given. After practice, participants simply asked to complete the task they just had practiced while wearing the occlusion googles, and all timing devices were kept out of sight of the test participant.

For study three where the objective was to induce a bias towards performance optimization, whenever a participant was failing on either of the two longer tasks in the task set, the test leader was instructed to tell the participant that s/he had failed to complete the task within the expected time. The test leader then would place a timer in clearly visible position to the test participant and ask if s/he would like to practice some more and try again. The test leader would also state that most participants had not experienced any difficulties in meeting the performance criteria.

2.4 Data Analysis

Based on the Total Shutter Open time for each test person, a pass percentage (i.e., the proportion of test participants meeting the NHTSA Occlusion testing criteria) was calculated for each tested task.

3 Results

Results show that all tasks which required 7 steps or less to complete passed in both groups. For those requiring 8 steps, 5 passed and 3 failed. For those requiring 9 steps, 1 passed and 3 failed. The two 10 step tasks both failed. Notably, the 13 step task, i.e., dialing a well known phone number, passed. In other words, except for number dialing, performance on tasks that took more than 7 interaction steps to complete declined with increasing step count.

There were no systematic effects of character input method; i.e., the different input methods did not have any systematic effect on test outcomes for the same tasks and groups (Table 2).

Table 2. Results for occlusion test in groups 1 and 2, sorted by step count

Group	Number of steps	Task	% of TP meeting criteria (%)	Keyboard	Scribble
1	3>	Radio	93.75		
2	4	Radio	100		
1	5	NAVI	100		
2	5	NAVI	100		
1	6	Media	100		X
1	6	Media	100	X	
2	6	Media	100		X
2	7	NAVI	100		X
2	7	NAVI	100	X	
1	7	Media	100		X
2	7	Contact	100		X
2	8	NAVI	96.15	X	
2	8	NAVI	100		X
1	8	NAVI	**75**		
2	8	NAVI	96.15		
1	8	NAVI	**81.25**	X	
1	8	NAVI	**75**		X
1	8	Contact	93.75	X	
2	8	Contact	100		X
2	9	NAVI	**73.08**		X
1	9	NAVI	92.31	X	
1	9	Message	**81.25**		X
2	9	NAVI	**61.54**		
1	10	Message	**37.50**		X
1	10	Contact	**81.25**	X	
1	13	Contact	100		

In the third study, where a bias toward optimal performance was induced, again the 13 step dialling task passed. Moreover, while all but one person failed on the 14 step navigation task in their first attempts, all participants passed after being coached by the test leader toward a more competitive behaviour (Table 3).

Table 3. test outcomes in group 3, before and after being coached toward a competitive mind set

Group	Number of steps	Task	% of TP meeting criteria	
			Before introducing bias (%)	After introducing bias (%)
3	2	Vehicle function	100	100
	2	Vehicle function	100	100
	2	Vehicle function	100	100
	3	Phone	100	100
	4	Media	100	100
	5	Navi	100	100
	13	Phone	100	100
	14	Navi	**4.16**	100

4 Discussion

The aim of the current study was to assess the robustness and performance of the Occlusion testing method for in-vehicle task assessment [8]. As the task performance above illustrates, performance was a lot more similar between drivers with Occlusion testing, compared to previous studies where EGDS was used [9–14]. Occlusion testing thus seems a lot less susceptible to between-driver glance variability, compared to EGDS based testing.

Furthermore, test outcomes were very robust. Results showed identical compliance performance for all tasks with a step count up to 7 interaction steps, after which an increase in step count results in a graceful decline in performance. Moreover, compliance up to 7 interaction steps was 100%; no participant failed on any of these tasks. This stands in stark contrast to EGDS based testing [12, 13] where step count only was a weak indicator of whether a task would pass or fail. Furthermore, Occlusion testing seems more permissive than EGDS testing. Here, all tasks at 7 interaction steps or less passed, whereas with EGDS, several tasks with lower step count than 7 failed [12, 13].

The highly robust outcome for tasks with 7 interaction steps or less leads to the tentative conclusion that under the Occlusion testing paradigm, tasks that fall within this range actually do not need to be formally evaluated, since outcomes seems highly predictable up to this threshold. However, for tasks with 8 steps or more, or if interaction design failures has made a task particularly difficult to navigate, empirical testing is necessary to understand if the task will meet the NHTSA performance criteria.

An important question that remains open and requires further study is how to apply a principle like a 7 interaction step threshold in practice when designing infotainment tasks. Recent work in the Swedish project SICS [17] shows that it is possible to develop simple tools that can help interaction designers understand if a task will meet the Occlusion criteria. In this project, an Occlusion Service Application was implemented, which when active blocks the infotainment screen with the same open-closed timing as the occlusion goggles would block the eyes. However, how to integrate step count thresholds

into function specification work is still not very intuitive, and will probably require additional protocols to be successful.

A notable exception to the 7 step threshold was dialing a known phone number. However, the analysis of why participants were successful at this task was quite straightforward. Participants simply did not dial a known number digit by digit; rather they dialed in chunks of 3–4 numbers per shutter open instance. Because of this strategy, they could complete what actually is a 13 step task within the Occlusion criteria. This insight led to the formulation of an exception rule. If a task comprises naturally chunkable step segments, those can be expected to performed in a single shutter open instance, and should count as a single step when one does the step count analysis. Naturally, this exception rule should only be considered a first approximation. While other manual-visual interactions than number dialing also might be natural for participants to perform in a single shutter open instance, a more sophisticated theoretical framework that predicts which these would be and when they would occur needs to be developed.

Now, while the 7 step threshold seems a very useful guideline to go by, especially in formative evaluations, the performance in the third group of participants illustrated quite clearly that any task assessment must avoid triggering optimizing response strategies in the participants. When people are coached to do their best and allowed to practice until they have optimized their strategies, a 14 interaction step task is obviously possible to complete in 8 shutter openings.

Since Occlusion outcomes already are quite "generous", in the sense that many tasks which failed under EGDS testing [12, 13] pass with Occlusion testing, the implications are important and a key factor for any practical setup for in-vehicle task assessment of compliance with NHTSA guidelines. Means to monitor and control for this factor probably need to be implemented as part of the NHTSA guideline test protocols.

Another important finding in the current study was when the prototype system search engine failed to deliver a search result within 1.5–2 s, participants had nothing to "work with" when the Occlusion shutters opened the next time, and instead had to wait until a search results was present at shutter opening before they could continue with the task. Since the Occlusion glasses keep ticking regardless of system search times, it follows that any task which involves longer search times will be punished, in the sense that they systematically will be allowed fewer interaction steps for completion to meet the pass/fail criteria.

This can be viewed as an artefact of the method and needs to be addressed. Just like for other systems where searches may take longer time, infotainment developers will include standard progress indicators which indicate that a search is ongoing. In this situation, it becomes unreasonable to assume that a real driver would keep looking at the in-vehicle screen for as long as the search continues. Hence, Occlusion testing either needs to be complemented with instructions for when to pause during long system searches, or whether shutter opening times where no interaction takes place should be discarded from analysis.

A limitation of the current study is that while comprehensive, it still only involved 67 test participants. Further studies that can corroborate these findings are thus needed. Also, since participant selection criteria did not strictly follow NHTSA guideline specifications, further testing with more controlled sampling should be performed.

References

1. Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., Perez, M.A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z.R., Jermeland, J., Knipling, R.R.: The 100-Car Naturalistic Driving Study, Phase II – Results of the 100-Car Field Experiment. National Highway Traffic Safety Administration, Report no: DOT HS 810 593 (2006)
2. Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J., Ramsey, D.J.: The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study data. Virginia Tech Transportation Institute, Blacksburg (2006)
3. Victor, T., Dozza, M.: Timing matters: visual behavior and crash risk in the 100-Car online data. Presentation at 2nd International Conference on Driver Distraction and Inattention, Gothenburg (2011). http://www.chalmers.se/safer/ddi2011-en/. Accessed 1 May 2012
4. Liang, Y., Lee, J.D., Yekhshatyan, L.: How dangerous is looking away from the road? Algorithms predict crash risk from glance patterns in naturalistic driving. Hum. Factors **54**(6), 1104–1116 (2012)
5. European Commission: Commission Recommendation of 26 May 2008 on Safe and Efficient In-Vehicle Information and Communication Systems; Update of the European Statement of Principles on Human-Machine Interfaces. European Commission (2008)
6. JAMA: Guideline for In-Vehicle Display Systems, Version 3.0. Japan Automobile Manufacturers Association, Tokyo (2004)
7. Alliance of Automobile Manufacturers: Statement of Principles, Criteria and Verification Procedures on Driver-Interactions with Advanced In-Vehicle Information and Communication Systems. Technical report, Driver Focus-Telematics Working Group, Washington (2006)
8. National Highway Traffic Safety Administration: Visual-manual NHTSA Driver Distraction Guidelines for in-Vehicle Electronic Devices. Docket no: NHTSA-2010-0053, National Highway Traffic Safety Administration (2013)
9. Broström, R., Ljung Aust, M., Wahlberg, L., Källgren, L: What drives off-road glance durations during multitasking – capacity, practice or strategy? In: Proceedings of the 3rd International Conference on Driver Distraction and Inattention, Gothenburg (2013)
10. Ljung Aust, M., Dombrovskis, S., Kovaceva, J., Svanberg, B. Ivarsson, J.: An empirically based suggestion for reformulating the glance duration criteria in NHTSA's visual-manual interaction guidelines. In: Proceedings of SAE 2013 World Congress, Paper No 2013-01-0444, Detroit (2013)
11. Ljung Aust, M., Rydström, A., Broström, R., Victor, T.: Better ways to calculate pass/fail criteria for the eye glance measurement using driving simulator test. In: Proceedings of the 4th International Conference on Driver Distraction and Inattention, Sydney (2015)
12. Rydström, A., Ljung Aust, M., Broström, R., Victor, T.: Repeatability of the NHTSA visual-manual guidelines assessment procedure. In: Proceedings of the 24th Enhanced Safety of Vehicles Conference, Paper No 15-0240, Gothenburg (2015)
13. Ljung Aust, M., Rydström, A., Broström, R., Victor, T.: Improving the reliability of the eye glance measurement using driving simulator test. In: Proceedings of the 24th Enhanced Safety of Vehicles Conference, Paper No 15-0938, Gothenburg (2015)
14. Broström, R., Bengtsson, P., Ljung Aust, M.: Individual glance strategies and their effect on the NHTSA visual manual distraction test. Transp. Res. F **36**, 83–91 (2016)

15. Angell, L., Perez, M.: Explanatory Material About the Definition of a Task Used in NHTSA's Driver Distraction Guidelines, and Task Examples. Report no: DOT HS 811 858, National Highway Traffic Safety Administration (2013)
16. Translucent Technologies. http://www.translucent.ca
17. Zhong, H., Larsson, P., Rydström, A., Vigneron, J.: SICS - Safe Interaction, Connectivity and State. ViP Technical report no 2012-04634, Linköping (2015). www.vinnova.se/PageFiles/751290059/2012-04634_EN.pdf

The Benefit of Touchless Gesture Control: An Empirical Evaluation of Commercial Vehicle-Related Use Cases

Michael Stecher[1(✉)], Britta Michel[2], and Andreas Zimmermann[2]

[1] Chair of Ergonomics, Technical University of Munich, Boltzmannstr. 15, 85747 Garching, Germany
Michael.Stecher@tum.de
[2] MAN Truck & Bus AG, Dachauer Str. 667, 80995 Munich, Germany
{Britta.Michel, Andreas.Zimmermann}@man.eu

Abstract. Touchless gesture control can be used for controlling various functions inside the vehicle. However, what added value can be created by touchless gesture control? Within a Wizard-of-Oz experiment in the driving simulator, we investigated the potential benefits of gesture control when switching on the cabin light, adjusting the roller blind, setting the radio volume, and calling up an information screen. For several functions, we found that in comparison to conventional controls touchless gesture control potentially increases driving safety, reduces the time required for input, and is physically less demanding. Due to its benefits, participants perceived touchless gesture control as useful and attractive, which makes touchless gesture control a serious alternative to conventional input modalities, provided that it works accurately.

Keywords: Gesture control · Touchless gesture interaction · Wizard-of-Oz experiment · Driving simulator study · Commercial vehicles · Added value · Benefit · Driver distraction · Workload · User acceptance · User experience

1 Introduction

Applying touchless gesture control in vehicles has become feasible due to progress in the development of sensor technology and pattern recognition. Today, in certain serially produced passenger cars, the driver can use touchless gestures to control selected functions of the infotainment system, like accepting or rejecting an incoming telephone call [1] or setting the radio to a new station [2]. The application of touchless gesture control in passenger cars enables a new and innovative way of input and is assumed to increase ease of use.

When it comes to commercial vehicles such as trucks and tractor units, we found no applications of touchless gesture control so far. In the domain of commercial vehicles, the application of new technologies has to provide a concrete benefit, since profitability is an important factor in this context. Hence, any application of touchless gesture control has to create added value in order to represent a serious alternative to conventional interaction modalities and common controls like knobs and pushbuttons. For instance, added value created by touchless gesture control could include:

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_38

- Immediate operation of controls beyond the driver's reach [3] and thus increase of input efficiency
- Decrease of driver distraction and thus increase of safety while driving [4, 5]
- Decrease of physical effort by allowing the driver to remain in his or her "comfort zone"

In the research project "Gesture Control for Commercial Vehicles", we developed several commercial vehicle-related use cases for touchless gesture control. In a first step, we identified functions for which touchless gesture control might create an appropriate benefit. For this purpose, we conducted different studies for user and workplace analysis. An overview of the methods applied within these studies is given in [6]. After identifying several functions, we derived appropriate gestures for those functions from [3] in a second step. Finally, this allowed us to define comprehensive use cases for touchless gesture control from the user's perspective.

Some of those use cases (as well as some other potential use cases for touchless gesture control) are evaluated in the experiment described in this paper. The focus of the experiment is to evaluate the extent to which touchless gesture control is capable of creating added value while driving. To this end, we compared touchless gesture control with conventional input modalities like pushbuttons, touch displays and voice control in a comprehensive driving simulator study.

2 Method

2.1 Functions and Use Cases

The following subsections give an overview of the functions and the corresponding use cases for which we compared touchless gesture control to conventional input modalities.

Switching on the Cabin Light. Each truck is equipped with a cabin light mounted at the headliner of the truck cabin. Usually, the cabin light is turned on and off with a switch or pushbutton. In most cases, switches and pushbuttons are located on the dashboard or the center console of the truck. In this experiment, we installed a push-button on the dashboard to the front right of the driver's seat (cf. Figure 1a).

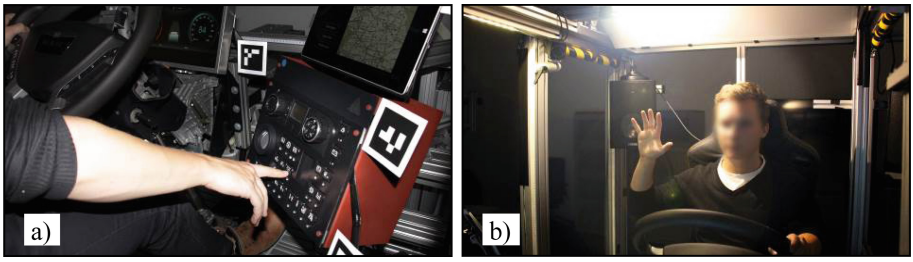


Fig. 1. Switching on the cabin light via the pushbutton on the dashboard (*left*) and via gesture control by pointing the forearm towards the light source and splaying out the fingers (*right*).

To switch on the cabin light via gesture control, participants had to point with their right forearm towards the light source while clenching their hand into a fist. By opening the fist and splaying out all five fingers, the light is turned on. Since the light source in the driving simulator is located right above the driver's seat, the gesture entailed the right arm posture shown in Fig. 1b.

Adjusting the Roller Blind. To prevent the sun from blinding the driver, trucks have a roller blind that can be rolled down along the front windowpane. Most commonly, the roller blind is pulled down by hand. To do this, the driver has to bend forwards and grab the roller blind before pulling it down. In this experiment, for conventional usage, participants had to pull down the roller blind by hand (cf. Fig. 2a).

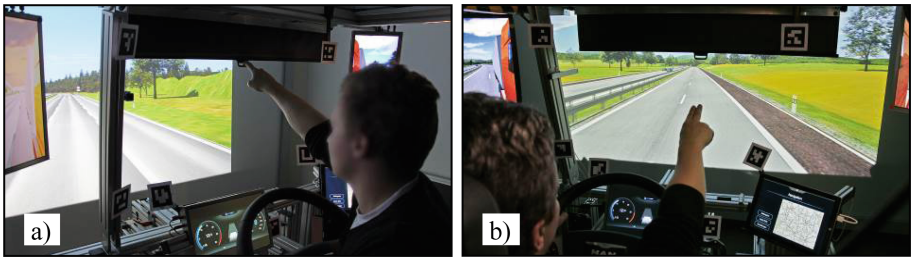


Fig. 2. Pulling down the roller blind by hand (*left*) and adjusting the roller blind electrically via gesture control by moving the hand downwards (*right*)

To adjust the roller blind via gesture control, the roller blind is motorized. In order to move the roller blind down, the participants had to stretch out their right arm and point with one or more fingers towards the roller blind. After bending the wrist downwards, the roller blind begins to move (cf. Figure 2b). By performing a left-to-right horizontal movement with the right hand and forearm above the steering wheel, the roller blind stops moving again.

Setting the Radio Volume. Depending on the vehicle's equipment, there are different ways to set the volume of the radio or the music playback. Usually, the volume is set with a rotary knob at the dashboard. For conventional usage within our experiment, we had two buttons on the steering wheel for setting the volume (cf. Fig. 3a). Pushing the "up" button increases the volume by 10%, while pushing the "down" button decreases the volume accordingly.

In order to set the volume via gesture control, participants had to move their right index finger in circles to the right-hand side of the steering wheel (cf. Figure 3b). The index finger had to point forwards while moving in clockwise circles to increase the volume and in counter-clockwise circles to decrease the volume. Each whole rotation changes the volume by 10%.

For both conventional control and gesture control, the current volume in terms of a percentage was displayed on-screen via the driving simulation, comparable to a head-up display.

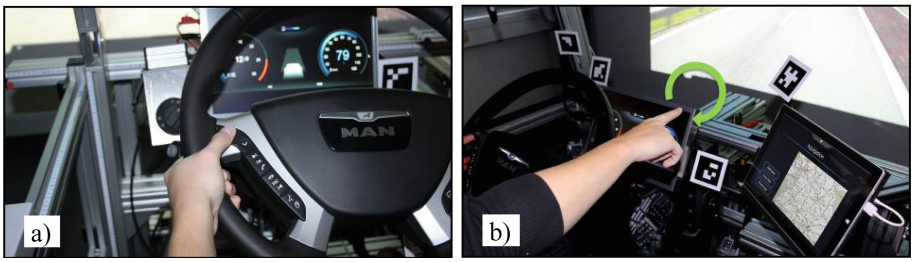


Fig. 3. Setting the radio volume by pressing a button on the steering wheel (*left*) and setting the volume via gesture control by moving the index finger in circles (*right*).

Calling Up an Information Screen. This use case involves a touch display, typically used for navigation and infotainment applications. On this display, participants could call up a so-called information screen that shows relevant information concerning the status of the vehicle, like tire pressure or oil temperature. When called up, the information screen slides in from the upper border of the display with a short animation.

In case of conventional usage, participants could call up the information screen by touching the display at its upper border with one finger and then dragging the finger downwards (cf. Fig. 4a).

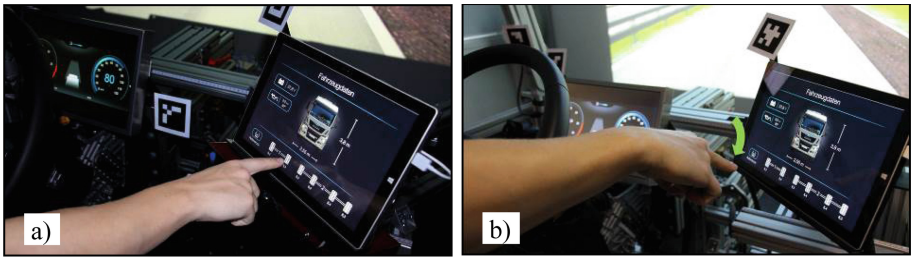


Fig. 4. Calling up the information screen by using the touch display (*left*) and by using touch-less gesture control (*right*).

As alternative to touch input, voice control could also be used to call up the information screen. In order to activate voice control, participants had to push and hold down the push-to-talk button at the steering wheel. The voice command "vehicle information" then prompts the information screen to slide in.

To call up the information screen via gesture control, participants had to place their right hand in front of the display, pointing with the index finger or alternatively with the whole flat hand towards the screen. After bending the wrist downwards, the information screen slides in (cf. Fig. 4b).

In each case, the information screen is displayed for five seconds before it disappears automatically.

2.2 Experimental Setup

Instead of implementing a real gesture recognition system, we set up the experiment as a "Wizard of Oz" experiment. This means that participants performed gestures under the belief that they are controlling the respective function, but in fact another person – the so-called wizard – interpreted all gestures. The wizard was located in a control room outside the driving simulator out of sight of the participants. From the control room, the wizard observed the participants via video connection and controlled all functions remotely according to the observed gestures (cf. Fig. 5a). Besides gesture recognition, the wizard also simulated voice control, for which we established an audio connection from the driving simulator to the control room. Setting up this experiment as a "Wizard of Oz" experiment allowed us to simulate accurate gesture and voice recognition without any misrecognition.

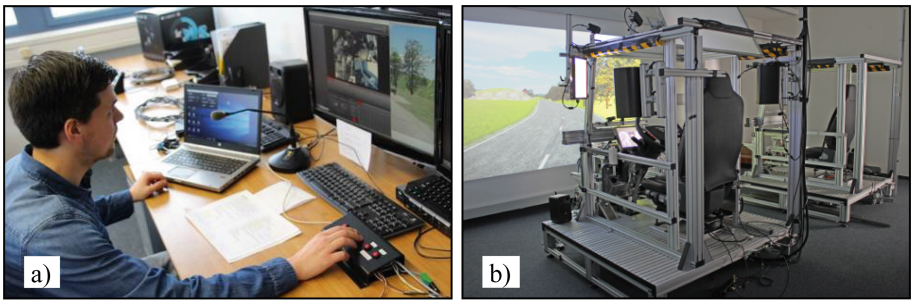


Fig. 5. From the control room (*left*), the wizard observes the participants in the driving simulator room (*right*) and controls the functions remotely.

The experiment was conducted in a static driving simulator consisting of two separate truck mockups (cf. Figure 5b). One of the two mockups includes all conventional controls, while the other one is set up for gesture control. In the latter, we installed a 3D camera to lead the participants to believe that they are interacting with a real gesture recognition system that is actually working. Except for use case-related controls, both mockups were configured identically. The driving scenario consisted of ordinary highway driving during which participants had to control the different functions via conventional controls and gesture control respectively.

2.3 Procedure

Each participant had to perform four driving tasks. The first two consecutive driving tasks were performed in one mockup, after which the participants changed to the other mockup, where they performed the second pair of driving tasks. Since the mockup in which each participant started was randomized, half of the sample used conventional controls first while the other half started with gesture control.

At the beginning of the experiment, participants completed a training session in order to gain familiarity with the simulator and the driving simulation. After that, before each driving task, the different functions were introduced to the participants with their respective controls or gestures. Participants could try out all the different functions and practice performing the gestures until they felt confident.

During each driving task, participants had to operate the functions according to specific tasks, which were described to them (e.g. "please set the radio volume to 80%" or "please turn on the cabin light"). In order to keep the tasks as realistic as possible, each function had to be operated only once for each mockup. After each driving task, participants filled in questionnaires in order to rate their impressions of operating each function. After completing all four driving tasks, we conducted a final interview in which we collected additional subjective impressions from the participants.

2.4 Objective and Subjective Data

Since one of the potential benefits of touchless gesture control is an increase of driving safety, we collected different data relating to the driving performance and driver distraction. With the driving simulator software SILAB [7], we recorded various objective driving data like the vehicle speed, lane-keeping accuracy (SDLP), and time to line crossing (TLC). In addition, we used an eye-tracker to investigate the visual distraction during each use case. In order to evaluate whether touchless gesture control can increase input efficiency, we additionally recorded the time required for each input.

To collect subjective data, we deployed several questionnaires for each use case. To investigate a potential decrease in physical strain by means of gesture control, participants filled in a Raw TLX (RTLX) questionnaire [8] in order to rate the physical demand and workload during the interaction.

Beyond this, we also investigated the appeal of gesture control and the pleasure of gestural interaction, since both are preconditions for the success of a gesture control system. For this purpose, we deployed a reduced UEQ questionnaire [9] to measure the attractiveness, stimulation and novelty as parts of the user experience, as well as the Van der Laan acceptance scale [10] to measure the usefulness and satisfaction of each use case.

3 Results

For each use case, we used objective and subjective data to compare gesture control with conventional control. For statistical comparison, we performed t-tests, Wilcoxon tests and ANOVAs, each with a significance level of 5%.

3.1 Participants

Twenty-eight professional truck drivers and people who drive trucks on a regular basis for professional reasons (e.g. truck mechanics, firefighters, etc.) participated in the experiment. The participants, three of whom are females, were between 24 and 72 years of age ($M = 41.8$, $SD = 14.1$). Twenty-two of them had no experience with touchless gesture control before participating in the experiment.

3.2 Analysis of Driving Performance

In order to analyze the driving performance, we calculated the standard deviation of the vehicle's lateral position (SDLP) for each use case (Fig. 6). The results show a significant difference between gesture control and conventional control when switching on the cabin light ($t(27) = -4.26$; $p < .001$) and when adjusting the roller blind ($t(27) = -2.81$; $p = .009$). Additionally, we found a significant difference in the SDLP when calling up the information screen ($F(2,46) = 5.12$; $p = .010$) due to a significantly lower SDLP for gesture control in comparison to voice control ($p = .010$). When setting the radio volume, we found no significant differences.

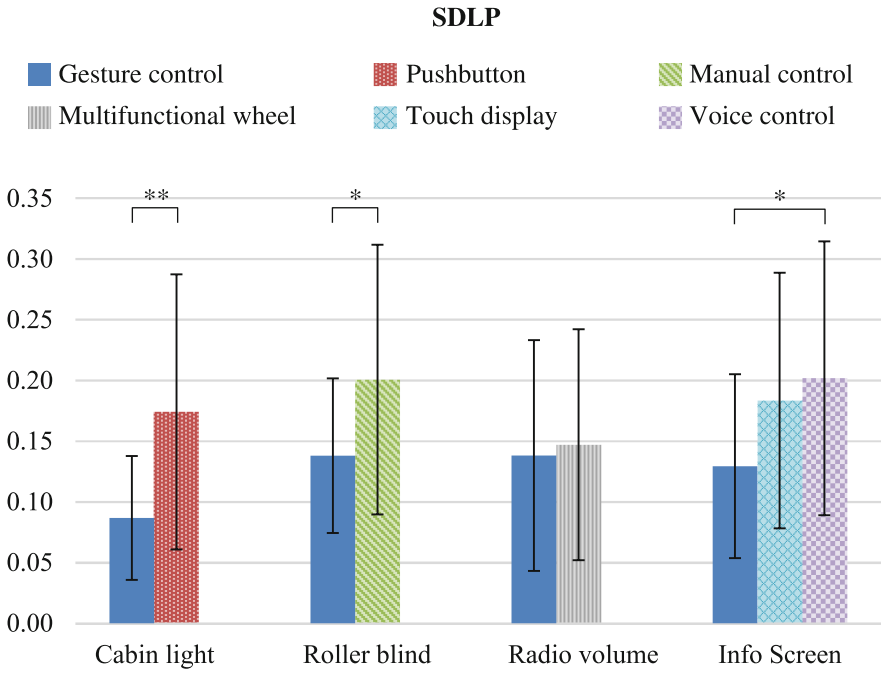


Fig. 6. Comparison of the standard deviation of the lateral position (SDLP) between gesture control and conventional control for each use case. Error bars correspond to ± 1 *SD*; *$p < .05$; **$p < .001$.

3.3 Analysis of Interaction Times

In terms of the mean input duration, we found significant differences in all use cases (Fig. 7). Both switching on the cabin light ($U = -4.56$; $p < .001$) and adjusting the roller blind ($U = -2.28$; $p = .023$) were significantly faster with gesture control, whereas setting the radio volume via gesture control took significantly longer ($t(26) = 5.99$; $p < .001$). When calling up the information screen, the differences were also significant ($F(2,46) = 83.16$; $p < .001$). Post-hoc tests show differences in all use cases ($p < .001$ each).

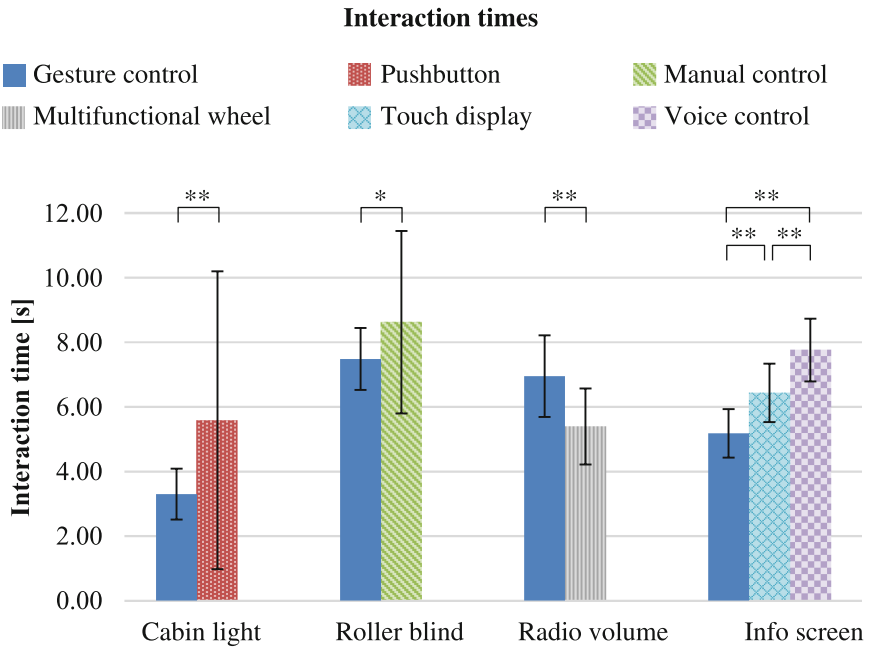


Fig. 7. Comparison of interaction times between gesture control and conventional control for each use case. Error bars correspond to ± 1 *SD*; *$p < .05$; **$p < .001$.

3.4 Analysis of Physical Demand and Workload

Concerning physical demand, the RTLX questionnaire shows significant differences (Fig. 8). Both switching on the cabin light ($U = -4.30$; $p < .001$) and adjusting the roller blind ($t(27) = -8.08$; $p < .001$) are physically less demanding when done with gesture control. Differences in physical demand are also significant when calling up the information screen ($F(2,54) = 12.51$; $p < .001$). The physical demand of calling up the information screen via touch display is significantly higher than via gesture control ($p < .001$) or voice control ($p = .004$). When setting the radio volume, the results show no significant differences.

The results also show significant differences in terms of the overall workload. The overall workload associated with switching on the cabin light is significantly lower with gesture control ($M_{GESTURE} = 4.15$; $SD_{GESTURE} = 2.26$) than with conventional control ($M_{CONV.} = 7.86$; $SD_{CONV.} = 4.51$; $U = -4.41$; $p < .001$). Similarly, when adjusting the roller blind, the overall workload is significantly lower when using gesture control ($M_{GESTURE} = 5.82$; $SD_{GESTURE} = 4.02$; $M_{CONV.} = 9.27$; $SD_{CONV.} = 4.86$; $t(27) = -8.08$; $p < .001$). With regard to calling up the information screen, differences in the overall workload are also significant ($F(2,54) = 6.72$; $p = .002$) due to a significantly lower workload associated with calling up the information screen using gesture control in comparison to interacting with the touch display ($M_{GESTURE} = 4.85$; $SD_{GESTURE} = 2.24$; $M_{TOUCH} = 6.88$; $SD_{TOUCH} = 3.75$; $p = .007$). However, we found no significant differences when setting the radio volume.

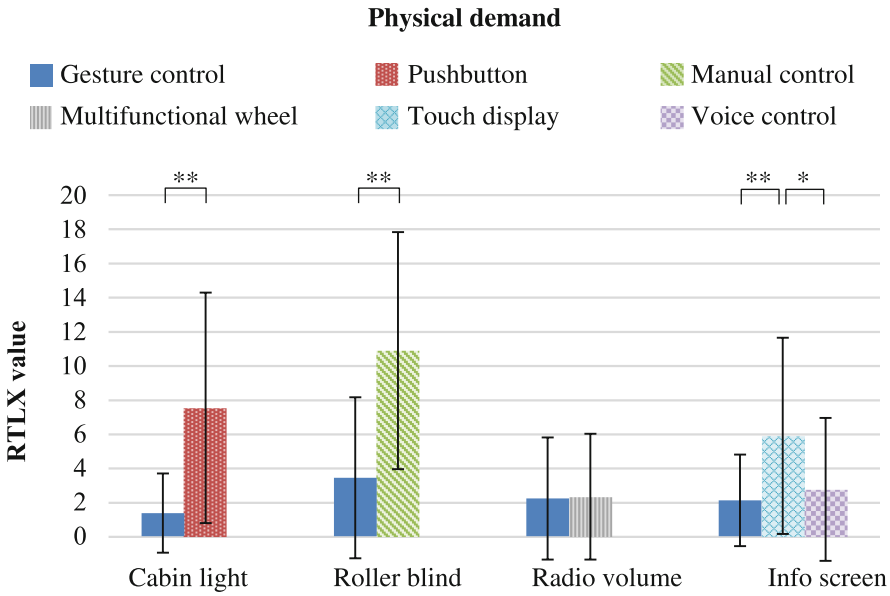


Fig. 8. Comparison of physical demand between gesture control and conventional control for each use case. Error bars correspond to ± 1 *SD*; *$p < .05$; **$p < .001$.

3.5 Analysis of Perceived Attractiveness and Usefulness

We deployed a reduced UEQ questionnaire and the Van der Laan acceptance scale in order to assess different dimensions of user experience and user acceptance. In the following, we report the results for attractiveness and usefulness, since attractiveness and usefulness are two of the most meaningful dimensions for improving user experience and user acceptance.

Concerning attractiveness, we found several significant differences (Fig. 9). In comparison to conventional controls, using gesture control seems to be significantly more attractive when switching on the cabin light ($t(27) = 6.15$; $p < .001$) and when adjusting the roller blind ($t(27) = 6.69$; $p < .001$). We also found a significant difference when calling up the information screen ($F(2,54) = 4.89$; $p = .011$). The attractiveness of calling up the information screen with gesture control was rated significantly higher than with the touch display ($p = .004$). When setting the radio volume, we found no significant differences.

When it comes to usefulness, similar results were found (Fig. 10). Both the usefulness of switching on the cabin light with gesture control ($t(27) = 4.69$; $p < .001$) and the usefulness of adjusting the roller blind with gesture control ($t(27) = 3.92$; $p = .001$) show significant differences compared to the respective conventional control. Furthermore, we found a significant difference when calling up the information screen ($F(2,54) = 4.21$; $p = .020$) due to the significantly higher usefulness of gesture control in comparison to touch interaction ($p = .018$). However, we found no significant differences when setting the radio volume.

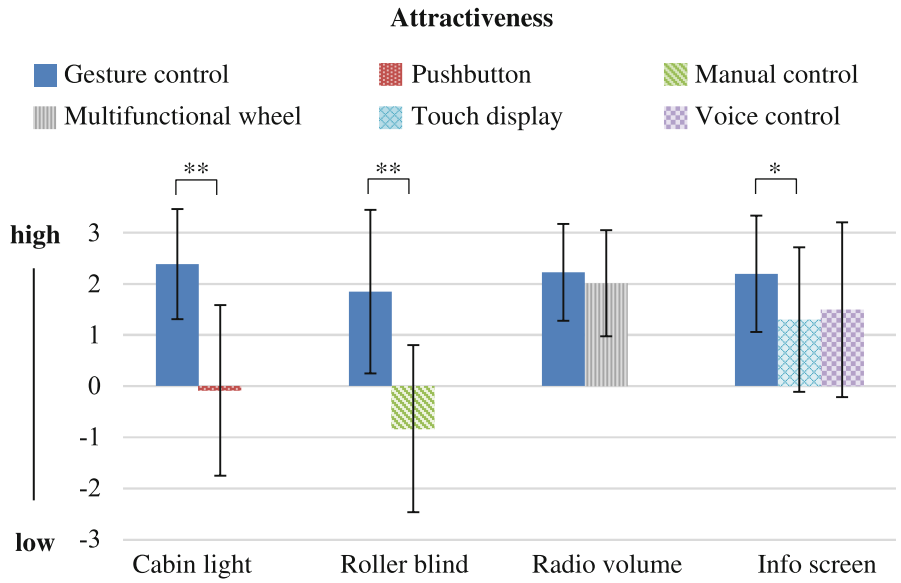


Fig. 9. Comparison of attractiveness between gesture control and conventional control for each use case. Error bars correspond to ± 1 *SD*; **p* < .05; ***p* < .001.

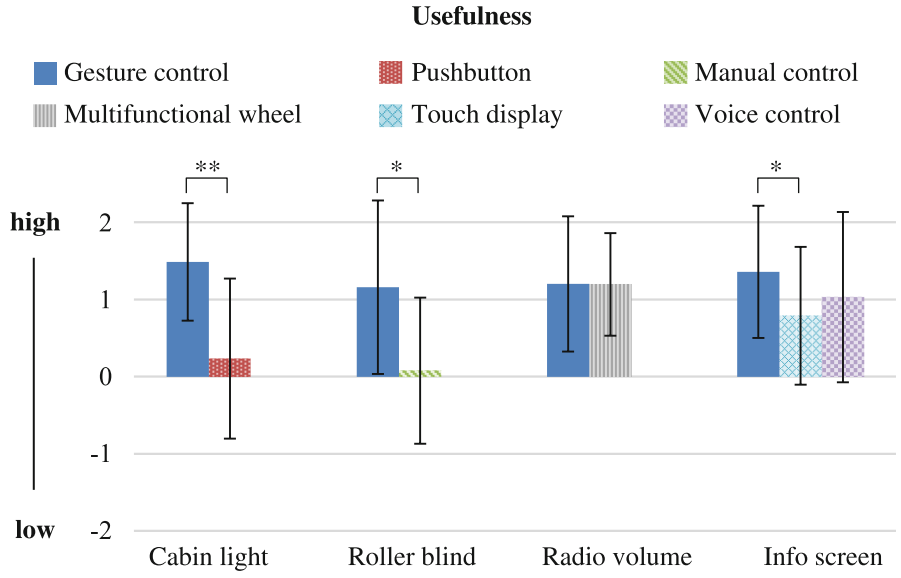


Fig. 10. Comparison of usefulness between gesture control and conventional control for each use case. Error bars correspond to ± 1 *SD*; **p* < .05; ***p* < .001.

4 Discussion

For several functions, the results of this experiment show that touchless gesture control can create added value, provided that the gestures are recognized accurately. Participants benefit from interacting via gesture control, which can be seen especially well in the data associated with switching on the cabin light and adjusting the roller blind. For both functions, gestural interaction is faster, safer, and physically less demanding.

In the case of the cabin light, participants did not need to visually identify the push-button and stretch out their arm to reach it. They simply perform the gesture while keeping their eyes on the road, which results in safer driving.

For the roller blind, gesture control makes driving safer, too. Since participants could remain in their seat and did not need to bend their whole upper body forwards in order to grab the roller blind, they had more control over the steering wheel. For the same reason, adjusting the roller blind with gesture control is physically less demanding and results in a more comfortable adjustment than by hand.

However, when setting the radio volume, we were not able to observe these or similar benefits. It took the participants longer to set the radio volume with gesture control than to set it with the button on the steering wheel. Input via button on the steering wheel is faster because the button can be reached directly with the thumb and does not require the hand to be taken off the steering wheel. At least gesture control is neither physically more demanding nor more distracting in this case.

In comparison to calling up the information screen via voice control, gesture control is faster and less distracting. On the one hand, voice control requires time for participants to push the push-to-talk button and speak out the voice command. Therefore, voice control takes longer than gestural input. On the other hand, the push-to-talk button is not directly reachable with the thumb like the button for setting the radio volume was. It may be the case that participants had first to identify the push-to-talk button visually, which could be the reason for higher distraction compared to gesture control. The pending analysis of eye-tracking data might lead to a clearer picture on this.

Considering the touch display, it was found that interaction via touch display requires more time and is physically more demanding than interaction via gesture control. Just like pushing the button when switching on the cabin light, touch display interaction requires participants to stretch out their arm and position their finger rather precisely on the display. Although descriptive data shows higher SDLP for the touch display than for gesture control, we found no significant differences in the driving performance. The analysis of eye-tracking data might give a more detailed answer on how distracting touch display input really is in comparison to gesture control.

The added value created by touchless gesture control is perceived as such by the participants. For functions where gesture control is beneficial, significant differences in usefulness can be found. Besides that, the usefulness of gesture control was rated quite highly throughout all use cases. This underlines the potential of touchless gesture control.

References

1. BMW 7 Series Sedan : Innovative functionality, http://www.bmw.com/com/en/newvehicles/7series/sedan/2015/showroom/innovative_functionality.html
2. Gestensteuerung < Technik-Lexikon < Technologie < Volkswagen Deutschland, http://www.volkswagen.de/de/technologie/technik-lexikon/gestensteuerung.html
3. Stecher, M., Baseler, E., Draxler, L., Fricke, L., Michel, B., Zimmermann, A., Bengler, K.: Tracking down the intuitiveness of gesture interaction in the truck domain. In: Ahram, T., Karwowski, W., Schmorrow, D. (eds.) Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics 2015 and the Affiliated Conferences, vol. 3, pp. 3176–3183, Elsevier (2015)
4. Akyol, S., Canzler, U., Bengler, K., Hahn, W.: Gesture control for use in automobiles. In: Proceedings of IAPR workshop on machine vision applications, pp. 349–352 (2000)
5. Riener, A., Ferscha, A., Bachmair, F., Hagmüller, P., Lemme, A., Muttenthaler, D., Pühringer, D., Rogner, H., Tappe, A., Weger, F.: Standardization of the in-car gesture interaction space. In: Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '13), October 28–30, 2013, Eindhoven, The Netherlands (2013)
6. Stecher, M., Kremser, F., Michel, B., Zimmermann, A.: Gesture control for commercial vehicles - methods for identifying useful use cases. In: Hirschberg, W., Fischer, P. (eds.) Proceedings of the 7th Grazer Nutzfahrzeug Workshop, May 12–13, 2016, Graz, Austria (2016)
7. Driving Simulation and SILAB. https://wivw.de/en/silab/
8. Byers, J.C., Bittner, A.C., Hill, S.G.: Traditional and raw task load index (TLX) correlations: are paired comparisons necessary? In: Advances in Industrial Ergonomics and Safety I, pp. 481–485 (1989)
9. Laugwitz, B., Schrepp, M., Held, T.: Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In: Heinecke, A.M., Paul, H. (eds.) Mensch & Computer 2006: Mensch und Computer im StrukturWandel, pp. 125–134. Oldenbourg, Munich (2006)
10. Van der Laan, J.D., Heino, A., De Waard, D.: A simple procedure for the assessment of acceptance of advanced transport telematics. In: Transportation Research - Part C: Emerging Technologies, pp. 1–10 (1997)

How Does Awareness Affect Performance in an Automotive Dual Task Condition?

Lee Skrypchuk[1,2(✉)], Patrick Langdon[2], Alex Mouzakitis[1], and P. John Clarkson[2]

[1] Jaguar Land Rover, Research & Technology, Coventry, UK
{LSKRYPCH,AMOUZAK1}@jaguarlandrover.com
[2] Engineering Design Centre, Engineering Department, Cambridge University, Cambridge, UK
{PML24,PJC10,LS578}@cam.ac.uk

Abstract. High levels of Situation Awareness (SA) have been shown to improve performance of an operator in a complex environment. This paper describes an experimental study looking at whether increases in awareness of a non-driving related activity can improve performance in an automotive multitasking condition. 50 participants drove a low-fidelity driving simulator whilst attempting to carry out a number of alternative activities using an in-vehicle interface. The findings demonstrate that increased awareness of the non-driving interface delivers improvements in both driving and non-driving activities. Differences in the driving scenario itself were unable create a difference in perceived workload and SA of the driving activity. Objective measures were sensitive enough to demonstrate effects due to differences in both non-driving task experience and driving scenario under multitasking conditions.

Keywords: Vehicle user interfaces · Situation awareness · Multi-Tasking

1 Introduction

1.1 Background

For many people driving a car is a routine, everyday activity. The experience we build over time increases our ability to deal with a wide variety of driving situations. Novel situations can occur whenever alternative activities disrupt this routine behavior making performance at driving unstable and requiring an ability to multitask within a challenging environment. In many instances, these alternative activities involve functions available on the vehicle that are not frequently used, if at all. This combined with the instantaneous demands of an ever-changing environment makes multi-tasking in car difficult. Therefore, awareness built over time towards this alternative activity whilst driving is required to avoid unnecessary demand.

Outside of the car, we engage in many routine activities such as eating, drinking, using a mobile phone or having a conversation. Over time, practice leads to proficiency at these activities by developing Situation Awareness (SA) toward them. That said, when we combine these activities with a task like driving, the additional demands of this unique combination lacks this same level of SA, potentially leading to poor performance.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_39

There are many examples in research literature, which point to experience over time leading to increased awareness toward driving. One experiment found that advanced driver training delivers improvements in driver SA [1]. Another demonstrated that when participating in a phone conversation whilst driving, both experienced and novice drivers performance was affected, experienced drivers to a lesser extent [2]. This increased awareness has been attributed to more extensive scanning of the roadway environment, leading to a greater knowledge of what is likely to happen [3]. There are also examples of how experience at driving can lead to reduced levels of workload [4] and more situationally aware behavior when carrying out alternative activities whilst driving [5]. Another study emphasized drivers anticipating when to carry out an alternative activity [6].

There are also examples of how alternative tasks, on their own, can be influenced by prior experience or awareness. One study focused on digital cameras, experience with a product was the main predictor of performance [7]. Another study found that training can reduce demographically induced effects of technology acceptance [8]. In a study focused on the effects of multi-modal interfaces, experienced users responded more favorable to a suite of different multimodal combinations than those less familiar with this approach [9]. These examples demonstrate enhancements in performance caused by increases in the level of awareness the user has with the product situation and in being able to anticipate what to do next with their personal goal in mind. There does however appear to be a lack of firm evidence of the effects of experience or increased awareness of a competing activity whilst driving, especially with in-vehicle interfaces that are not typically used frequently. Does experience only improve performance with the competing activity or are performance increments seen for both activities? This paper will address this issue through an experimental study.

1.2 What Is an Activity in an Automotive Context?

The vehicle environment can be divided into two groups of activities. The first are those directly related to driving; Driving Related Activities (DRA). These concern the goal of operating a car safely from one place to another and can be broadly broken down into sub goals such as hazard detection, maneuvering and navigation [10]. The DRA is continuous activity and driven by mixture of endogenous (planned destination) and exogenous (traffic, road condition) stimuli.

The second group are activities unrelated to driving or Non-Driving Related Activities (NRA). NRAs can be anything else that the driver may want to do; or think about, whilst driving that does not directly affect the main goal of getting to the intended destination. NRAs are typically carried out by in-vehicle devices but also include interactions with passengers, mobile phones, food or drink. NRAs are periodic and are mostly endogenous (such as climate and music preferences), although some exogenous motives exist (such as passenger interaction, billboards and scenery). Tasks carried out using integrated devices can be generalized using typical HCI methods of task classification such as GOMS [11]. The main distinction being that the most challenging situation for these devices to be used is whilst the vehicle is in motion. This combination can place significant demand on the driver and isn't considered by HCI classification techniques.

There are a number of factors confounding this. Firstly, the amount of functionality available in the vehicle and the variation in frequency of use of many functions means that some are not used frequently if at all. This leads to out dated mental models and inefficient multitasking. Secondly, multitasking has become a social norm, especially amongst younger people [12, 13]. The amount of active requests for real time communication has led to the term "media multitasking" [14]. Thirdly, drivers; young people especially, are less risk adverse and so are more likely to demonstrate risky driving behavior [15, 16]. The combination of these aspects means that drivers are increasingly likely to want to engage in these combinations of activities whether vehicle OEM's offer suitable functionality or not.

1.3 Situation Awareness in the Vehicle

SA is a proposed theory that discusses how operators deal with complexity when operating a system. An example would be pilots dealing with the significant demand of safely flying a plane [17]. The most common definition is *"the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"* [18]. Perception, followed by comprehension and subsequently the ability to anticipate future events explains how an operator manages a multitude of information types to build an understanding of the surrounding environment. A recent modification of this approach towards driving has proposed that because of a number differences between driving and the professionally oriented domains (such as aviation) there is a key difference in how SA exists within the vehicle environment [19]. This proposes that competing activities (NRAs) reduce awareness to the DRA leading to reduced performance. When attention is directed towards an NRA, awareness to the NRA is built. Fluctuating levels of awareness to the two competing activities (DRA and NRA) leads to inefficient performance. Therefore, to optimize performance on both, maintaining awareness to both is essential.

As discussed, one way of creating awareness is through experience. Pilots, for example, train for a significant period meaning that error free operation is achievable. Novice users make more mistakes and take longer to carry out what should be routine tasks, especially in a multitasking condition [20]. As previously discussed, it has been identified that experience with the driving activity leads to a greater awareness of the driving environment and subsequently higher performance however, there is little evidence of the effect of experience or increased awareness towards an NRA. Therefore, this study will look at how experience with a NRA affects task performance for the both groups of activities in a vehicle environment.

2 Study Design

2.1 Experimental Design and Hypothesis

The experiment used a 2×2 design with independent variables (IV) designed to modulate awareness for two competing activities (DRA and NRA). The first IV was the driving scenario which had two levels that differed by having varying traffic density and

overtaking instances between high (busy), and low (quiet). This IV used independent groups and so each participant only experienced one driving scenario. Awareness to the NRA was modulated using experience to the vehicle interface (Novice-Low and Experienced-High). All participants experienced both NRA conditions. The hypothesis was largely the same as in [19]. The first would be that performance would be highest overall when awareness is considered high for both activities (the Quiet and Experienced combination). Secondly, the lowest overall performance would be seen when awareness is considered low for both activities (the Busy and Novice combination).

2.2 Equipment and Setup

The study was carried out in a low fidelity-driving simulator. The testing platform consisted of a half-vehicle interior, with representative layout and an 85" LCD TV used to present the driving environment with a 90-degree field of view of the forward road scene including side and rear view mirrors. The vehicle interior used a left hand drive half-vehicle property configured with an automatic transmission. The DRA was configured using STISIM (version 3) as a typical highway driving condition. Each driving scenario lasted approximately 10–15 min. The interior contained a Liquid Crystal Display (LCD) instrument cluster, displaying vehicle speed, and a centre console mounted 15" portrait oriented touchscreen upon which all NRA activities were carried out. There were no supplementary steering wheel controls or driving aids. The setup contained a two-camera SmartEye eye tracking solution that recorded moment-to-moment glance locations. All data was collected on a PC allowing for time synchronization. The setup is shown in Fig. 1.

Fig. 1. Low fidelity simulator setup used for the experiment. The setup included driving simulator software, touchscreen interface and data collection platform

2.3 Task Design – NRA and DRA

All NRAs were carried out on the touchscreen. They were created using Qt (v5.7) which also logged the interactions for post processing. Five tasks were implemented to broadly match the type of interactions seen in a modern day in-vehicle touchscreen interface. List search, map manipulation, alphanumeric entry, menu navigation and option selection cover the main types of interaction found in the vehicle. The task steps are captured in Fig. 2.

Fig. 2. Task structure for each of the individual NRA task types used within the experiments. The circle's represent touch points and the arrows instances of swipe interaction

The two driving scenarios developed modeled a typical UK three-lane highway. The scenarios were identical in-terms of road layout but differed both in terms of the amount of traffic and overtaking. In the quiet scenario, overtaking was limited to 2–3, whereas in the busy scenario required 7–8 depending upon the length of the drive. In the busy scenario, the amount of traffic in lanes 2 and 3 was also greater meaning the driver had to be more vigilant when carrying out overtaking maneuvers. The DRA consisted of the participant driving for approximately 10–15 min. They were asked to remain in the left hand lane of the motorway and maintain a speed of 70 MPH. If they needed to overtake a vehicle in order to maintain speed then were expected to do so before immediately returning into the left hand lane. All tasks were carried out once during the driving scenario with 1-min driving only period preceded each multitasking period.

2.4 Dependent Variables

Two subjective measures were taken, Workload using NASA TLX [21] and Situation Awareness Rating Technique (SART) [22]. Both were asked twice, once for each activity to try to aggregate the effect of each on the user. A number of dependent variables were used. For the DRA, steering wheel angle, speed and pedal (throttle and brake) actuations were measured. For the NRA task time, button presses, number of glances and mean total glance time were recorded.

The focus of this study was to explore the effects of awareness towards the NRA. Therefore, a detailed analysis was carried out on measures that indicate the presence of awareness, the first of these is *glances per button push*. Those aware of the next step should be able to find a button with just a single glance, locate and execute. The second is *mean time between button pushes*. A shorter, consistent time would indicate reduced visual search time and thus a greater overall awareness. The variation in driving environment should affect both but if awareness is enhanced, a reduction between the novice and experienced conditions is expected.

2.5 Participants and Procedure

Sixty participants were recruited and split into the two groups, thirty for each driving condition (Busy condition, Mean age = 40.63, SE = 1.66/Quiet condition, Mean age = 41.07, SE = 1.66) with a total of 32 Males and 28 Females. Five participants did not turn up and one had to stop immediately due to feeling the effects of simulator illness. A further four participants had data collection problems during the experiment and were excluded leaving fifty full data sets. All participants were compensated for participating in the trial.

The procedure started with the participants being welcomed to the facility and asked to sign a consent form for their participation. The participant was then asked to sit in the simulator and make themselves comfortable by adjust the seat and steering wheel. The remote eye tracker was then calibrated by asking the participant to look at designated zones around the cab. To ensure that the participant was a novice their first experimental run was tested. In-between runs one and two there was a short training period. This was required to increase awareness of the participant toward the NRA interface. The training was carried out to criteria using a protocol whereby the facilitator walked the participant through each step of each task. The participant then had to carry out each task successfully five times in a row to move onto the next. Any failure during this sequence and the success counter was reset and continued until five successful completions were achieved. The study followed five experimental conditions carried out in the following order:

(1) Pre-trial questionnaire – collection of participant information.
(2) Familiarisation Run – a 5-minute drive to get used to the simulator dynamics. If they required more time they were given the opportunity to repeat this step.
(3) Drive 1 – Novice condition, Followed by NASA TLX and SART.
(4) Training Period – The participants trained to criteria (training criteria below)
(5) Drive 2 – Experienced condition, Followed by NASA TLX and SART.

3 Results and Analysis

The data was processed and significant outliers removed. The data passed normality testing to meet the criteria for Two Factor Analysis of Variance (ANOVA). Factor A was Driving Scenario (Busy, Quiet) and Factor B was NRA Experience (Novice, Experienced). The data was analyzed using Minitab V17. There were no significant effects of age or gender ($P > 0.05$). Only significant results are reported.

3.1 Subjective

For DRA workload there were significant effects of Factor B, [*F (1, 108) = 13.55, P-Value = 0.000*] with the Novice condition creating higher workload. For NRA workload there were significant differences caused by both Factor A [*F (1, 109) = 5.03, P-Value = 0.027*] and Factor B [*F (1, 109) = 32.32, P-Value = 0.000*] with Novice and Busy conditions reporting higher workload. For DRA SART there was a significant effect of Factor B with the Experienced condition creating higher subjective awareness [*F (1, 105) = 29.36, P-Value = 0.000*]. For NRA SART there was a significant effect of Factor B [*F (1, 109) = 39.35, P-Value = 0.000*] again with the Experienced condition promoting greater awareness to the activity.

3.2 Objective

For the performance measures, the data was split into driving only and multitasking epochs. A general overview of both DRA and NRA measures for both driving only and multitasking epochs only are summarized below with only significant results reported [*P < 0.05*].

DRA Measures – Driving Only. Steering wheel position showed significant effect of Factor A [*F (1, 96) = 7.21, P-Value = 0.009*] and Factor B [*F (1, 93) = 28.09, P-Value = 0.000*] with the Novice and Busy conditions generating a greater variation in wheel position. Figure 3 shows the steering wheel variation for driving only and

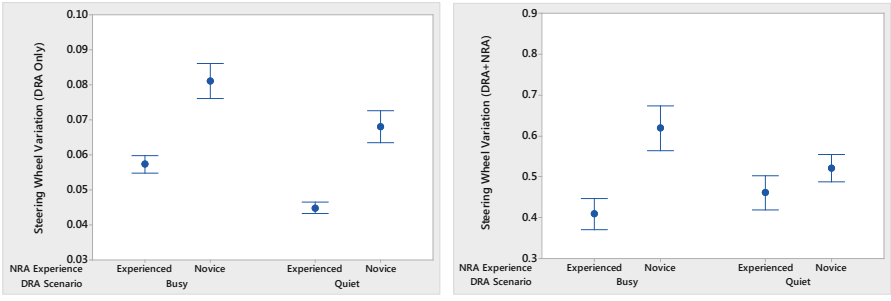


Fig. 3. Steering wheel variation between driving only (left) and multitasking epoch conditions (right). Note a difference between the y scale of the two graphs of a factor of 10 showing significantly less steering variation in the single task epochs.

multitasking epochs. Variation in speed proved significant for Factor B [*F (1, 97) = 4.74, P-Value = 0.032*] with a lower mean speed for the Novice group. Throttle activations also proving significant for Factor B [*F (1, 102) = 8.33, P-Value = 0.005*] with more activations seen by the Novice group.

DRA Measures – Multitasking. Steering wheel position showed significant effect of Factor B [*F (1, 93) = 9.40, P-Value = 0.003*] with the Novices condition generating a greater variation in wheel position. Variation in speed provided a similar result with Factor B proving significant [*F (1, 101) = 5.61, P-Value = 0.020*] with a lower mean speed for the Novice condition. Standard deviation of speed proved significant for Factor B [*F (1, 102) = 52.74, P-Value = 0.000*] the Novice group providing greater variation. Throttle activations also proving significant for Factor B [*F (1, 97) = 8.33, P-Value = 0.005*] with more activations created by the Novice condition.

NRA Measures – Multitasking – All Tasks. Only mean task time proved significant for Factor B [*F (1, 103) = 260.96, P-Value = 0.000*] with Experienced group reporting significantly lower task times than the Novice group.

NRA Detailed Task Analysis. An initial 2-way ANOVA looking at the effects of Driving scenario (Factor A) against the five task types (Factor B) proved insignificant ($P > 0.05$) for all measures. A 2-way ANOVA was carried out looking at the effects of Experience (Factor A) against the five task types (Factor B). All measures results proved significant for Factor A, Factor B and A*B ($P < 0.05$). Post hoc Tukey HSD pairwise comparisons were used to find out where task differences existed. Table 1 shows the groupings for the various measures used.

Table 1. Tukey HSD Comparisons for Objective Measures versus Task Types. The various groups for each measure are donated by the capitalized letters. Tasks with different letters were significantly different from the others

	Task 1	Task 2		Task 3	Task 4	Task 5
Glances per Button Push	A	A		B	A	C
Mean Time Between Pushes	A	A		B	A	C
Task Time	A	B		B	C	D
Button Pushes	A	B	C	C	B	C
Number of Glances	A	B		C	B	D
Mean Total Glance Time	A	B		C	B	D

4 Discussion

The subjective results expose two key findings; firstly, that experience to an NRA in-vehicle interface increases perceived SA and reduces perceived workload when in a multitasking condition. Only workload was significant however, likely due to the significant reduction in task time meaning less effort was required to complete a task. The ability to complete the NRA leads to a benefit for both activities. As awareness on the

NRA task improves, so does the performance with alternative activity; in this case driving, in a multitasking situation. The two different driving scenarios were unable to significantly vary drivers' perception of SA or workload. One explanation for this could be that the difference was not noticeable and so people felt they were able to manage equally as well, in other words their available resources easily met the demand in both scenarios. To confirm this, a closer look at the constituent parts of the SART scores for the DRA (Supply, Demand and Understanding) showed differences within the Demand scale.

For the objective measures during the driving only epochs, unlike the subjective results there was a significant effect of driving scenario, particularly for steering wheel variation and throttle activations. In other words, the user had to make more steering wheel movements during the Busy scenario than the Quiet scenario. This was clearly down to the difference in number of maneuvers carried out between the two scenarios. This does show evidence of a difference being present within the environment even if the participants were unable to report it. This could be due to both steering and accelerator usage being routine in driving and so because of its more implicit nature, may not be available for recall or creating sufficient workload that their SA felt challenged [23].

Experience with the NRA produced a significant effect on most of the driving performance measures during the driving only epochs. This is may be counterintuitive as this is when there is no active alternative task. Closer inspection reveals that this was due to an increased amount of activity in the Novice condition. This was due to more confirmation glances outside of the task epochs, confusion as to whether the task was completed and familiarisation glances because it was the first time the participant had seen the interface. The participants were all experienced with high yearly mileage and so the baseline driving performance was consistent yet the NRA awareness level clearly had an effect on driving control.

Figure 3 shows how much of a difference there was in steering wheel behavior between single and multi-tasking conditions. As expected, variation in the driving only epochs is low, albeit with differences caused by the four experimental groupings. This increased significantly during multitasking epochs, demonstrating the effect on steering behavior of the additional task. During multitasking steering behavior was no different between Quiet and Busy conditions. The active NRA possibly made the users revert to a default behavior that in order to maintain lane position that was consistent regardless of t he driving scenario. This mean they were less likely to attempt overtaking when an NRA was active. This shows a mentality to reduce focus on driving in order to carry out the NRA and agrees somewhat with the subjective findings. For NRA measures, none of the parameters demonstrated significant effect of driving scenario. Experience with the task however demonstrated significance differences for task times, which were significantly shorter with the experienced group. Parameters such as glances and buttons presses however proved insignificant. It is not always the case that fewer glances or buttons presses leads to best performance, especially when scrolling a list or typing a destination.

The post hoc comparison of NRA experience and Task type found differences between the five tasks used for all measures. Table 1 shows consistency of grouping of the tasks across the various measures. Task 2 (typing) and Task 4 (map manipulation)

were the same for all but one parameter. Task 1 (list search) was similar to Task 2 and Task 4 in three of six measures with Task 3 (Option selection) and Task 5 (climate setting) being broadly different from the other grouping as well as different from each other. Tasks 1 (list search), 2 (typing) and 4 (map manipulation) were consistently grouped and were all tasks that had some level of uncertainty associated with them. Task 1 was a list where the target item differed between runs so the user could not repeat exactly what they did previously. A certain amount of visual search was required each time the user attempted the task. Task 2 was the text entry task where the target street name varied. Task 4; the map manipulation, was consistent between runs but was a difficult task to carry out. The uncertainty was mostly associated with interaction, requiring specific actions such as zoom and two dimensional swipe scroll as well as anticipation of where they needed to target on the map. Task 3 was the sub-menu option task buried in the menu hierarchy but because it was consistent improved significantly between runs. Task 5 (climate adjust) was the simplest and improved significantly in the experienced condition. This task was also always present on the screen and so was accessible at any point in time.

This analysis demonstrates the importance of novelty or expectancy within the interface. In this study, Interface tasks with an element of novelty, took longer and showed much less of an improvement than those that contained a level of consistency. This is confirmed by the relative number of button pushes per glance not being significantly different between conditions but the task time being considerable longer in the Novice condition. For a novice essentially all tasks are novel or new and so a certain amount of exploration is required before attempting a task. This is symptomatic of novice behavior where a large amount of visual search is required. This searching period results in attention narrowing toward the in-vehicle interface whilst the user locates the target element and gains experience with the activity [24]. When multitasking, if the user is aware of the interface elements, they can locate and execute an action much more efficiently. As experience is gained, the amount of novelty reduces due to increasing awareness towards the task. Increases in awareness to tasks that contain uncertainty (list selection or typing) compared with those that don't (climate setting) result in smaller performance decrements, just as in [19] with the fixed and variable keypad tasks.

5 Conclusions

For the main hypothesis of the experiment, it was predicted that when multi-tasking, experienced users in the quiet driving scenario would demonstrate highest performance with the two activities because of greater awareness towards them. Whilst this was not confirmed through the subjective measures, there is plenty of evidence from the objective measures that this is the case. The objective results lead to the conclusion that increased awareness of an activity has benefits to both activities when multitasking. The subjective measures however highlighted that users can find it difficult to subjectively perceive differences within the driving environment tested when they contain routine behaviors. Sub goals, such as monitoring traffic or carrying out overtaking were not distinct enough to create a difference between the two scenarios. The overall lack of

significance here could have been down to the methodology. Participants only experienced one of two driving scenarios and so had no real comparison. This could have been counteracted by having a 3rd experimental drive, which involved the alternative condition allowing them to make a comparison between the conditions.

The secondary hypothesis stated that the lowest performing condition would be the Novice user in the Busy condition. This was also confirmed as the lowest performance of all of the groups tested being associated with this condition. The lack of awareness to a task combined with a busy scenario meant performance toward both suffered and thus tasks took longer and there was a greater variation toward the driving activity as a result.

The task-by-task analysis proved that certain interface tasks could help to promote greater awareness over time than others. This has important implications in terms of practical interface design in that the more novel the interface is, the more likely the user to spend time attending to it and thus the greater the implication towards the DRA. Focus should be given in reducing the amount of search time and unique interface elements thus trying to eliminate any novice effect. This could be done by trying to replicate interface techniques that are already familiar to a user. The key implication being that interface systems that can help the user develop awareness towards them inherently or help a user to maintain a level of awareness whilst driving will lead to improved task performance and a greater ability of the user to balance the risks associated with multitasking in the car.

Acknowledgements. This work is funded by Jaguar Land Rover Research through the Centre for Advanced Photonics and Electronics at Cambridge University.

References

1. Walker, G.H., et al.: Does advanced driver training improve situational awareness? Appl. Ergon. **40**(4), 78–87 (2009)
2. Kass, S.J., et al.: Effects of distraction and experience on situation awareness and simulated driving. Transp. Res. F **10**(4), 321–329 (2007)
3. Underwood, G., et al.: Visual search while driving: skill and awareness during inspection of the scene. Transp. Res. F **5**(2), 87–97 (2002)
4. Patten, C.J.D., et al.: Driver experience and cognitive workload in different traffic environments. Accid. Anal. Prev. **38**(5), 87–94 (2006)
5. Schömig, N., Metz, B.: Three levels of situation awareness in driving with secondary tasks. Saf. Sci. **56**, 44–51 (2013)
6. Schömig, N., et al.: Anticipatory and control processes in the interaction with secondary tasks while driving. Transp. Res. F **14**(6), 525–538 (2011)
7. Langdon, P., et al.: The effects of prior experience on the use of consumer products. Univers. Access Inf. Soc. **6**(2), 179–191 (2007)
8. Roberts, S.C. et al.: An evaluation of age, gender, and technology experience in user performance and impressions of a multimodal human-machine interface. In: IIE Annual Conference. Proceedings. Institute of Industrial Engineers, Delhi (2011)
9. Jacko, J., et al.: The effects of multimodal feedback on older adults' task performance given varying levels of computer experience. Behav. Inf. Technol. **23**(4), 247–264 (2004)

10. Gugerty, L., et al.: Situation awareness in driving. In: Fisher, D.L., Rizzo, M., Caird, J.K., Lee, J.D. (eds.) Handbook for driving simulation, pp. 19-1–19-9. CRC Press, Boca Raton (2011)
11. Kieras, D.E., John, B.E.: The GOMS family of user interface analysis techniques: comparison and contrast. ACM Trans. Comput. Interact. **3**(4), 320–351 (1996)
12. Lerner, N., Boyd, S.: On-road study of willingness to engage in distracting tasks, pp. 1–102. National Highway Safety Administration, Washington, DC (2005)
13. Zhang, W., Zhang, L.: Explicating multitasking with computers: gratifications and situations. Comput. Hum. Behav. **28**(5), 1883–1891 (2012)
14. Foehr, U.G.: Media multitasking among American youth: prevalence, predictors, and pairings, pp. 1–39. The Henry J. Kaiser Family Foundation, Menlo Park (2006)
15. Jonah, B.A.: Accident risk and risk-taking behaviour among young drivers. Accid. Anal. Prev. **18**(4), 255–271 (1986)
16. Olsen, E.C.B., et al.: Intersection behaviour of novice teen drivers and experience adult drivers. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 1907–1911 (2005)
17. Endsley, M.R., Bolstad, C.A.: Individual differences in pilot situation awareness. Int. J. Aviat. Psychol. **8414**(2014), 37–41 (2009)
18. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Factors J. Hum. Factors Ergon. Soc. **37**(1), 32–64 (1995)
19. Skrypchuk, L., et al.: Creating inclusive automotive interfaces using situation awareness as a design philosophy. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) **9739**, 639–649 (2016)
20. Grier, R., et al.: The red-line of workload: theory, research, and design. In: Proceedings of the Human Factors Ergonomics Society Annual Meeting, vol. 52, no. 18, pp. 1204–1208 (2008)
21. Hart, S., Staveland, L.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Hum. Ment. Workload **1**, 139–183 (1988)
22. Selcon, S.J., Taylor, R.M.: Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. In: Situational Awareness in Aerospace Operations (1989)
23. Lo, J.C., et al.: Explicit or implicit situation awareness? Measuring the situation awareness of train traffic controllers. Transp. Res. F **43**, 325–338 (2016)
24. Chapman, P., et al.: Visual search patterns in trained and untrained novice drivers. Transp. Res. F **5**(2), 157–167 (2002)

Intelligent Driving Interface Layout and Design Research

Linghua Ran[1(✉)], Hong Luo[1], Yunhao Yan[2], Mengli Yu[2], and Xin Zhang[1]

[1] Ergonomics Laboratory, China National Institute of Standardization, Beijing, China
{ranlh, luohong, zhangx}@cnis.gov.cn
[2] Beihang University, Beijing, China

Abstract. In this paper, we systematically studied the Layout Preference of Intelligent Driving Vehicles. Interviews on layout preference of intelligent driving vehicles were made to collect users' desired zone for the driving information elements. And the subjects were asked to quantify the layouts of six experimental interfaces in terms of the overall satisfaction by two rating ways. After analyzing and dealing with the above-mentioned interviews and experimental results, the interface elements in the experiment are abstracted and the general conclusion of intelligent driving interface layout is obtained, which can help to determine the scientific and rational interface layout.

Keywords: Intelligent driving · Interface layout and design · Focus group · Likert summated rating scale

1 Introduction

As the future of automobile industry, the market volume of intelligent driving industry is expected to reach CNY 121.4 billion in 2020. Intelligent driving sets free the hands of drivers, and thus the driver may be more engaged in other tasks than driving. The vehicle system, as one of the car systems directly accessible by drivers in the cab, provides very complicated functions such as video and entertainment features, information service in addition to fundamental vehicle navigation. Users involve many of their perceptions when working on the vehicle system, and therefore the vehicle system must be specially designed to match the human cognitive characteristics, to provide guidance in many aspects so that the users can do their jobs effectively. Because of the complicated operating environment and functions, the usability and user interface of vehicle information system become significantly important. Among all the factors, interface layout becomes the focus of many researches, because it has direct influence on the efficiency and intuitive feelings of users, which make reasonable layout a critical part in interface design.

This paper carries out systematical studies on information and features on the intelligent driving systems, and further carries out empirical studies through interviews of focus groups and Eprime-based user interface preference analysis, in order to obtain qualitative and quantitative data, so as to provide basis for layout design of intelligent driving vehicles.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_40

2 Interviews on Layout Preference of Intelligent Driving Vehicles

2.1 Interview Topics

By interaction behaviors of drivers, the in-car HMI of intelligent cars may be classified as: main driving interface, auxiliary driving interface, In/ex-car information, interaction and entertainment interface as well as integration of mobile devices and car. It is found through analysis of related documents [1, 2] and investigations into currently available intelligent driving assistance systems [3, 4], the intelligent devices and systems are mainly located in five zones, as shown in Fig. 1. Through abstraction of the 5 zones and presentation of fundamental driving information elements, users can choose their desired zone for the driving information elements in the focus group interview.

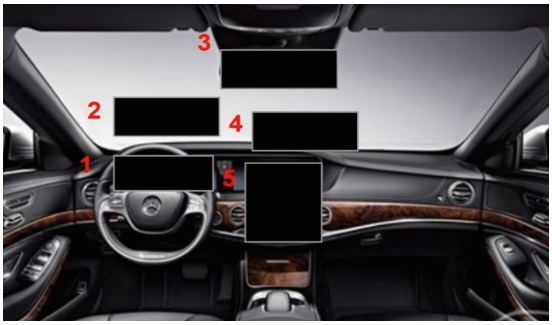


Fig. 1. 5 zones of intelligent driving equipment

Common information elements are derived from collation of related documents [5–7] and basic features of currently available products (such as Tesla and Volvo). According to the findings, such information includes 12 basic functional information and 11 auxiliary functions. As shown in Table 1.

Table 1. Functional information and auxiliary functions

No.	Basic function	Auxiliary functions
1	Speed	Navigation
2	Mileage	Blind spot information system (BLIS)
3	Fuel level	Auto parking
4	Time information	Auto-follow
5	Auxiliary functions	Car-front external conditions
6	Unexpected/error information	Car -rear conditions
7	3D driving vision	Night vision
8	Road conditions/external environment information	Forward-collision warning (FCW)
9	Map and navigation	Auto Lane Change
10	Audio play	Traffic sign recognition
11	Weather conditions	Driver fatigue detection
12	Phone call answering	

2.2 Participants

Eight participants aged 30 to 45 are involved, with male/female ratio of 3:1. Each participant has accumulated driving mileage of over 50 thousand kilometers and has driving experiences in at least 2 days in a week, besides all of them have their private cars of CNY 150 thousand or higher and are strongly interested in intelligent driving.

2.3 Interview

This study is carried out by Focus Groups, concerning the understanding of drivers to layout of various elements in relation to the basic features and auxiliary functions provided on the intelligent driving interface.

Figure 1 was presented to users via PPT, in which 12 basic functions and 11 auxiliary functions are listed and users are required to select the presentation position of important information. In this study, data are collected through questionnaires. Participants are required to deliver their opinions on each question, in order to collect qualitative data.

2.4 Interview Findings

The percentage of presentation positions of various basic features and auxiliary features selected by the participants is shown in Table 2.

Table 2. Percentage of presentation positions of various basic features and auxiliary features selected by the participants

Information Type	Location 1	Location 2	Location 3	Location 4	Location 5
Percentage of presentation positions of basic information selected by participants	Mileage 88% Fuel level 88%	Speed 63% Time 50% Unexpected information 75% Map and navigation 63%	Time 50% Weather 75% Unexpected information 25%		Auxiliaries 88% 3D diving vision 75% Audio play 100% Phone call answering 100% Road conditions 25%
Percentage of presentation positions of auxiliary functions selected by participants	Auto follow 50% Traffic sign recognition 63% Driver fatigue detection 75%	Navigation 63% Night vision 88% Forward collision warning 75% Auto Lane Change 75%	Car-front external conditions 50%	Blind Spot Detection 75% Auto follow 40% Car-front external conditions 50% Car-rear conditions 37%	Auto-park 75% Car-rear conditions

It can be found from Table 2 the presentation positions selected by participants are as follows:

(1) Speed information: 63% participants prefer that speed information is presented in Zone 2, and they widely believe that speed information is important and requires attentions from time to time, while Zone 2 is relatively prominent and is easily accessible by driver without needs to look down.
(2) Mileage information: 88% participants prefer Zone 1 and believe that mileage information is not so important and does not require constant attentions, even if it is related to driving. As such they prefer Zone 1.
(3) Fuel level: 88% participants prefer Zone 1 and believed that it agrees with their usual habits; while some others prefer Zone 2 and believe that fuel level is also very important and should be seen from time to time.
(4) Time information: 50% participants prefer Zone 2 and expect to see it from time to time; while the other 50% prefer Zone 3 and believed that time does not need their constant attentions;
(5) Auxiliary functions: 88% participants prefer Zone 5 and believed that Zone 5 conforms to their usual habits.
(6) Unexpected/error information: 75% participants prefer Zone 2 and believed that such information is relatively important and will be easily accessible in Zone 2; while 25% prefer Zone 3 and wish that all passengers in the car can see it so that they can warn the driver.
(7) 3D driving vision: 75% participants prefer Zone 5 and believe that this function is not frequently used and Zone 5 agrees with their usual habits, while some prefer Zone 2 and said that this may be a help to understand the current road conditions.
(8) Road condition/external environment information: 25% participants prefer Zone 2 and believe that Zone 2 is prominent and easily accessible; another 25% participant prefer Zone 3 since they want it to be seen by all passengers in the car; while still another 25% prefer Zone 5 and believe that Zone 5 agrees with their habits.
(9) 63% participants prefer Zone 2 for map and navigation and believe that map and navigation are closely related to driving and Zone 2 will be very prominent and helps driver to understand the current road conditions; while several other prefer Zone 5, since they believe that the screen is large and the map is complicated and may block the eyesight of driver if located in Zone 2.
(10) Audio play: All participants choose Zone 5 and believe Zone 5 agrees with their habits.
(11) Weather: 75% participants prefer Zone 3 and believe that weather conditions are not very important and may be presented somewhere that can be easily visible; while some other prefers Zone 4 and classified weather as inferior information.
(12) Phone call answering: All participants choose Zone 5 and believe that it agrees with their habits and especially easy to operate by the right hand.

It can be found from Table 2 that the zones where auxiliary functions should be located are as follows:

(1) 63% participant prefer Zone 2 for navigation features which should be presented in 3D vision;
(2) 75% participants prefer Zone 4 for blind spot detection and believes it should be presented in real-time images;
(3) 75% participants prefer Zone 5 for auto parking and believe it should be presented in real-time images;
(4) 50% participants prefer Zone 4 for auto follow, while the other 50% prefer Zone 1 in 2D images.
(5) 50% participants prefer Zone 4 for presentation of external conditions outside the car, while the other 50% prefer Zone 3 and believe 2D, 3D or real-time images are all acceptable.
(6) 63% participants prefer Zone 5 for presentation of rear conditions, while 37% prefer Zone 4 and 3D images;
(7) 88% participants prefer Zone 2 for night vision functions and all participants prefer real-time images;
(8) For forward collision warning, 75% participants prefer Zone 2 and 88% prefer 3D images;
(9) 75% participants prefer Zone 2 for auto lane change, while the other 25% prefer Zone 4, and they prefer 2D or 3D images in terms of the presentation;
(10) 63% participants prefer Zone 1 for traffic sign recognition, while 25% prefer Zone 4; In terms of the form of presentation, 63% prefer 2D images and 25% prefer 3D images;
(11) 75% participants prefer Zone 1 for driver fatigue detection, while the other 25% prefer Zone 4; all participants prefer 2D images in the presentation;

It can be concluded from the above-mentioned data that the basic principle of participants in choosing their preferred zones for related intelligent driving functions is that drivers prefer important information that needs to be seen but does no need any operation to be presented in Zone 1; relatively important information should be presented in Zone 2 so that driver can see it in real time; inferior information that does not require real-time attentions or is expected to be seen by others should be presented in Zone 3 and 4; and multimedia information should be presented in Zone 5 so that drivers can operate easily by hand. 88% participants believe that speed and time information should be presented in intuitive digital form; mileage and fuel level should be presented in digits or dashboard; while other information should be presented in texts, dashboards or images.

The following conclusion can be drawn as for presentation position of five types of information and features in relation to intelligent driving:

(1) Zone 1: preferred to present mileage and fuel level information; auto following, traffic sign recognition (2D) and driver fatigue detection features.
(2) Zone 2: preferred to present speed, time, unexpected/error information, maps and navigation; and best for the navigation features, night vision, forward collision warning and auto lane change features.

(3) Zone 3: preferred to present road conditions/external environment information, weather conditions; and best for car-front external conditions.
(4) Zone 4: preferred to present information of blind spot detection, auto follow and external conditions.
(5) Zone 5: preferred to present auxiliary features, 3D driving images, road conditions/external environment information, audio play and phone call answering; and best for functions such as auto parking and car-rear conditions.

3 Preference Studies on Layout of Intelligent Diving Cars

3.1 Experiment Scheme

Prototype design was made on intelligent driving interface by taking from Tesla and Volvo, one of the most mature intelligent driving products currently available, as a reference, in order to carry out further layout studies on intelligent driving. The optimal interface layout is determined by driver's satisfaction to content positions of various elements on a same interface.

3.2 Materials

Five typical elements, namely speed, road conditions, audio player, fuel level and driving status, are abstracted from the functional interface of currently available products with intelligent driving (Tesla and Volvo). Among them, fuel level and driving status are status information; speed stands for car status, road conditions stand for auxiliary function and monitoring information, while audio player stands for entertainment information. Such information is updated in real time; therefore the presentation position of such three elements adopted alternative designs.

In order to the brand effects created by physical products, fuel level and driving status are indicated in fixed form on the top of the interface, road conditions, audio and speed icons are designed in hypostatized manner with unchanged pattern. By changing the positions of such 3 elements, 6 interface prototypes are designed. See Fig. 2.

3.3 Participants

Totally 24 participants aged 22 to 35 were involved in the interview, with male/female ration of 1:1. The participants aged 24 by average, with 2-8-year driving experience averaging 4 years and 2.5-day drive per week. Participants accepted pre-experiments before the experiments officially begin, in order to get familiar with the experimental environments, and images are presented in a random manner. Finally, data obtained from the experiment were analyzed and studied.

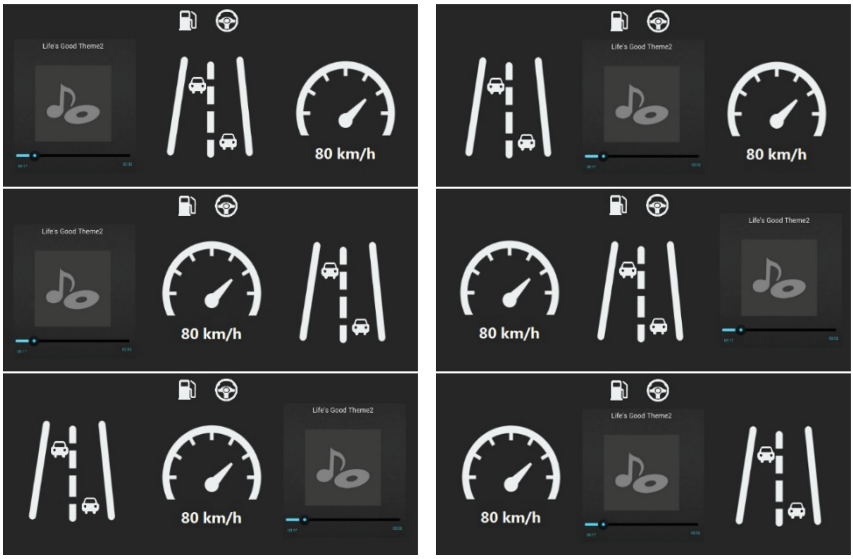


Fig. 2. Experiment materials

3.4 Experiment Process

E-prime software is employed to present the 6 prototype interfaces, to which participants are required to rate their satisfactions to the interface layout. Ratings were made in two ways. In the first way, participants are required to rate their overall satisfaction to all 6 experiment interface layouts immediately after they are presented, by following the Likert 7-scale rating method, where 1 stands for very dissatisfied while 7 stands for very satisfied. In the second way, two random interfaces are presented to participants who are required to choose the better one on basis of their own judgments. 15 paired comparisons are made and finally the rating of each interface will be summed up, where larger rating indicating higher satisfaction. In these ways, the plans are rated and ranked.

3.5 Findings

We have some statistics analysis on effective data and have the following findings:

Obvious difference in Likert 7-scale ratings can be found, the images are ranked from high to low satisfaction as: P4, P5, P1, P3, P6 and P2. The ratings of images in paired comparison from high to low satisfaction are: P4, P1, P5, P3, P6 and P2.

It may be found from Table 3 that the preference of 6 interface designs are basically consistent, of which P4 has the highest satisfaction rating, while P2 the lowest. There is slight difference in ratings of P1 and P5 but they are still very close, as such it is generally concluded that the overall satisfaction rating and paired comparison rating results agree with each other.

Table 3. Ratings of interfaces

Rating method	P1	P2	P3	P4	P5	P6
Likert summated rating scale, mean	4.6	3.1	4.2	5.5	4.7	4.0
Paired comparison of mean ratings	2.8	1.6	2.5	3.7	2.6	1.8

On the interface, the speedometer reflects important information in driving status, as the participants are used to reading left to right, they prefer that speed information reflecting driving status is presented on the left part of the dashboard. At initial stages of the studies, speed information is preferred at Zone 1 and Zone 2 as the interfaces used in focus group interviews. Road conditions stand for auxiliary functions and monitoring information and are located, as less important information, to the right of top important information, which is something like Zone 3 and Zone 4 in the focus group interview. For entertainment information, presented as audio player, participants prefer such information is located in the rightmost part of the area.

4 Conclusion

It is concluded from studies herein that the speedometer that stands for vehicle driving status is of top importance to driving and is preferred at the left part of the driving interface. Audio player may stand for entertainment information or auxiliary functions that are less relevant to driving. Such information is inferior in terms of importance and emergence, and is therefore preferred at the right part of driving interface, which agree with the layout designs of mature products, such as Tesla and Volvo. Less important information may be presented between the driving status and entertainment information, since it does not require constant attentions or is expected to be seen by others.

Acknowledgments. This work is supported by the National Key Technology R&D Program (project number: 2014BAK01B01).

References

1. Maria J.: Visual assistance HMI for use of video camera applications in the car. ISSN 1402–1617/ ISRN LTU-EX–09/003–SE/ NR 2009:003. Luleå University of Technology, (2009)
2. Akhlaq, A.: Smart-dashboard. Augment. Safe Smooth Driv (2010)
3. Tesla. https://www.tesla.com/support, (2016)
4. Volvo (2016)http://www.volvocars.com/intl/about/our-innovation-rands/intellisafe/autonomous-driving, 2016
5. Harvey, C., Stanton, N.A.: Usability Evaluation for In-vehicle Systems. Crc Press, Boca Raton (2013)
6. Rydström, A., Broström, R., Bengtsson, P.A.: comparison of two contemporary types of in-car multifunctional interfaces. Appl. Ergon. **43**(3), 507–514 (2012)
7. Stanton, N.A., Young, M.S., Walker, G.H.: Automating the driver's control tasks. Int. J. Cognit. Ergon. **5**(3), 221–236 (2001)

Human-Machine Interface Design Development for Connected and Cooperative Vehicle Features

William Payre(✉) and Cyriel Diels

Centre for Mobility and Transport, Coventry University, Coventry CV1 5FB, UK
{william.payre,cyriel.diels}@coventry.ac.uk

Abstract. This paper discusses the design and evaluation of a number of connected and cooperative vehicle sign designs which have been developed within the UK Connected Intelligent Transport Environment project (UKCITE). As part of the first phase of the project, the sign design of four different applications were developed and evaluated: Emergency Electronic Brake Lights, Emergency Vehicle Warning, Traffic Condition Warning, and Road Works Warning. Whereas some of the feature made use of existing signs (e.g. road works warning), other applications required new signs. Appropriateness of the signs were evaluated by 21 participants who were shown videos of relevant traffic scenarios with the different signs displayed at appropriate moments. Results are discussed in the context of their appropriateness and suggest that existing standard signs may not appropriately represent new connected vehicle features requiring new design solutions.

Keywords: Human-Machine Interface · Connected vehicles · Virtual gantry · Sign design

1 Introduction

This paper discusses the design and evaluation of connected and cooperative vehicle HMI concepts which have been developed within the first phase of the UKCITE project. As part of the first phase of the project, four different applications were developed and evaluated: Emergency Electronic Brake Lights, Emergency Vehicle Warning, Traffic Condition Warning, and Road Works Warning.

The UK Connected Intelligent Transport Environment (UKCITE) is a project to create the most advanced environment for testing connected and automated vehicles (www.ukcite.co.uk). It involves equipping 40 miles of urban roads, dual-carriageways and motorways with combinations of different "talking car technologies" including LTE-V, DSRC, and Wi-Fi. It will test Vehicle to Vehicle (V2V) and Vehicle to infrastructure (V2I) communications and interoperability between vehicles manufactures and technology providers and will trial the use of a ITS connected "app" for Virtual Road side and driver messaging. Finally, it will also test street level Wi-Fi to understand if it is a viable technology to provide convenience and or automotive related services.

Driver information provision is the critical interface for any V2X application and, hence, the design and impact of driver information is critical for such applications.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_41

Nowadays, vehicles are increasingly equipped with In-Vehicle Information Systems (IVIS). By providing drivers with information such as in-vehicle signage, traffic information or advance warning signals, these systems have the potential to improve road capacity, road safety, and user comfort. However, there is also the risk that these systems may lead to driver distraction. Indeed, IVIS are getting more complex and may require or tempt drivers to set and monitor them, which may not only negatively affect user acceptance but also road safety. Hence, this expansion of functions and systems introduces new challenges for HMI designers and a balance has to be found between HMI effectiveness in affecting drivers' behaviour, system acceptance, and driver distraction.

To date, most IVIS functions are provided via nomadic devices (e.g. mobile phones, SatNavs). In future, these functions are expected to become also available via in-vehicle displays [1] including the instrument cluster, centre stack display, or Head Up Display. Within UKCITE, driving simulator studies are planned to determine the most appropriate location and presentation format. It is expected that this will be feature specific and a function of criticality and urgency, and resulting information priority. However, in this paper we first consider the design of V2X signs to convey four safety related-features: Emergency Electronic Brake Lights, Emergency Vehicle Warning, Traffic Condition Warning, and Road Works Warning. Each of the four features are briefly described below.

The Road Works Warning (RWW) feature informs drivers when approaching road works areas. Via roadside infrastructure, it provides information on current valid roadwork and associated constraints. The Traffic Condition Warning (TCW) feature provides the driver with information regarding the traffic situation, e.g. congestion ahead. The Emergency Electronic Brake Lights (EEBL) feature aims at improving safety in a dense driving environment, by avoiding rear end collisions, which can occur if a vehicle ahead suddenly brakes, specifically in situations with low visibility. It displays a warning to the driver to make them aware of a vehicle ahead performing an emergency braking. Finally, the Emergency Vehicle Warning (EVW) feature warns drivers about an approaching emergency vehicle and provides them with advanced warning to take timely and appropriate actions, e.g. lane change. Whereas some of the sign designs are based on existing roadside signs (i.e. RWW, TCW), other applications require the design of new signs (e.g. EEBL and EVW).

The human factors issues for these applications relate to the question as to what information to present and should take into consideration the need for safe, usable, and effective interfaces. Previous studies have shown that traffic signs symbols are not always understood by drivers [2, 3] and that the comprehension of traffic signs is related to three ergonomic principles: compatibility (i.e. spatial, conceptual, physical), familiarity (i.e. the frequency of the sign on the road) and standardization (i.e. the consistency and homogenous representation of forms, colours, symbols and directions and so forth, in all traffic signs for displaying a similar message) [3, 4]. Based on these principles, the present study focuses on the symbols representing the RWW, TCW, EEBL and EVW applications, and aims at designing two new signs for the EVW and EEBL features. Along with the ergonomic principles, the UK Highways Agency standards have been taken into consideration to design the EEBL and EVW symbols.

As mentioned, the aim of this study was to develop and evaluate these sign designs focusing on users' comprehension and understanding of the expected response. It was expected that signs following the ergonomic principle of standardization, compatibility and familiarization will be considered the most appropriate by individuals.

2 Method

2.1 Participants

The sample consisted of 24 participants, however three of them were removed from the analysis since they did not fill in the questionnaire entirely. They were free to withdraw from the study at any time. Participants were 13 male and 8 female students and staff members from Coventry University. They were 22.5 year-old on average (SD = 5.4). Sixteen were licensed drivers, with 0 to 18 years of driving experience either in the UK or in the European Union (M = 5.1, SD = 5.73). They drove on average 4706 miles a year (min = 0, max = 12000, SD = 4484).

2.2 Materials

The demographic questionnaire contained questions on participants' gender, age, whether they have a driving license, where did they get their driving license and average yearly mileage.

The signs shown to participants for the RWW, TCW features were the same as those used on the UK roadside (see Figs. 1, 2).

Fig. 1. Signs 1 (left) and 2 (right) for the Road Works Warning (RWW) feature

Fig. 2. Signs 1 (left) and 2 (right) for the Traffic Condition Warning (TCW) feature

EEBL sign 1 originated from the DRIVE C2X project [5] and was slightly redesigned to match UK Highways England road gantry standards. The ISO 7000-2569 Passenger car X.02 symbol was also used to illustrate the two vehicles represented. EEBL sign 2 signifies a generic *danger* warning as also used in the UK Autodrive project [6]. EEBL sign 3 was based on existing car features such as *press clutch pedal* in order to be able to start the engine. A pedal below the shoe was added to indicate the expected driver response (Fig. 3).

Fig. 3. Signs 1–3 for the Emergency Electronic Brake Lights (EEBL) feature

Figure 4 shows the EVW signs used. 1 & 2 were based on the ISO-2565-2004 Emergency first aid vehicle symbol. On top of that symbol, sign 2 had the ISO-7000-1421 Interior compartment illumination symbol with a 180° rotation. EVW sign 3 was designed using a front car symbol with the ISO-7000-1421 symbol on top of it.

Fig. 4. Signs 1–3 for the Emergency Vehicle Warning (EVW) signs

2.3 Experimental Design and Procedure

After a short introduction, explaining that the study aimed at investigating the perception of new safety-related traffic signs displayed in connected vehicles, the 24 participants were shown 12 videos. These videos illustrated situations in which signs were displayed in the vehicle while driving manually at 70 mph on a UK motorway. The first perspective was used so participants could imagine themselves driving the vehicle. All signs were displayed at the bottom part of the screen during 3 secs (Fig. 5). The first video was an example to familiarize the participants with the material of the study. The following eleven videos were displayed in a randomised order. Six of them illustrated the EEBL feature, among which the three sign design alternatives were displayed respectively with or without a vehicle performing a harsh braking ahead.

Fig. 5. Screenshot of the video displaying the Road Works Warning (RWW) feature

Three videos illustrated the three different EVW design signs, and one video for both the TCW and RWW. All the videos were presented in a randomized order. After each video, participants answered three questions on a paper questionnaire:

- *What do you think this sign mean?*
- *What would you do in response to seeing this sign?*
- *Following the sign you were shown, what would you expect to see on the road next?*

Following the presentation of the videos, they were provided the definitions for each of the four features illustrated by the eleven signs they saw on the videos. They were then asked to fill out a further questionnaire to rate how appropriate they considered each of the signs for each of the different features. A 7-point Likert scale was used which ranged from *1: Not at all appropriate* to *7: Very appropriate*. Finally, participants filled in a demographic questionnaire. In the current paper, only the results related to the appropriateness of the signs are presented.

3 Results

3.1 Emergency Electronic Brake Lights (EEBL)

A one-way within subjects ANOVA was conducted to assess which sign design alternative was considered the most appropriate to illustrate the EEBL feature. There was a significant effect of the EEBL design alternatives on their appropriateness, Wilks' Lambda = .43, $F(2,19) = 12.88$, $p = .000$. Three paired samples t-tests were conducted to make post hoc comparisons between conditions. A first paired samples t-test indicated that there was a significant difference in the scores for EEBL sign 1 ($M = 5.05$, $SD = 1.77$) and 2 ($M = 2.33$, $SD = 1.8$); $t(20) = 4.83$, $p = .000$. A second paired samples t-test indicated

that there was not a significant difference in the scores for EEBL sign 1 (M = 5.05, SD = 1.77) and 3 (M = 4.05, SD = 1.96), t(20) = 1.57, $p = .13$. A third paired samples t-test showed that there was a significant difference in the scores for EEBL sign 2 (M = 2.33, SD = 1.8) and 3 (M = 4.05, SD = 1.96), t(20) = −3.31, $p = .003$.

3.2 Emergency Vehicle Warning (EVW)

Similarly, a significant mean effect was also found for the EVW feature, Wilks' Lambda = .56, F(2,19) = 7.39, p = .004. Three paired samples t-tests were conducted to make post hoc comparisons between conditions. A first paired samples t-test indicated that there was a significant difference in the scores for EVW sign 1 (M = 4.19, SD = 1.86) and 2 (M = 5.86, SD = 1.32); t(20) = −3.47, $p = .002$. A second paired samples t-test indicated that there was not a significant difference in the scores for EVW sign 1 (M = 4.19, SD = 1.86) and 3 (M = 4.43, SD = 1.83), t(20) = −.4, $p = .69$. A third paired samples t-test showed that there was a significant difference in the scores for EVW sign 2 (M = 5.86, SD = 1.32) and 3 (M = 4.43, SD = 1.83), t(20) = 2.8, $p = .011$.

3.3 Traffic Condition Warning and Road Works Warning (TCW, RWW)

Two paired samples t-test were conducted to compare the two design alternatives of the RWW signs and the TFC signs. There was not a significant difference in the scores for RWW sign 1 (M = 6.24, SD = 1.3) and 2 (M = 5.1, SD = 2), t(20) = 1.91, $p = .07$. Similarly, there was also not a significant difference in the scores for TCW sign 1 (M = 5.57, SD = 1.66) and 2 (M = 5.57, SD = 1.69), t(20) = .00, $p = 1$.

Finally, no significant impact of gender, driving experience and whether or not being a licensed driver on the appropriateness of the signs was found.

4 Discussion

Participants estimated that the EEBL sign 1 was the most appropriate to illustrate this feature, although no significant difference was found with EEBL sign 3. This result could be explained by the ambiguousness of the suggested action to perform by EEBL sign 3. Indeed, the shoe is above a pedal, but it could be either the brake or the gas pedal. Besides, using the car engine to brake could also be an adequate action to avoid a collision consecutive to an emergency braking performed by a car ahead on the same lane. With regard to EEBL sign 2, the exclamation mark could be too elusive to suggest an emergency braking. It may also not properly match the physical compatibility principle to illustrate properly this warning.

Concerning the EVW feature, participants declared that sign 2 was the most appropriate to illustrate this feature. The light on top of the roof and the ambulance cross on the side of the car could be considered more representative than the ambulance only (i.e. sign 1) or a regular car with a light on top (i.e. sign 2). Additional qualitative data should be investigated in another paper to assess to what extend an ambulance is representative

of all emergency vehicles. The meaning of the ISO-7000-1421 Interior compartment illumination symbol could also be confusing and should be investigated further.

Finally, no significant differences were found for both RWW sign 1 and 2 and TCW sign 1 and 2. However, RWW sign 1 (i.e. without any legend) was better evaluated than the RWW sign 2 (i.e. legend only). Overall, participants declared that these signs were appropriate to illustrate RWW and TCW features, which could be explained by the familiarity with these signs. Moreover, driving experience did not have a significant impact on the estimated appropriateness either. All these results were congruent with previous research [7].

Note that in the current study no acoustic warning signals were displayed when signs popped up in the videos. Pairing visual and acoustic warning signals could help drivers detecting and understanding the criticality of a situation [8]. Therefore, the comprehension of the signs might be easier when they are displayed along with an appropriate acoustic warning, depending on the criticality of the situation. For example, in this study EEBL warning is more critical than TCW and an acoustic signal could underline this difference.

The signs were shown only 3 secs in each videos. Further research is needed to investigate what is the most appropriate duration to display such warning in the vehicle in order to drivers' perceptive capacities and acceptance. Besides, the location in the vehicle of these signs (e.g. on an instrument cluster, a Head-up display, a centre console or a mobile phone) could also have an impact on the perception and acceptance of these feature.

Whether having a driving licence did not seem to have an influence on drivers' evaluation of the appropriateness of the signs. Nonetheless, it could be argued that experience with these signs could enhance their comprehension and should therefore be considered in future research.

Finally, adding a legend for each sign could improve drivers' understanding. This is common practice when displaying this type of information on overhead gantries and road-side signs, and it can also help showing additional information such distance or location. Nevertheless, adding legends to signs could lead to crowding, which is the difficulty to recognise familiar objects when located in our visual periphery [9], which may lead to longer response times.

5 Conclusion

The existing RWW and TCW signs were considered appropriate whereas the EEBL and EVW were considered less adequate. Indeed, the results showed differences on the appropriateness of the signs according to their definition. Nonetheless, apart from the conceptual ergonomic guidelines, little is known about the reasons explaining such differences. Qualitative data, such as participants verbatim, is needed to further understand how individuals understand and perceive these signs. This is important since a danger warning can potentially be interpreted in different ways, hence hindering the positive impact of additional in-vehicle information on traffic safety. Findings of the

present study are important for designers, because they show that some existing standards signs do not illustrate appropriately new in-vehicle safety-related information features yet.

References

1. Everis connected car report (2015) www.everis.com/global/WCRepositoryFiles/everis%20connected%20car%20report.pdf, 2015
2. Al-Madani, H., Al-Janahi, L.R.: Role of drivers' personal characteristics in understanding traffic sign symbols. Accid. Anal. Prev. **34**, 185–196 (2002)
3. Shinar, D., Dewar, R.E., Summala, H., Zakowska, L.: Traffic sign symbol comprehension: a cross-cultural study. Ergonomics **46**, 1549–1565 (2003)
4. Ben-Bassat, T., Shinar, D.: Ergonomic guidelines for traffic sign design increase sign comprehension. Hum. Fact. **48**(1), 182–195 (2006)
5. DRIVE C2X – Accelerate cooperative mobility. http://www.drive-c2x.eu/use-06
6. UK Autodrive - http://www.ukautodrive.com/
7. Ng, A.W., Chan, A.H.: The effects of driver factors and sign design features on the comprehensibility of traffic signs. J. Saf. Res. **39**(3), 321–328 (2008)
8. Winkler, S., Kazazi, J., Vollrath, M.: Driving with a multi stage warning system in the head-up display–How do drivers react upon it? In: Proceedings of the human factors and ergonomics society Europe (2016)
9. Whitney, D., Levi, D.M.: Visual crowding: a fundamental limit on conscious perception and object recognition. Trends Cognit. Sci. **15**(4), 160–168 (2011)

How Does Eye-Gaze Relate to Gesture Movement in an Automotive Pointing Task?

Bashar I. Ahmad[1(✉)], Patrick M. Langdon[1], Lee Skrypchuk[2], and Simon J. Godsill[1]

[1] Engineering Department, University of Cambridge, Trumpintgon Street, Cambridge CB2 1PZ, UK
{bia23, pml24, sjg30}@cam.ac.uk
[2] Research & Technology, Jaguar Land Rover, Whitley, Coventry CV3 4LF, UK
lskrypch@jaguarlandrover.com

Abstract. With the recent advances in the gesture and eye-gaze tracking technologies, such trackers are increasingly finding applications in the automotive domain. They not only offer the means to measure the driver's workload (manual, visual and cognitive), their outputs can also be used as input modalities to improve the usability of in-vehicle interactive displays, such as touch-screens. This paper presents a preliminary study on analyzing the relationship between the motor and visual attention associated with the secondary task of interacting with an in-car display whilst driving. In particular, the response time and misalignment between the motor and visual attention during the pointing-selection tasks are considered. They are shown to be highly affected by the driving conditions. This study serves the purpose of devising effective approaches to combining data from gesture and eye-gaze trackers to simplify interacting with displays in vehicles.

Keywords: Touch-screen · Eye gaze tracking · Gesture · Interactive display

1 Introduction

Interactive displays, such as touch-screens, are becoming increasingly prevalent in modern vehicles [1, 2]. This is predominately motivated by the proliferation of the touch-screen technology as well as the ability of such displays to present large quantities of information related to the In-Vehicle Infotainment Systems (IVIS). They also promote intuitive interactions via free hand pointing gestures, particularly for novice users [1–4]. Additionally, other types of displays, such as Head-Up Displays (HUDs) and general 3D displays or projections, have the potential of providing a more immersive driving experience and are becoming increasingly commonplace in the automotive environment [5, 6]. For instance, the Jaguar Land Rover HUD windscreen incorporates laser holography [7]. Devising effective means to easily interact with an in-car HUD or projections is an open research problem.

Nevertheless, interacting with an in-vehicle display can require dedicating a considerable amount of attention (visual, cognitive and manual) that would otherwise be

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_42

available for the primary task of driving [4]. Due to road and driving conditions, the user's input can also be subject to high levels of perturbations leading to erroneous on-display selections [8]. Attempts to rectify an incorrect selection or adapt to the noisy environment can tie-up more of the driver's attention. Consequently, such distractions can significantly affect the driving performance and can have potential safety implications [9, 10].

Current advances in sensory technologies has led to the widespread use of gesture trackers (e.g. in the gaming industry) as well as their introduction to automotive applications [11–13]. This is driven by their potential to extend Human Computer Interaction (HCI) beyond traditional keyboard/cursor, rotary knobs, touch input, etc. They offer new opportunities to simplify interactions with the IVIS, for example BMW 7 Series cars have gesture control for some features. Predictive in-vehicle displays [14–16], also utilize a gesture tracker in conjunction with suitable probabilistic endpoint prediction algorithms, to determine, well in advance, the icon the user intends to select on the in-car display via a free hand pointing gesture. They enable *mid-air* selection where an interface item is automatically selected by the system once an inference certainty criterion is met [15], e.g. prior to the user physically touching the display. Thereby, this predictive solution can substantially improve the usability of in-car interactive displays by reducing distractions-workload of a pointing-selection task [15].

On the other hand, eye-gaze trackers are primarily used to examine human performance (e.g., measuring workload) and a corpus of literature exists in this area [17]. This includes modelling visual demand via Fitt's law, which is discussed in the next section. Nonetheless, eye-gaze measurements offer vital information on the user's intent and can serve as an input modality to an interactive display as in [18, 19]. Such studies employ eye-gaze observations with other input modalities (e.g. gesture) in a sequential manner where only one modality is utilized at any point in time. Thus, the eye-gaze and gesture data are not fused or used simultaneously.

This paper presents a preliminary study on analyzing the relationship between the eye-gaze and pointing gesture movements during the secondary task of acquiring (pointing and selection) on-screen interface components whilst driving. It is the first step towards realizing a multimodal automotive solution, which can effectively fuse data from in-vehicle gesture and eye-gaze trackers; both continuously monitoring the driver behavior-movements. This aims to reduce the effort and distractions of interacting with an in-vehicle display. In particular, the focus here is on the driver's visual and motor response times as well as the misalignment between the motor and visual attention during the pointing-selection tasks. Therefore, this study mainly addresses the following two question from data pertaining to ten participants and collected in a fixed mid-fidelity simulator:

Question 1: How does the divided attention scenario and driving task (lane changing or free driving) affect the motor and visual response times?
Question 2: What is the time delay (misalignment) between the first on-screen glance and the start of the pointing gesture during a target acquisition task?

The remainder of this paper is organized as follows. After providing a brief background of key related models from the HCI area in Sect. 2, the method is described in Sect. 3. The results are then discussed in Sect. 4 and conclusions are drawn in Sect. 5.

2 Background

Whilst the objective of this paper is *not* to model or predict the visual and/or motor demand of the secondary task of interacting with an in-vehicle display, below we briefly introduce seminal human models, namely Fitts' and Hick-Hyman law as well as their combination as in [20].

According to the human motor movement model, Fitts' law [21], which is widely used in the HCI area, the index of difficulty ($\mathcal{J}_D$) and the movement time (T_P) associated with acquiring an interface component are given by:

$$\mathcal{J}_D = \log_2(1 + L/W) \tag{1}$$

and

$$T_P = a + b\mathcal{J}_D, \tag{2}$$

such that W and L denote the width of the target Graphical User Interface (GUI) item and its distance from the starting position (e.g. steering wheel) of the pointing object (e.g. pointing finger), respectively [21–23]. Parameters a and b are estimated empirically from data. Thus, Fitts' law stipulates that the pointing time T_p scales linearly with the index of difficult. It is dictated by the ratio between the amplitude of the motor movement, which is represented by L, and the simplicity/ease of the required motor movement, which is proportional to the width W of the target icon. The selection task can therefore be simplified and expedited by increasing the target size (i.e. larger W) or moving it closer to the pointing apparatus (i.e. smaller L); these observations form the basis for the pointing facilitation strategies [22]. It is noted that recent studies such as [24] apply Fitts' law to quantify the visual demand of selecting GUI icons on an in-vehicle display during driving, e.g. the total and mean glance durations of a target acquisition task.

Whereas, the Hick-Hyman law [25, 26] expresses the decision time of a human response to a visual stimulus as a function of the information content of this stimulus. In the context of this paper (i.e. the task of interacting with a display), Hick-Hyman law states that the time required to choose an interface component (including search, identify and possibly plan action), otherwise dubbed response time, is given by:

$$T_R = c + d\log_2(1/p), \tag{3}$$

where p is the probability of the stimulus being present, i.e. the likelihood of a given GUI icon being the target of the pointing task [20]. Constants c and d are empirically derived. Hence, Eq. (3) stipulates that the response time, prior to the onset of the pointing gesture, is a linear function of the information content defined by $\log_2(1/p)$.

If the interface has N selectable components that have an equal probability of being the target (i.e. being highlighted in the study below), then the reaction time reduces to: $T_R = c + d\log_2(N)$. As per (3), Hick-Hyman law supports the intuition that the required response (reaction) time T_R for the pointing-selection task declines as the user becomes more familiar with the interface; familiarity here implies that the probability p of a given stimulus/item being chosen as a target (e.g. highlighted by the interface) is higher (i.e. more likely to be known to the user) and consequently the $\log_2(1/p)$ component in (3) decreases.

Since the period required to accomplish a pointing-selection task, in general, comprises of the response period T_R in (3) and the duration of the pointing action T_P in (2), Hick-Hyman and Fitts' laws can be combined to predict the overall duration T of the target acquisition [20, 27], e.g. $T = T_R + T_P$. In [20], these two laws were combined and the crucial learning or expertise factor was incorporated to model the average overall time of acquiring an interface component.

3 Method

In this section, we first outline the experimental set-up and the data collected from the employed sensors. The graphical user interface for the secondary task of interacting with an in-vehicle display is then described. The participants' information and the considered measures are also listed.

3.1 Experimental Set-up and Logged Data

The study is conducted in a fixed mid-fidelity simulator at Jaguar Land Rover laboratories where a projector is used to display the road for the driving task.

- *Driving Tasks:* Lane-change-test driving simulation according to ISO 26022 is utilized for the primary task of driving and various driving performance related metrics (e.g. lane deviation, speed, etc.) are recorded. Road signs, separated by unequal distances, instruct users to change between the three lanes. No other road users are present and the closed circuit has bends of varying radii.
- *Display:* an 11.5″ Windows touch-screen is mounted to the dashboard in a portrait orientation; it displays an experimental GUI, which is discussed in more details below. All interactions with the touch-screen, e.g. mouse-up events, are logged with their respective time stamps.
- *Audible cue:* The system produces an audible cue (namely, a short and sharp beep sound signal) to indicate the start of a new target acquisition task. Thus, it informs the driver than a new stimulus is present on the touch-screen. Let t_{Sel} be the time instant a participant successfully selects the target on-screen icon. The next audible cue (i.e. a new pointing task) will be produced-presented after T_d seconds from the completion of the last selection at t_{Sel} ; this delay is chosen randomly such that $T_d \in \mathbb{D}$ and $\mathbb{D} = \{2, 3, 4, 5, 6\}$ seconds.

- *Switch for Pointing Gestures:* The steering wheel is instrumented with a switch to detect when the user hand leaves the wheel, which signals the start of the free hand pointing gesture, aimed at acquiring the highlighted on-screen icon.
- *Eye-gaze tracker:* a Smart Eye system [28] with two cameras is employed to continuously track the eye-gaze movements throughout the conducted experiments and according to a predefined "*physical-world*" model of the simulator interior and exterior. This model facilitates mapping the 3D environment into a 2D space. Consequently, the tracker provides the Cartesian coordinates of the gaze location in 2D in mm, with respect to a given reference point. The utilized system can reliably track the eye-gaze in the entire region of interest, including the road ahead and the touch-screen on the dashboard, at a rate of approximately 60 Hz.

3.2 Graphical User Interface

The displayed experimental interface has $N = 12$ square shaped selectable icons as depicted in Fig. 1 such that:

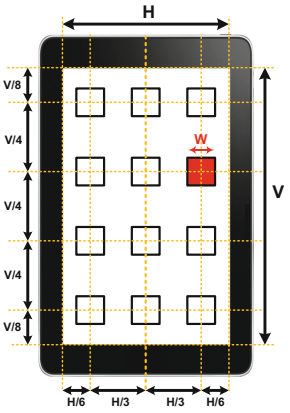


Fig. 1. GUI displayed on the touch-screen with $N = 12$ selectable icons. The target of the current pointing-selection task is highlighted in red; icon sizes change from one task to the next.

– The N icons are always simultaneously present; they are distributed across the touch-screen in a square formation, where H and V in Fig. 1 are the width and length of the screen in mm, respectively.
– At any point in time, all of the selectable GUI icons have the same size/width denoted by W; however W can change from one selection task to the next.
– The positions of the centers of all of the GUI icons are fixed (as per Fig. 1) and do not change during trials.
– During a pointing-selection task, only one icon (out of 12) is highlighted in red. This GUI item is chosen randomly such that all of the N interface items are

highlighted at least once every 12 selections. All of the experimental GUI icons have an equal probability of being highlighted. This set-up resembles the ISO 9241 pointing task, often referred to as the Fitt's law task.
– The width of the 12 GUI icons can have any of the values $W \in \mathbb{W}$, such that $\mathbb{W} = \{15, 20, 25, 30, 35, 40\}$ mm representing a range of typical sizes used in automotive interfaces [29].

3.3 Participants

Ten participants (3 females), ranging in age from 22 to 60 years old (mean is 34.6 and standard deviation is 13.4), are recruited from Jaguar Land Rover and Cambridge University staff. All participants had a prior experience in using a touch-screen (1–20 years, mean is 5.6 and standard deviation is 5.8). All participants are right handed and had driving licenses for 4–40 years (mean is 13.9 and standard deviation is 12.2). Their annual mileage ranges from 1,000–20,000 (mean is 10333.3 and standard deviation is 6500). None of them had any known impeding motor or visual impairment; two were wearing corrective glasses.

3.4 Driving Conditions and Procedure

In this study, two driving conditions are considered: (1) the driver follows the lane-changing signs, and (2) the driver does not have to change lanes, i.e. free driving in the circuit. The former represents a relatively high workload scenario compared with the free driving. This is more tangible with the steering wheel not being accurately calibrated per participant, rendering the driving task more demanding as noted by most participants. The order of the presentation of these two conditions to participants is counterbalanced.

After signing the consent form and being briefed about the experiment, each participant first undertook a training session; mainly to be acquainted with the lane changing instructions and how to perform the secondary task of on-screen target acquisitions. The trials start only after a participant confirms that he/she is familiar with the set-up and ready to proceed.

Whilst a participant continuously performs the primary task of driving, he/she undertakes the total of 72 GUI target acquisitions (pointing and selection) per each of the two considered driving conditions; breaks are introduced whenever necessary to mitigate the effects of fatigue. Each pointing-selection task commences as follows:

1. an audible cue, namely a short ping sound signal, is produced by the system to inform the driver that a new stimulus is present on the touch-screen, i.e. an icon is highlighted on the display,
2. the driver is ought to select the highlighted GUI item as soon as possible via a free hand pointing gesture, without compromising his/her driving performance,
3. after the participant successfully selects the on-display target, no GUI is highlighted until the initiation of the next pointing-selection task, i.e. at/after the next audible cue.

It is noted that participants were instructed to place their right hand on the steering wheel at all times, in particular over the area instrumented with a switch, except when they are undertaking the pointing gestures.

3.5 Considered Measures

For the n^{th} pointing-selection task, where W_n denotes the width of all displayed GUI icons, Table 1 lists the notation for relevant time instants. Here, we consider the following measures:

- ***Motor Response Time:*** $T_{M,n}$ (in seconds) is the delay between the audible cue and the start of the pointing gesture, i.e. $T_{M,n} = t_{\text{P0},n} - t_{A,n}$.

Table 1. Timestamps (in seconds) of various events related to the n^{th} pointing-selection task.

Notation	Description
$t_{A,n}$	Time the audio cue is produced
$t_{GD,n}$	Time the first eye-fixation on the touchscreen, after $t_{A,n}$, is detected by tracker
$t_{\text{P0},n}$	Start time of the pointing gesture, i.e. hand-off the steering wheel
$t_{\text{Sel},n}$	Time the user selects the highlighted on-screen icon

- ***Visual Response Time:*** $T_{V,n}$ (in seconds) is the delay between the audible cue and the user shifting his/her visual attention to the pointing task, $T_{V,n} = t_{GD,n} - t_{A,n}$.
- ***Visual and Motor Misalignment:*** $T_{VM,n}$ (in seconds) is the delay between the start of the pointing gesture and the first glance on the display, after the audible cue, thus, $T_{VM,n} = t_{\text{P0},n} - t_{GD,n}$.

4 Results and Discussion

To illustrate the sequence of events in a given experiment, Fig. 2 shows the eye-gaze behavior during four pointing-selection actions. It can be noticed that during these tasks the driver's eye-gaze is fixed on the touchscreen throughout the target acquisition engagement, i.e. the pointing-selection task is accomplished in one glance. Whilst this is a commonly observed behavior, it is not the only strategy adopted by the drivers who took part in the presented study, for instance two or multiple glances before and during the pointing gesture were also observed. This aspect will be the subject of further data analysis and experimental work in the near future. It can also be seen in Fig. 2 that the driver does not always look *first* at the display (e.g. to locate the on-screen target) prior to starting the free hand pointing gesture. This aspect is addressed in more details below.

Figures 3 and 4 depict the normalized histogram for the visual response time ($T_{V,n}$) and motor response time ($T_{M,n}$) under the two studied driving conditions, i.e. free driving to represent a low-effort driving scenario and lane changing task, which

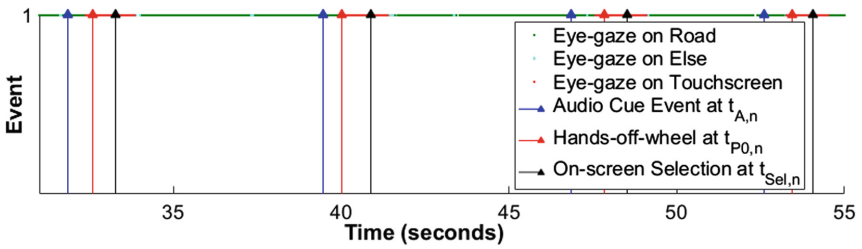
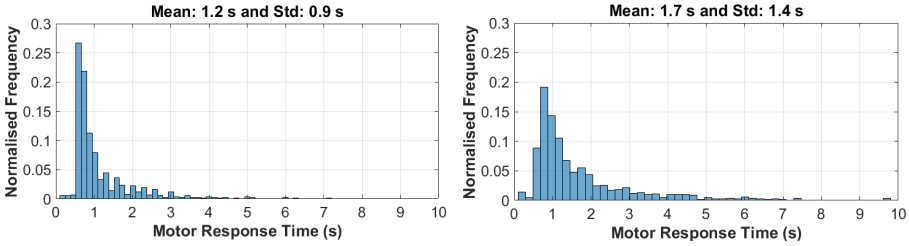


Fig. 2. Timeline of selected events from four pointing-selection tasks undertaken by one of the participating drivers, in the studied divided attention scenario.

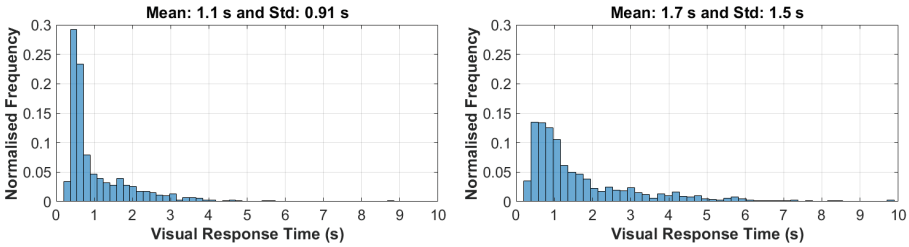
imposes "relatively" higher demand on the driver. It can be noticed from these two figures that the increase in driving effort/workload leads to substantial increases in the:

(a) average visual and motor response times (see the means in Figs. 3 and 4), and



(a) Free driving **(b)** Lane changing task

Fig. 3. Histogram of the motor response time $T_{M,n}$ for all drivers under two driving conditions. (a) Mean: 1.2 s and standard deviation: 0.9 s. (b) Mean: 1.7 s and standard deviation: 1.4 s.



(a) Free driving **(b)** Lane changing task

Fig. 4. Histogram of the visual response time $T_{V,n}$ for all drivers under two driving conditions. (a) Mean: 1.1 s and standard deviation: 0.91 s. (b) Mean: 1.7 s and standard deviation: 1.5 s.

(b) variability in the measured response times as captured by the displayed variances and/or the notable spread of the distributions in Figs. 3 and 4.

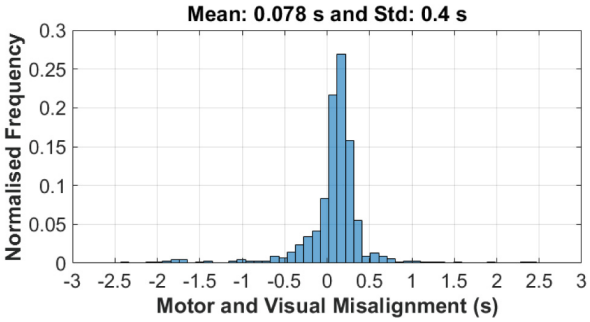
Therefore, the demands of the primary task of driving significantly impacts the driver's visual and motor behavior during the secondary task of acquiring interface items.

Another key observation is that the distribution of $T_{V,n}$ and $T_{M,n}$ are one sided and heavily skewed. The latter is more visible in the lane-changing driving task. It is noted that all of the GUI has an equal probability of being highlighted (i.e. being the stimulus). It is also intuitive to assume that a parallel visual search will be applied by the driver to locate the on-screen target. Subsequently and according to Hick-Hyman law or its various adaptations, the response time, which is deterministic in nature as per (3), is expected to be a fixed value, possibly with small variations. Whereas, the data shown in Figs. 3 and 4 indicate that the effective fusion of measurements from in-vehicle gesture and eye-gaze trackers should adequately take into account the significant variabilities in the drivers' responses-behaviors in the considered divided attention scenario. This can be captured by adopting a statistical model that adequately represents the randomness in the response times.

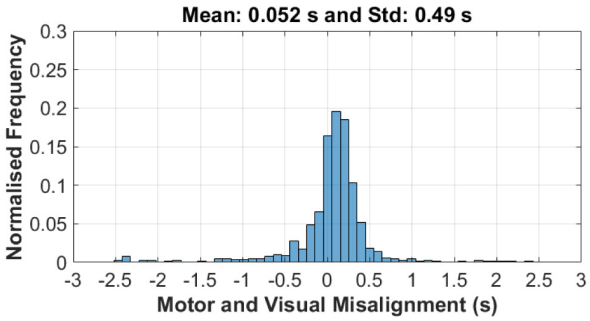
On the other hand, Fig. 5 displays the misalignment between the visual and motor attention/response, i.e. $T_{VM,n} = t_{\mathrm{P0},n} - t_{GD,n}$, for the secondary task of interacting with an in-vehicle display, whilst driving. It can be noticed that negative $T_{VM,n}$ values are prevalent. This implies that for a large number of pointing-selection tasks, the driver starts the pointing gesture *before* glancing at the screen. This is attributed to the driver's prior knowledge of the location of the display on the dashboard and his/her visual attention shifts to the pointing-selection task at a later stage to guide the free hand pointing gesture movements (e.g. to locate the highlighted icon and possibly fine-tune the required motor action). Nevertheless, it can be seen from the figure that the average (mean) delay between the onset of the visual and motor actions is positive, i.e. on average drivers first look at the screen before pointing. Interestingly, the distributions of $T_{VM,n}$ are less one sided and skewed (distribution in Fig. 5b is nearly symmetric) compared with those of $T_{V,n}$ and $T_{M,n}$; they are also centered close to zero.

Additionally, Fig. 5 demonstrates that the increase in the demand/workload of the performed driving task has less visible effect on the average misalignment between the visual and motor attention for the pointing-selection tasks. However, the variability of $T_{VM,n}$ under the lane changing condition is more visible as per Fig. 5b. Overall, Fig. 5 illustrates that statistically modelling this misalignment, which is intrinsically important to relating the gesture and eye-gaze tracker observations to the current pointing-selection task, is a promising approach.

Finally, it is noted that the pointing time as per Fitts' law is not discussed here as the observed $t_{P,n} = t_{\mathrm{Sel},n} - t_{\mathrm{P0},n}$ indicated a consistent conforming motor behavior (i.e. between the user starting the pointing gesture and selecting the on-screen target).



(a) Free driving.



(b) Lane changing task.

Fig. 5. Histogram (normalized) of the misalignment between the motor and visual attention $T_{VM,n}$ (in seconds) for the all drivers under two driving conditions. (a) Mean: 0.078 s and standard deviation: 0.4 s. (b) Mean: 0.052 s and standard deviation: 0.49 s.

5 Conclusions

This preliminary study shows that the demands imposed by the primary task of driving substantially impact the motor and visual behavior of a driver during the secondary task of interacting with an in-vehicle display. It gives an insight into the relationship between the gesture and eye-gaze movements during this divided attention scenario, which is fundamental to developing effective strategies to fuse pointing gesture and eye-gaze measurements in a multimodal automotive interactive display technology. Future work will address the underlying motor and visual behavior under a multiple hypothesis testing procedure.

References

1. Burnett, G.E., Porter, J.M.: Ubiquitous computing within cars: designing controls for non-visual use. Int. J. Hum Comput Stud. **55**(4), 521–531 (2001)
2. Harvey, C., Stanton, N.A.: Usability Evaluation for In-Vehicle Systems. CRC Press, Boca Raton (2013)
3. Burnett, G., Lawson, G., Millen, L., Pickering, C.: Designing touchpad user-interfaces for vehicles: which tasks are most suitable? Behav. Inf. Technol. **30**(3), 403–414 (2011)
4. Jaeger, M.G., Skov, M.B., Thomassen, N.G.: You can touch, but you can't look: interacting with in-vehicle systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1139–1148, April (2008)
5. Bark, K., Tran, C., Fujimura, K., Ng-Thow-Hing, V.: Personal Navi: benefits of an augmented reality navigational aid using a See-Thru 3D Volumetric HUD. In: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 1–8, September (2014)
6. Broy, N., Guo, M., Schneegass, S., Pfleging, B., Alt, F.: Introducing novel technologies in the car: conducting a real-world study to test 3D dashboards. In: Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 179–186, September (2015)
7. Jaguar Land Rover: XE in-car technology. http://www.jaguar.co.uk/jaguar-range/xe/features. Accessed 18 Mar 2016
8. Ahmad, B.I., Langdon, P.M., Godsill, S.J., Hardy, R., Skrypchuk, L., Donkor, R.: September. Touchscreen usability and input performance in vehicles under different road conditions: an evaluative study. In: Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 47–54 (2015)
9. NHTSA: Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. NHTSA-2010-0053 (2013)
10. Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J.: The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic driving study data. Natl Highw. Traffic Saf. Adm. HS **810**, 5942006 (2006)
11. Ohn-Bar, E., Trivedi, M.M.: Hand gesture recognition in real time for automotive interfaces: a multimodal vision-based approach and evaluations. IEEE Trans. Intell. Transp. Syst. **15**(6), 2368–2377 (2014)
12. Doshi, A., Trivedi, M.M.: On the roles of eye gaze and head dynamics in predicting driver's intent to change lanes. IEEE Trans. Intell. Transp. Syst. **10**(3), 453–462 (2009)
13. Zheng, R., Nakano, K., Ishiko, H., Hagita, K., Kihira, M., Yokozeki, T.: Eye-Gaze tracking analysis of driver behavior while interacting with navigation systems in an urban area. IEEE Tran. Hum. Mach. Syst. **46**(4), 546–556 (2016)
14. Ahmad, B.I., Murphy, J.K., Langdon, P.M., Godsill, S.J., Hardy, R., Skrypchuk, L.: Intent inference for hand pointing gesture-based interactions in vehicles. IEEE Trans. Cybern. **46** (4), 878–889 (2016)
15. Ahmad, B.I., Langdon, P.M., Godsill, S.J., Donkor, R., Wilde, R., Skrypchuk, L.: You do not have to touch to select: a study on predictive in-car touchscreen with mid-air selection. In: Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 113–120, October (2016)
16. Ahmad, B.I., Murphy, J.K., Godsill, S.J., Langdon, P.M., Hardy, R.: Intelligent interactive displays in vehicles with intent prediction: a Bayesian framework. IEEE Signal Process. Mag. **43**(2), 82–94 (2017)
17. Wilson, J.R., Sharples, S. (eds.): Evaluation of Human Work. CRC Press, London (2015)

18. Biswas, P., Langdon, P.: Multimodal intelligent eye-gaze tracking system. Int. J. Hum. Comput. Interact. **31**(4), 277–294 (2015)
19. Stellmach, S., Sellen, A., Blake, A.: The costs and benefits of combining gaze and hand gestures for remote interaction. In: Human-Computer Interaction, pp. 570–577. Springer International Publishing, Cham
20. Cockburn, A., Ahlström, D., Gutwin, C.: Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. Int. J. Hum Comput Stud. **70**(3), 218–233 (2012)
21. Fitts, P.M., Peterson, J.R.: Information capacity of discrete motor responses. J. Exp. Psychol. **67**(2), 103 (1964)
22. McGuffin, M.J., Balakrishnan, R.: Fitts' law and expanding targets: Experimental studies and designs for user interfaces. ACM Trans. Comput. Hum. Interact. (TOCHI) **12**(4), 388–422 (2005)
23. Soukoreff, R.W., MacKenzie, I.S.: Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. Int. J. Hum Comput Stud. **61**(6), 751–789 (2004)
24. Large, D.R., Crundall, E., Burnett, G., Skrypchuk, L.: Predicting the visual demand of finger-touch pointing tasks in a driving context. In: Proceedings of the International Conferenece on Automotive User Interfaces and Interactive Vehicular Applications, pp. 221–224, September (2015)
25. Hick, W.E.: On the rate of gain of information. Q. J. Exp. Psychol. **4**(1), 11–26 (1952)
26. Hyman, R.: Stimulus information as a determinant of reaction time. J. Exp. Psychol. **45**(3), 188 (1953)
27. Hoffmann, E.R., Lim, J.T.: Concurrent manual-decision tasks. Ergonomics **40**(3), 293–318 (1997)
28. Smart Eye: http://www.smarteye.se. Accessed 10 Mar 2017
29. Kim, H., Kwon, S., Heo, J., Lee, H., Chung, M.K.: The effect of touch-key size on the usability of In-Vehicle Information Systems and driving safety during simulated driving. Appl. Ergon. **45**(3), 379–388 (2014)

Determination of Auditory Stimuli for the Auditory Variant of the Detection Response Task Method

Kristina Stojmenova(✉) and Jaka Sodnik

Faculty of Electrical Engineering, University of Ljubljana,
Trzaska Cesta 25, 1000 Ljubljana, Slovenia
{kristina.stojmenova,jaka.sodnik}@fe.uni-lj.si

Abstract. The Detection Response Task (DRT) is a method for measuring the attentional effects of secondary tasks on driver's cognitive load and was standardized by ISO in 2016 [1]. It observes response times and stimuli detection performance (hit rates) to different stimuli types – visual, tactile and auditory – as indirect indicators of changes in cognitive load. ISO 17488 [1] standard provides instructions for technical implementation and measurement methods only for the visual and tactile versions (use of visual and tactile stimuli) of the DRT method. In this study, we try to find the most appropriate auditory stimulus for technical implementation also for the auditory version of the method. This paper presents results of a user study in which responses to different auditory DRT stimuli were compared while inducing users' cognitive load with a modified *n* – back task, also known as Delayed Digit Recall task [2].

Keywords: Auditory stimuli · Detection Response Task · Cognitive load · Delayed Digit Recall Task

1 Introduction

The Detection Response Task (DRT) is a method for assessing attentional effects of secondary tasks on driver's cognitive load [1]. In this method, the driver is asked to respond to either visual, tactile or auditory stimuli while driving and performing various secondary tasks (e.g. typical interaction with in-vehicle information systems - IVIS). Visual stimuli are presented in a form of a red dot and can be placed on the left side of the steering wheel on the car's windshield (Remote DRT) or can be placed on the driver's head in front of his left eye (Head-mounted DRT). Tactile stimuli are presented with a small vibration motor placed on the driver's collarbone on his/her left shoulder. The stimuli can last from 1 to 2 s and are presented to the driver in random 2–5 s time intervals. The driver is asked to respond to the stimuli as soon as possible by pressing a button attached to his/her left hand index finger (the driver presses the button against the steering wheel). Response times and hit rates are interpreted as the indicators of the attentional affect of secondary tasks on driver's cognitive load. Response times are measured simply as response times to stimuli while stimuli detection performance, described as hit rate is calculated as percentage of correct responses (occurring from 100 ms to 2500 ms after introduction of stimulus) out of all presented stimuli.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_43

Although the current version of the ISO 17488 standard mentions the possibility of using auditory stimuli, it provides guidelines for technical implementation only for visual and tactile presentation of DRT stimuli. It includes instructions on how to gather and interpret data collected with this method, and provides instructions on how to choose among the DRT variants depending on the type of secondary task used in the study. It is suggested that all DRT variants are suitable for evaluating the attentional effects on driver's cognitive load of auditory-vocal tasks (e.g. *n* – back task). However, for more realistic multi-modal tasks, which require also visual attention (e.g. voice operated IVIS, where content is presented visually), the tactile DRT should be preferably used. In such cases the tactile DRT is the best option since it avoids visual sensory/perceptual interference. For tasks that require manual interaction, a motor interference is present for all DRT variants as they all use the same response method.

However, most of in-vehicle infotainment systems today present information visually, and are operated manually. When using visual DRT, regardless of the stimuli location, the same perceptual output channel is used for perceiving both DRT stimuli and secondary tasks information. Similarly, when using the tactile DRT in combination with visual-manual tasks, the same input modality is used for answering to DRT stimuli and for performing secondary tasks. The use of tactile warning and information systems (e.g. seat/steering wheel vibrations for lane deviations or line changing) in today's vehicles is gaining popularity, suggesting the use of tactile version to be less appropriate for assessing the attentional effects to driver's cognitive load also for this kind of tasks. Warning systems on the other hand are mostly auditory, due to the fact that sound alerts usually implicate emergency/danger and consequently evoke fast responses, on the contrast to e.g. the tactile sense, which is known to be the slowest. These observations have inspired us to explore the use and possible positive effects of the auditory version of the DRT method. We believe this variant could expand the method's usability (because it doesn't interfere with the input/output channels of visual-manual secondary tasks) and could be a valuable addition to the published ISO standard. In particularly, we would like to explore the attentional effects of the auditory DRT method on cognitive load of auditory-vocal cognitive tasks. In this paper, we therefore try to find the most suitable and effective stimulus also for the auditory DRT variant, providing the same level of sensitivity to changes in cognitive load as the other variants presented in the ISO standard.

2 Related Work

Auditory signals can be described in two ways, by their intensity, and by their frequency. Psychoacoustic researches have shown a strong relationship between response time, auditory stimuli intensity and frequency. More than a century old researches from Cattell [3], Pieron [4] and Chocholle [5], have shown that response time decreases with the increase of auditory stimuli intensity. In a simple but rather extensive research study, Chocholle [5] compared response times for different intensity levels and frequencies of auditory stimuli. His results showed that response times systematically decreased with

the corresponding increase in signal intensity, while stimuli with equal loudness produced equal response times regardless of the stimuli frequency.

The human hearing system is sensitive only to stimuli containing frequencies from 20 Hz to around 20 kHz, and is most sensitive to frequencies between 2 kHz and 5 kHz [6]. The upper limit of the hearing system decreases with age because after a certain age people cannot hear high frequencies. This is due to the fact that frequencies play a significant role in perceived loudness. Fletcher and Munson first introduced a definition of the term loudness and proposed an empirical formula for determining the level of loudness based on the intensity and frequency of the signal's components [7]. The formula is based on equal loudness curves obtained by comparing loudness levels of many sounds (pure tones and complex signals). Their measurements were later repeated by many research groups, which led to some modifications in the original curves. The latest revision was made in 2003 [8], and it is shown in Fig. 2. Regardless of the loudness, a local minimum for all of the curves can be seen between 2 kHz and 5 kHz. These frequencies also represent the most sensitive audible range, and therefore most interesting to be explored.

Stojmenova and Sodnik [9] performed a study with a similar purpose to the presented study in this paper earlier in 2016. In their study, the authors observed the possible differences in sensitivity to changes of cognitive load of different auditory signals: single-frequency signals (pure tones), multi-frequency signals (consisting of a sum of three frequencies), and noise based signals. The results suggest that the use of white noise resulted in fastest absolute response times, however, smallest difference in response times were found for this type of signal when comparing trials with and without cognitive task. However, this difference in response time is the most important indicator of increased cognitive load. Bigger the difference, higher the sensitivity of the method to changes in cognitive load. Therefore white noise stimuli could not be used as DRT stimuli. They also observed a noise band 1 kHz to 4 kHz, which showed better results, however the participants reported noise stimuli to be very unpleasant to listen to and also often masked with environment sounds in the simulator (engine sound, surrounding traffic, etc.). The authors found biggest differences in response times for single-frequency stimuli. Participants also reported that single-frequency auditory signals were more pleasant to listen to for a longer period of time.

These results inspired us to take a closer look primarily at single-frequency stimuli due to their high sensitivity to differences in cognitive load and positive feedback from the participants. We focused on 4 kHz signal and its neighbourhood based on the previously presented results and because it is in the range of most sensitive frequencies according to the Equal-loudness curves (Fig. 1).

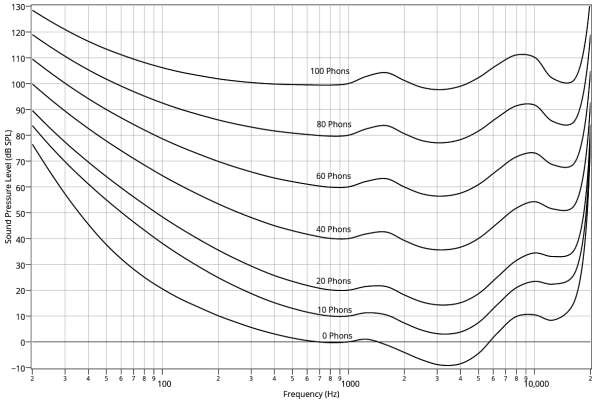


Fig. 1. Equal – loudness curves [8].

3 Methods

A surrogate driving experimental set-up was used in this study. This kind of experimental set up was used in order to reduce visual-manual distractions and to try to reduce any additional attentional distractions that may occur as a result of operating a real vehicle or drive in a simulator. Participants were therefore asked to watch a video of a real world driving on a highway in Germany, recorded from a driver's perspective. The video was played on a curved 48′′ Samsung LED TV while the participants set in racing car seat with a Fanatec steering wheel and pedals attached to the construction to recreate a feeling of driving. A simple Java application, played on a Mac Book Pro, was developed to generate and play auditory DRT stimuli and also to record participant's response times. Participants responded to the stimulus by single click on an Apple Magic Mouse held in their right hand. The mouse was placed on a table on the right side of the driver, just as if the driver would place its arm on an armrest. The auditory DRT stimuli were played through Sennheiser HDR 180 headphones with 16-bit/44.1-kHz resolution system, in order to avoid acoustic delays and to reduce nonlinear distortion [6]. The surrounding sounds from the driving video presented to the drivers were played trough the same Samsung TV used for presenting the visuals.

Each participant performed 12 trials – two trials with each of the six tested auditory signals. For every sound signal they would perform one trial where they would watch the video and answer to auditory DRT stimuli, and one trial where they had to watch the video and perform two tasks simultaneously: answer to DRT stimuli and perform an additional cognitive task. Participants were instructed to keep their eyes on the road at all times, and to try to answer to the DRT stimuli promptly and perform the cognitive task as accurately as possible.

3.1 Cognitive Task

The cognitive task used in the experiment is the Delayed Digit Recall task [2, 10], where participants are presented a random single digit number and have to answer by recalling number presented moments earlier. There are three possible levels: 0, 1 and 2 – back. 0 – back asks the driver to recall the last presented number, and 2 – back to recall 2 numbers before the last one. In a NHSA paper by the US Department of Transportation, the 2 – back task level is suggested to be starting point for setting a limit for acceptable amount of cognitive distraction [11]. In order to impose the highest acceptable cognitive load, the 2 – back task was used in this study. Participants were presented single digit numbers vocally by the experimenter every 5 s and were asked to repeat the second last number (i.e. 2 – back task), as illustrated in Table 1.

Table 1. Consequential order of presentation numbers in the cognitive 2 – back task

Order of								
Presented numbers	3	4	5	5	4	1	9	2
Answered numbers	No answer	No answer	3	4	5	5	4	1

3.2 Auditory DRT Stimuli

6 different auditory DRT stimuli were tested in this experiment: 5 single frequency and 1 multi-frequency signal:

- 2 kHz pure tone,
- 3 kHz pure tone
- 4 kHz pure tone,
- 5 kHz pure tone,
- 8 kHz pure tone, and
- sum of pure tones (2 kHz + 4 kHz + 8 kHz).

The duration of the stimuli was always one second. We played all of the signals in all of the trials the at same dB intensity. We did not play every sound at same loudness, which is a subjective level of perceived intensity for each individual, because according to Chocholle [5], we would get same duration of response times regardless of the change in frequencies.

3.3 Design

This user study was conducted in a within subject counterbalanced measures design. 24 participants (9 female) aged 21–33 years old, all with valid driving licenses, participated in the study. Upon arrival, they were informed about the purpose of the study, and were introduced to the DRT method and the *n* – back task. Each participant first practiced performing the *n* – back task and the DRT task separately, and than simultaneously to get to know each of the tasks and to reduce the novelty effect.

After the introduction part, participants performed all 12 trials with 30-second pause between each trial. One trial lasted approximately 1 min, in which participants were presented with 12 DRT stimuli.

4 Results

Shapiro-Wilk normality test showed that the data was not normally distributed, and a Levene's test showed the data is not homogeneous. We therefore used non-parametric tests to explore the data. A Friedman test showed there was a statistically significant difference in response times for response times without and with the cognitive *n* – back task, $\chi^2(1) = 1149$, $p < 0.01$ (Fig. 2.). Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.025$. There was statistically significant increase in response times for all of the observed sound stimuli ($p < 0.01$).

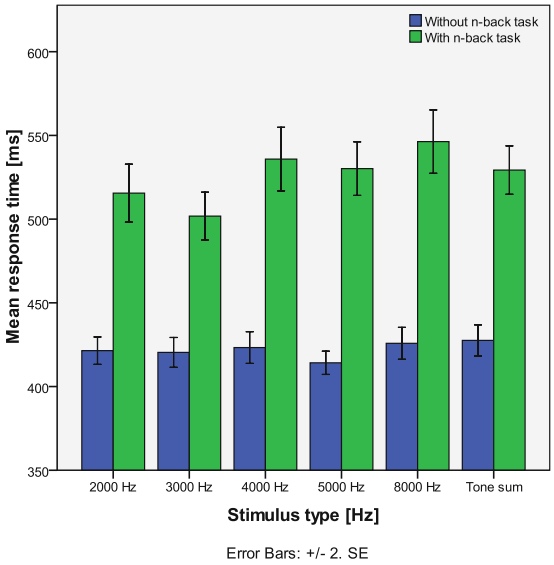


Fig. 2. Mean response times for each signal for trials performed without and with the 2 – back task.

Figure 3 presents only the differences in mean response times (i.e. increase in response times) between trials without and with a cognitive task. Post hoc analysis with Wilcoxon signed-rank tests showed highest statistically significant differences ($p < 0.025$) in response times for 3 kHz tone compared to the 4 kHz tone, 5 kHz, 8 kHz tone and the tone sum, and for the 2 kHz tone compared to the 8 kHz tone.

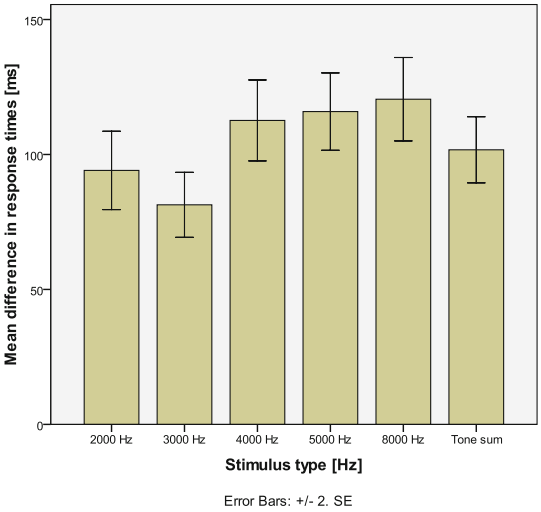


Fig. 3. Mean difference in response times for each sound signal between trials performed without and with the additional secondary 2 – back task.

Results of hit rates and secondary task performance are not presented graphically, because in all trials the miss rate score was less than 5%, and the results of performance rate of secondary tasks did not show any significant differences.

5 Discussion

Greatest differences in response times were found for auditory stimuli of 4 kHz, 5 kHz and 8 kHz. Similarly to Stojmenova and Sodnik [9], using a 4 kHz auditory signal as DRT stimulus resulted in significant changes in response times for trials without and with a cognitive task. These results suggest that the 4 kHz signal could be a strong candidate as an appropriate stimulus for the auditory version of DRT method. Similar differences in response times were found also for 5 kHz and 8 kHz auditory signals. As 4 kHz, 5 kHz is also in the most sensitive audible area these results are not surprising. It is interesting though that 3 kHz and 2 kHz, which are also in this area, showed significantly smaller differences in results 8 kHz on the other hand, which is out of this range showed similar results to 4 kHz and 5 kHz. We can speculate that these results are due to the perceived loudness of the stimuli. Even when played at same intensity levels, higher frequencies are perceived as louder, which may have influenced the processing speed and therefore evoked faster response times. Niessen [12] found out that increased intensity of stimulus speeds up the processing course of action and in that way affects alertness and the time needed to direct attention to a stimulus. This effect might have influenced also our results. The sum of pure tones showed poor results, similar to the 2 kHz signal, which also make it less appropriate candidate for DRT stimulus.

6 Conclusion

Consistent significant differences in response times for trials with a cognitive task compared to trials without a cognitive task prove that the auditory DRT variant is also sensitive to changes in cognitive load. This is in line with the results of Stojmenova et al. [13] in which the authors compared all three variants of the DRT method. The mean differences in response time for 4 kHz, 5 kHz and 8 kHz signals are approximately 110 ms, which is more than 25% of absolute response time when comparing trials with and without a cognitive task. If we take into consideration Nissen's findings [12], we could probably increase this difference by playing sound stimuli at higher intensity, but only up to a point until it does not mask the environment or becomes unpleasant and disturbing. We would however suggest to exclude the pure tone of 8 kHz from further research due to the fact that upper limit of hearing decreases with age. It is therefore reasonable to choose a sound with lower frequency like 4 kHz or 5 kHz since both single frequency signals showed similar differences in response times. In our future work, we will test the 4 kHz and 5 kHz also with visual-manual tasks, and observe the sensitivity of auditory DRT also for this kind of tasks. In the future, we want to compare also response times of the auditory and tactile DRT variants together with visual-manual and cognitive auditory-vocal tasks.

Acknowledgements. This study was funded by Slovenian Research Agency.

References

1. Road vehicles – Transport information and control systems – Detection-Response Task (DRT) for assessing attentional effects of cognitive load in driving. ISO/DIS 17488
2. Mehler, B., Reimer, B., Dusek, J.A.: MIT AgeLab Delayed Digit Recall Task (n-back). MIT AgeLab white paper 2011-3B (2011)
3. Cattell, J.M.: The influence of the intensity of the stimulus on the length of the reaction time. Brain **8**, 512–515 (1886)
4. Pieron, H.: NouvelIes recherches sur l'analyse du temps de latence sensorielle et sur la loi qui relie le temps al'intensite d'excitation. Annee Psychologique **21**, 58–142 (1920)
5. Chocholle, R.: Variation des temps de reaction auditif en function de l'intensite it diverses frequences. Annee Psychologique **41**, 6S–124 (1940)
6. Fastl, H., Zwicker, E.: Psychoacoustics: facts and models, vol. 22, pp. 8–22. Springer, Berlin (2006)
7. Fletcher, H., Munson, W.A.: Loudness, its definition, measurement and calculation. Bell Syst. Tech. J. **12**(4), 377–430 (1933)
8. Acoustics – normal equal-loudness-level contours. ISO 226:2003
9. Stojmenova K., Sodnik, J.: Določitev zvočnega dražljaja za zvočno različico metode zaznavno-odzivne naloge. In: Zbornik petindvajsete mednarodne Elektrotehniške in računalniške konference ERK 2016, IEEE Region 8, part A, pp. 329–332
10. Mehler, B., Reimer, B., Coughlin, J.F.: Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: an on-road study across three age groups. Hum. Factors **54**(3), 396–412 (2012)

11. Ranney, T.A., Baldwin, G.H., Parmer, E., Domeyer, J., Martin, J., Mazzae, E.N.: Developing a test to measure distraction potential of in-vehicle information system tasks in production vehicles (No. HS-811 463) (2011)
12. Niessen, M.J.: Stimulus intensity and information processing. Percept. Psychophys. **22**, 338–352 (1977)
13. Stojmenova, K., Jakus, G., Sodnik, J.: Sensitivity evaluation of the visual, tactile, and auditory detection response task method while driving. Traffic inj. prev. **18**(4), 431–436 (2016)

Perceptions Towards the Use of Intelligent Transport System Technologies by Earthquake Victims

Izza Anwer(✉) and Susan Grant-Muller

Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, UK
Tsiam@leeds.ac.uk, s.m.grant-muller@its.leeds.ac.uk

Abstract. Islamabad (Pakistan) is a well-developed community that is experienced in natural disasters, particularly earthquakes. The problem faced by the people of Islamabad (Pakistan) is the lack of timely information immediately prior to, during and following earthquakes resulted in many causalities and other major losses in the past. Due to a lack of relevant literature, the perceptions of earthquake-affected people regarding the use of ITS technologies (that can facilitate broken communication and missing information links between different earthquake-affected people and earthquake management authorities) are unknown. In total 102 randomly selected households of Islamabad (Pakistan) were recruited to complete a questionnaire, which had been carefully designed and piloted. A Multinomial Logistic Regression approach was employed to analyse the data. In total 15 different statistically significant models were developed. The results will help in designing ITS oriented policies and management plans for the rescue, search, evacuation operations, and emergency logistics for the earthquake affected people.

Keywords: Disaster · Rescue · ITS technologies · ICT technologies · Transport systems · Earthquakes

1 Introduction

ITS technologies are deployed in a number of ways to manage traffic, logistics, disaster management and even in combining different data sources [1]. The role of ITS technologies is more important when it comes to planning and managing transport systems, and activities related to earthquakes such as evacuations and rescue and search operations [2]. We define ITS technologies as all those technologies that can be used in any capacity to support transport related activities, services and information [1]. ITS technologies also involve all types of Information and communication (ICT) as well as transport system technologies. The effective flow of transport information and communication can contribute towards sustaining the resilience of transport systems in critical situations [3, 4] such as earthquakes.

Earthquakes are unpredictable disasters and their management requires active planning approaches for transport-earthquake activities that are diverse, difficult and are stretched over the time period for the three stages i.e. pre, during and post- earthquake. The major problem encountered during earthquakes is that of broken communication

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_44

links, that cause hurdles in timely information flow [5]. An absence of transport-earthquake related communications on multiple platforms, amongst transport-earthquake management authorities and transport-earthquake affected people causes many causalities and losses. It can leave many people in misery that could otherwise have been located and rescued in time [6]. ITS technologies have unexplored potential benefits that need to be trialled and tested especially in disaster scenarios, and for different communities who are potential users of ITS technologies.

This research paper focuses on such aspects of information and communication flow through the use of ITS technologies that are available to, and usable by, earthquake affected communities for the three stages of an earthquake. The objective is to understand the depths of perceptions formed by the primary victims of earthquakes who used ITS technologies whenever they encountered (and may still encounter) earthquakes. These individuals have faced transport-earthquake problems resulting from a lack of exchange of transport-earthquake information. In addition, the influencing variables behind particular perceptions are also analysed. It is important to gain better knowledge of what types of ITS technologies are, or can be, utilized by earthquake victims, which technologies that they comfortable to use and under what circumstances. The main question however, is which technologies are operational under changing earthquake circumstances and what are the influencing variables that affect their use by individuals in earthquake prone communities. This leads to a further investigation of those variables that influence earthquake victims in making choices among different ITS technologies over the three stages of earthquakes.

This research paper further investigates the role of the ITS technologies that are used at a humanitarian level to manage transport-earthquake information [5] by the use of ICT technologies [3] especially in disasters. Different levels of participation can contribute to the different dimensions of communications such as decision making by government bodies in crisis and resource management. Transport oriented ICT technologies are in use to improve the transport system in a crisis by providing real time data to some extent. The research question addressed here will provide insights into the perceptions that a developed community of a developing country have with regard to the use of ITS technologies for the said purpose, as well as how information and communication systems work for an earthquake-affected community [5, 6].

This research paper is a part of an extended study [7] that aims to answer the main research question. One of the case studies is presented here i.e. the perceptions of the community of Islamabad. The residences in the vicinity of Margala towers is an area (in Islamabad) which was directly affected by earthquake (and aftershocks), es-pecially so from 2005 onwards. Data was collected from residents in this area, who have experienced earthquakes and aftershocks many times. The rationale behind choosing Islamabad as a case study was that it has been badly affected by earthquakes in the past, the infrastructure (road, traffic and transport, and information and communication system) has been almost completely demolished by some disasters. A lack of timely information and communication flow regarding transport-earthquake activities between authorities and victims meant that individuals were left in misery and unaware of the situation. This was despite all the available resources and the expertise to deploy those resources [8]. To date, no significant study of the case of Islamabad has been published concerning the exchange of transport-earthquake information through ITS technologies

between the different strata of community and the authorities [8], even though worldwide practices emphasize the need [9]. This paper assumes that any technology (or medium of information and communication, including Information and Communication Technologies and social media) that are, or can be, used for the purpose of facilitating decision making (based on available travel and traffic information for transport users and any transport related activity) should be termed as an ITS technology. Data collection to investigate the research questions is described in the next section.

2 Sampling Method and Data Collection

The earthquake victims and earthquake experienced community members were recruited as a sample to analyse the effect of different variables on choosing ITS technologies under different stages of earthquake. In total, data was collected from 102 household samples (exclusively earthquake-affected people). Households in the vicinity of Margala towers were randomly surveyed to collect data. A questionnaire was designed carefully to test 18 Independent variables (IDV) that could possibly influence choosing ITS technologies. The questionnaire was piloted and was ethically approved by the Ethical Approval Committee of University of Leeds. The quality of the data collected was given priority. It took around 3–4 months to collect data. The collected data was analysed through Multinomial Logistic Regression technique, explained in the next section.

3 Multinomial Logistic Regression (MNLR) Analysis approach

Multinomial logistic regression (MNLR) is considered as an advanced regression method, used to model relationships between DV and IDV. It is particularly used where DV have more than two discrete outcomes. A MNLR model is a generalized form of logistic regression and binomial regression [10, 27, 29], best suits the quantitative data having discrete multiple categories for DV as output [11, 12]. MNLR is not new to ITS system technologies [1, 13, 26]. However, the application of MNLR in the context of emerging transport system technologies in floods is yet to be investigated further. This research provides another opportunity [7] to the use of MNLR at humanitarian level to identify the extent, use, availability and choice of ITS technologies under certain circumstances. Along with many other literature resources, considerable guidance is taken from [14, 15] to understand MNLR technique better and interpretation of results effectively. For a DV with k categories, the probability that the i'th case falls in category k is given by π_{ik} where:

$$\pi_{ik} = \frac{e^{z_{ik}}}{e^{z_{i1}} + e^{z_{i2}} + \ldots . . e^{z_{ik}} + 1} \tag{1}$$

In this analysis, "K" varies for each DV e.g. categories of"simple ITS technologies" are do not know/none, landline/mobile phone, radio, TV and more than one. zik is the value of the k'th unobserved continuous variable for the i'th case and is assumed to be linearly related to the predictor variable.

$$z_{ik} = b_{k0} + b_{k1}x_{il} + \ldots b_{kj}z_{ij} \tag{2}$$

x^{ij} is the j'th predictor for the i'th case, bkj is the j'th coefficient for the k'th unobserved variable and J is the number of predictors.

Z^k is arbitrarily set to zero creating a reference case against which others are compared. Here the reference case for choosing specific ITS technology out of a set of ITS technologies under certain situations while considering 'do not know/none' as a reference category. The model does not require normality, linearity, and homogeneity of variance in the predictor variables; however it is based on assumption that the DVs with groups being discrete, non-overlapping and identifiable. Also, the assumptions considered in MNLR is that within DV categories the choice of one category is independent of the choice of the other category. Tables given below, show dependent and independent variables used to build models. Forward step-wise method of regression is used to build model.

4 Results

In total 18 IDVs were tested through MNLR, however, only 10 of them showed their effect on the DVs (can be observed in Tables 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15). ITS technologies were grouped into seven (see tables). The reference category for all DVs is "Do not know" whilst for each IDV it varies and is mentioned in the Tables 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15 as RC (reference category). Each DV and IDV have two or more than two categories, which is the main reason to choose MNLR as the analysis technique that it can easily deal with multiple categories (usually more than two). Categories of DVs are mentioned in each (Tables 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15) model's title separately. Categories of IDVs that are distinguishing the DV's relevant category with respect to reference category of DV (i.e. Do not know). 15 statistically significant models are presented in Tables 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15. Each table presents results from 102 samples. Marginal percentage criteria (MPC) are satisfied for each model. In each table, * = statistically significant at the 95% confidence level, and ** = statistically significant at the 90% confidence level. Also, "^" shows merged categories (of DV) due to lack of variation of responses from participants in certain categories. In the first column, independent variable (IDV) is presented and in the second column, DV's category that is distinguished with respect to the reference category of DV is presented. In the third column, the statistically significant category of IDV is presented with respect to reference category (RC) of IDV along with respective parameter. In the next three columns, "OR" stands for "Odds Ratio", "LCI" stands for "Lower confidence interval" and "UCI" stands for "Upper confidence interval".

Table 1. Model 1: Pre-Earthquake Islamabad. DV Simple (Conventional) Intelligent Transport System Technologies *(Do not know/None^, Landline/mobile phone, Radio, TV, All of above)*

IDV	DV's Sign. category (β)	IDV's Sign. category (β)	OR	LCI	UCI
Emergency/rescue Authorities.	(Landline/mobile phone)	Not to mildly prefer (−3.16)	.04*	.003	.53
		RC: Prefer to strongly prefer			
Dependency status	(Landline/mobile phone)	Alone (2.372)	10.7*	1.02	112
		RC: Not alone			

Table 2. Model 2: During-Earthquake Islamabad. DV Simple (Conventional) Intelligent Transport System Technologies *(Do not know/None^, Landline/mobile phone, Radio, TV, Loudspeaker, All of above)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Emergency/rescue Authorities.	(Landline/mobile phone)	Not to mildly prefer (-1.819)	.16**	.023	1.12
	TV	Not to mildly prefer (-2.277)	.06**	.103	1.05
		RC: Prefer to strongly prefer			

Table 3. Model 3: Pre-Earthquake Islamabad. DV Subscription Services *(Do not know/None^, Email Subscription Services/Newsletters/Brochures^, Mobile phone Subscription Services, All of above)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Vehicle ownership	(Email…brochures)	Yes (1.314)	3.7*	1.18	11.8
	(Mobile phone subscriptions)	Yes (4.291)	73*	7.8	682.3
	(All of above)	Yes (2.554)	13.*	2.84	58.3
		RC: No			
Police	(All of above)	Not to mildly prefer (−3.292)	.04*	.007	.206
		RC: Prefer to strongly prefer			

Table 4. Model 4: During-Earthquake Islamabad. DV Subscription Services *(Do not know/None^, Email Subscription Services/Newsletters/Brochures^, Mobile phone Subscription Services, All of above)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Gender	(Mobile ph. subscriptions)	Male (2.239)	9.4*	1.08	81.9
		RC: Female			
Police	(All of above)	Not to mildly prefer (−1.858)	.16*	.040	.602
		RC: Prefer to strongly prefer			

Table 5. Model 5: Post-Earthquake Islamabad. DV Subscription Services *(Do not know/None^, Email Subscription Services/Newsletters/Brochures^, Mobile phone Subscription Services, All of above)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Gender	(Mobile ph. subscriptions)	Male (1.5)	4.5**	.775	26.1
		RC: Female			

Table 6. Model 6: Pre-Earthquake Islamabad. DV Smart Phone Applications *(Do not know/None^, Contributory websites/Alerts/updates Apps.^, GPS trackers, All of above)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Police	(All of the above)	Not to mildly prefer (−1.55)	.21*	.066	.68
		RC: Prefer to strongly prefer			

Table 7. Model 7: During-Earthquake Islamabad. DV Smart Phone Applications *(Do not know/None^, Contributory websites/Alerts/updates Apps.^, GPS trackers, All of above)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Evacuation Vehicles	(Contributory/updates Apps)	Do not know/Not required or used (−2.762)	.06*	.005	.74
	(Contributory/updates Apps)	Boat/Helicopter or Airplane (−2.906)	.06*	.005	.60
		RC: All of above			
Police	(GPS trackers)	Not to mildly prefer (−1.516)	.22*	.049	.98
	(All of the above)	Not to mildly prefer (−2.133)	.12*	.03	.47
		RC: Prefer to strongly prefer			

Table 8. Model 8: Pre-Earthquake Islamabad. DV Social Media Public Forums *(Do not know/None^, Facebook, Twitter, Both)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Government Authorities	(Both)	Not to mildly prefer (−2.3)	.01*	.011	.83
		RC: Prefer to strongly prefer			

No statistically significant model for "post-earthquake simple ITS technologies", "pre, during and post-earthquake Messengers", "post-earthquake smart phone applications" and "post-earthquake social media websites" were developed.

Table 9. Model 9: During-Earthquake Islamabad. DV Social Media Public Forums *(Do not know/None^, Facebook, Twitter, Both)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Emergency/rescue Authorities.	(Facebook)	Not to mildly prefer (−1.941)	.14*	.043	.48
	(Both)	Not to mildly prefer (−1.445)	.24*	.071	.78
		RC: Prefer to strongly prefer			

Table 10. Model 10: Post-Earthquake Islamabad. DV Social Media Public Forums *(Do not know/None^, Facebook, Twitter, Both)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Education level	(Facebook)	School/college graduated (−2.71)	.07*	.006	.7
		RC: University grads./Professionals			

Table 11. Model 11: Pre-Earthquake Islamabad. DV Social Media Websites *(Do not know/None^, YouTube, Contributory/Information websites and blogs^, All of above)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Police	(YouTube)	Not to mildly prefer (2.1)	8.2*	1.52	44.
		RC: Prefer to strongly prefer			

Table 12. Model 12: During-Earthquake Islamabad. DV Social Media Websites *(Do not know/None^, YouTube, Contributory/Information websites and blogs^, All of above)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Experience (Earthquake)	(Contributory.. websites/blogs)	Yes (2.729)	15.*	1.68	140.
		RC: No			
Government Authority	(YouTube)	Not to mildly prefer (−2.816)	.06*	.006	.58
	(All of above)	Not to mildly prefer (−3.451)	.03*	.003	.34
		RC: Prefer to strongly prefer			

Table 13. Model 13: Pre-Earthquake Islamabad. DV News Channels *(Do not know/None^, Local news channels, International news channels, Both news channels)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Vehicle tech. (In-built)	(Both)	Yes (1.93)	6.9*	1.954	24.5
		RC: No			

Table 14. Model 14: During-Earthquake Islamabad. DV News Channels *(Do not know/None^, Local news channels, International news channels, Both news channels)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Emergency/rescue Authorities	(Both)	Not to mildly prefer (−1.455)	.23*	.071	.76
		RC: Prefer to strongly prefer			

Table 15. Model 15: Post-Earthquake Islamabad. DV News Channels *(Do not know/None^, Local news channels, International news channels, Both news channels)*

IDV	DV's category	IDV's Sign. (β)	OR	LCI	UCI
Emergency/rescue Authorities	(Both)	Not to mildly prefer (−2.431)	.09*	.020	.38
		RC: Prefer to strongly prefer			

5 Interpretation of Results

Pre-earthquake is a phase when no one knows that an earthquake may happen. Even though there have been many attempts to predict the happening of earthquake, there has been little success so far. In regions, where earthquakes occur frequently, there are no surprises for people and they can expect it at any time and to some extent they are prepared too. Earthquakes often have aftershocks that give a relative idea to the affected people and some time too, to adopt and adapt to a certain situation that they have experienced at least once. The duration among earthquakes and their aftershocks may span over from some seconds to many hours or even a couple of days. As a sample, results from Tables 1 and 2 (models 1 and 2) are interpreted below. The rest of the models are summarized and discussed in Sect. 6.

In pre and during-earthquake stages, IDV "emergency/rescue authorities" have commonly played a role in choosing from available simple (conventional) ITS technologies. In both cases, "landline/mobile phones" were used frequently when transport-earthquake affected people preferred to seek information from "emergency/rescue authorities". In pre-earthquake phase, along with "emergency/rescue authorities", "dependency status" also contributed in developing statistically significant model. Transport-earthquake affected people used "landline/mobile phones" when they were "alone" in earthquake to deal with situations on their own. During-earthquake two ITS technologies were identified i.e. "landline/mobile phones" and "TV" that have been used to seek transport-earthquake.

6 Discussion

Pre-earthquake is a period prior to an earthquake or after time (before the next earthquake) between closely occurring aftershocks. Seeking transport-earthquake information from "emergency/rescue authorities" under prevailed situations i.e. handling an earthquake "alone" increases the use of "landline/mobile phones". To share this information with other people of community shows the sense of care of one effected person for other people with in their community. The pattern of using simple ITS technologies remain same from pre to during earthquake stage probably because of the sudden change in dynamics with an addition of TV in it, to get visuals of the affect that earthquake have caused. During-earthquake phase, though is very short duration for decision making but effected people decided to seek transport-earthquake information through "landline/mobile phones" and "TV" as authentic sources of information that is backed up by "emergency/rescue authorities". It shows that the community is relied on "emergency/rescue authorities" for seeking information and required help. "Dependence on authorities" and "situational circumstance" in earthquakes are the influencing IDVs that are affecting the choice of simple (conventional) ITS technologies.

In earthquakes, available facilities such as vehicles, dependence on authorities and personal characteristics are the influencing IDVs that compelled the transport-earthquake affected people of Islamabad to seek information through subscription services. People relied upon police to facilitate them though they had their own resources (vehicles) too. Male members of the society dominated over female members of the society in using subscription services to gain transport-earthquake information. In pre and during-earthquake phase all types of subscriptions were used but in post-earthquake phase only mobile subscription were operational. The possible reason is that in post-earthquake phase internet network stopped working due to localized damage to the operating infrastructure but mobile phone network still stayed operational because it was operated from a distant and undamaged infrastructure. However, the services could be availed until battery of mobile phones die.

Transport-earthquake affected people significantly relied on police authorities to be rescued from the earthquake-affected area and owned their own travel means too, to evacuate the earthquake affected area. GPS trackers (application) guided through the routes to evacuate the area along with other applications of similar specifications [28]. From pre to during-earthquake stage, more applications were added in use to achieve a broader range of information. There was a great trust building between police and transport-earthquake affected people. Resources were sufficiently available to the affected people in various earthquake circumstances. Transport-earthquake information flow through "contributory/updates apps" shows the bridging of information and communication gap between authorities and affected people. The population of the affected area, however, appeared to be quite dependent on and pampered by the authorities.

Dependency on authorities is influencing the use of social media public forums in pre and during-earthquake stages. In pre-earthquake stage, "both" (i.e. Facebook and Twitter together) and in during-earthquake stage "both" (i.e. Facebook and Twitter together) as well as "Facebook" forums were used. In post-earthquake stage, only "Facebook" was used that shows that people tried to share information with in their closer social circles

compared to pre and during-earthquake stages where people relied on authorities and other open sources (information shared on Twitter is from authorities and by open sources whilst information shared on Facebook is from other known members of the community and specific authorities through specially created information pages). Highly educated people (i.e. university graduates /professionals) used "Facebook" more compared to the "school/college graduates" probably because the later explored further sources of information too and did not just rely on Facebook solely.

Compared to all other models in which transport-earthquake affected people relied on authorities every time, Table 11, shows an opposite pattern in which people did not rely on police and took an initiative to share information through YouTube probably for an unknown audience (national and international both). However, an opposite pattern is observed in Table 12, again, in which people used YouTube and relied on government authorities probably assumed that their shared information would be of any help to any authority. The use of YouTube in pre-earthquake and during earthquake shows a difference that in former case no expectation from "police" is observed but in later case "government authorities" were relied to take notice of information shared through YouTube. "Experience" of being into earthquakes however, trained affected people over the period of time to learn about possible information sources and their effectiveness. Relying and trusting on the information shared by government authorities however shows that there is a trust building between affected people and authorities. In all three stages both news channel sources were followed by the transport-earthquake affected people. Affected people relied on in-built ITS technologies with in their own vehicles to access both types of news channel sources. In during and post-earthquake stages.

The overall perception is that people are more pampered by the authorities or in fact expect too much from the authorities and their expectation are returned too. This gives an impression that the community lacks an independent decision making to deal with situations on their own.

7 Conclusions

A number of conclusions can be drawn from the research, as follows. Landline/mobile phones and TV are conventional ITS technologies that are used by people even though they are significantly dependent on the infrastructural conditions [16]. A rapid source of information is to call emergency authority to acquire immediate help from emergency/rescue authorities, as is practiced by the transport-earthquake affected people of Islamabad. Any transport-earthquake information in the form of subscription services was helpful to victims, as individuals were confident that the information was informed by the city police. People relied on, and trusted, the police to a large extent. Studies show that general and repeated information that is system generated is not very effective in disasters, in comparison to texts and human-generated information [17]. However the trend was found to be different in the case of Islamabad.

Male members were more active in using subscription services compared to female members of the community in the study. This indicates that females may be less familiar as ITS users and therefore less well-equipped to deal with transport-earthquake

situation independently, possibly due to cultural factors [18]. In the particular case of the study area, women may be affluent and part of a higher socio-economic group. Vehicle ownership motivates individuals to stay updated with transport-earthquake information. They will probably have to evacuate the area and in that case it is important to know specific transport-earthquake information such as situations on different routes, blocked roads, alternative passages, available public-emergency buildings (hospitals) etc.

The trust and understanding between police (authorities) and earthquake victims served to enhance the sharing of information through "contributory/updates apps". People were also self-sufficient to utilize their own resources to evacuate the area, or to approach medical facilities near-by and where available, with the help of guided applications such as GPS trackers. It is a usual practice that people use GPS trackers in disasters [19]. However, further detail to add to the information obtained from GPS trackers would most probably need to be acquired from the police.

Facebook seems more popular in the community because it was used in all three stages of an earthquake. Facebook is used where people try to connect more to their closer social circle [20]. Twitter however, is an open source of information and is used in disasters [21, 22]. University graduates /professionals are frequent users of Facebook (compared to school /college graduates) shows that a higher level of education increases the possibility of using Facebook in the post-earthquake stage. Dedicated Facebook and Twitter information pages especially created by the police and other authorities to share updated information and to give guidelines that are not yet established. People used YouTube to share information but it may not be accessed (watched) by the relevant audience because they might not be aware that where and when it was shared and by whom [23].

There is a need for reforms that make people more independent and ready to deal with such situations, rather than relying heavily on authorities. The burden to the authorities need to be lowered and their resource needs to be directed to other areas where effected people will otherwise not be approached in time. People utilized all available resources to seek information where they could. There is need to train them further to deal with the negative impacts of earthquakes independently similar to some other practices [24].

The overall image is of a society where the needs are well attended to when earthquakes occur. Individuals are personally well resourced as well as provided for by the authorities. Very little community strength or public participation was observed. No significant public participation outside of closer social circles was observed and there was no statistically significant use of smartphone messages.

8 Application of the Work

This work has highlighted different ITS technologies that are useful under different earthquake circumstances and for different users with different capabilities. This study has also highlighted the training needs of affected people in order to better cope with transport-earthquake scenarios, as well as the shifts and deficiencies in roles and responsibilities among the authorities. The findings will help transport-earthquake

managers, planners and policy makers to establish plans and policies according to the needs of the transport-earthquake affected people ensure a fair distribution of responsibilities among different authorities and generate guidance plans and comprehensive training for community members.

9 Future Directions

Future work can be done on designing the guidelines and templates of information that should be directed towards transport-earthquake affected people through ITS technologies. There is a need to work on engaging people with social media, subscription services and information websites to exchange transport-earthquake information. There is need to apply the same model to another case-study which is less dependent on authorities for support in a time of natural disaster [11].

References

1. Grant-Muller, S., Usher, M.: Intelligent transport systems: the propensity for environmental and economic benefits. Technol. Forecast. Soc. Change, (0)
2. Zear, A., Singh, P.K., Singh, Y.: Intelligent transport system: a progressive review. Indian J. Sci. Tech. **9**(32) (2016)
3. Shafiq, F., Ahsan, K.: An ICT based early warning system for flood disasters in Pakistan. Res. J. Recent Sci. ISSN **2277**, 2502 (2014)
4. Faturechi, R., Miller-Hooks, E.: Travel time resilience of roadway networks under disaster. Transp. Res. Part B **70**, 47–64 (2014)
5. Grant-Muller, S., Usher, M.: Intelligent transport systems: the propensity for environmental and economic benefits. Technol. Forecast. Soc. Change **82**, 149–166 (2013)
6. Kolosz, B., Grant-Muller, S., Djemame, K.: Modelling uncertainty in the sustainability of intelligent transport systems for highways using probabilistic data fusion. Environ. Model Softw. **49**, 78–97 (2013)
7. Litman, T.: Lessons from Katrina and Rita: what major disasters can teach transportation planners. J. Transp. Eng. **132**(1), 11–18 (2006)
8. Eisenman, D.P., et al.: Disaster planning and risk communication with vulnerable communities: lessons from Hurricane Katrina. Am. J. Public Health **97**(1), S109–S115 (2007)
9. Nemoto, Y., Hamaguchi, K.: Resilient ICT research based on lessons learned from the Great East Japan Earthquake. Commun. Mag. IEEE **52**(3), 38–43 (2014)
10. Comfort, L.K., Ko, K., Zagorecki, A.: Coordination in rapidly evolving disaster response systems the role of information. Am. Behav. Sci. **48**(3), 295–313 (2004)
11. Anwer, I.: Factors limiting the scope of ITS technologies in disaster relief operations. In: European Transport Conference 2016Association for European Transport (AET) (2016)
12. Phister P.W., Allen, D., Barath, J., Brandenberger, U., Bruehlmann, R., Burton, A., Farrell, P.W., Marien, G.: Pakistan earthquake case study (2009)
13. Bull, S.B., Donner, A.: The efficiency of multinomial logistic regression compared with multiple group discriminant analysis. J. Am. Stat. Assoc. **82**(400), 1118–1122 (1987)
14. Field, A.: Discovering statistics using IBM SPSS statistics. Sage, Thousand Oaks (2013)

15. Muller, A., Whiteman, G.: Exploring the geography of corporate philanthropic disaster response: a study of Fortune Global 500 firms. J. Bus. Ethics **84**(4), 589–603 (2009)
16. Kumar, N., Vragov, R.: Active citizen participation using ICT tools. Commun. ACM **52**(1), 118–121 (2009)
17. Sahami Shirazi, A., et al.: Large-scale assessment of mobile notifications. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, (2014)
18. Broos, A.: Gender and information and communication technologies (ICT) anxiety: Male self-assurance and female hesitation. CyberPsychol. Behav. **8**(1), 21–31 (2005)
19. Zakia, U., et al.: A Navigation system for rescue operation during disaster management using LTE advanced network and WPAN. In: Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016 IEEE 7th Annual. IEEE (2016)
20. Yin, J., Lampert, A., Cameron, M., Robinson, B., Power, R.: Using social media to enhance emergency situation awareness (2012)
21. Panagiotopoulos, P., et al.: Social media in emergency management: twitter as a tool for communicating risks to the public. Technol. Forecast. Soc. Change **111**, 86–96 (2016)
22. Dufty, N.: Twitter turns ten: its use to date in disaster management. Aust. J. Emerg. Manag. **31**(2), 50 (2016)
23. Lange, P.G.: Publicly private and privately public: Social networking on YouTube. J. Comput.-Med. Commun. **13**(1), 361–380 (2007)
24. U.S. Department of Health and Human Services, S.A.a.M.H.S.A.T.m.f.m.h.a.h.s.w.i.m.d., Crisis Counseling Program Toolkit–phases of disaster phases of disaster: collective reactions (2000)
25. Palen, L., Liu, S.B.: Citizen communications in crisis: anticipating a future of ICT-supported public participation. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, New York (2007)
26. Jarašūniene, A.: Analysis of possibilities and proposals of intelligent transport system (its) implementation in Lithuania. Transport **21**(4), 245–251 (2006)
27. Starkweather, J., A.K. Moske, Multinomial logistic regression. Consulted page at September 10th: http://www.edu/rss/class/Jon/Benchmarks/MLR_JDS_Aug2011.pdf (2011)
28. Chatterjee, K., et al.: Driver response to variable message sign information in London. Transp. Res. Part C **10**(2), 149–169 (2002)
29. Hosmer, D.W., Lemeshow, S.: Introduction to the Logistic Regression Model, pp. 1–30. 2nd Edition, Applied Logistic Regression (2000)

Blind Driving by Means of a Steering-Based Predictor Algorithm

Pavlo Bazilinskyy(✉), Charles Beaumont, Xander van der Geest, Reinier de Jonge, Koen van der Kroft, and Joost de Winter

Department of BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands
P.Bazilinskyy@tudelft.nl

Abstract. The aim of this work was to develop and empirically test different algorithms of a lane-keeping assistance system that supports drivers by means of a tone when the car is about to deviate from its lane. These auditory assistance systems were tested in a driving simulator with its screens shut down, so that the participants used auditory feedback only. Five participants drove with a previously published algorithm that predicted the future position of the car based on the current velocity vector, and three new algorithms that predicted the future position based on the momentary speed and steering angle. Results of a total of 5 h of driving across participants showed that, with extensive practice and knowledge of the system, it is possible to drive on a track with sharp curves for 5 min without leaving the road. Future research should aim to improve the intuitiveness of the auditory feedback.

Keywords: Road safety · Driver support · Auditory display · Human–machine interface · Driving simulator

1 Introduction

Road traffic crashes are a major public health problem. If no action is undertaken, road crashes will become the seventh leading cause of death by 2030 [1]. About 95% of crashes are caused by driver error, in particular inattention and distraction [2].

One way to avoid crashes caused by poor driver behavior is automated driving [3]. An important technology, which may be seen as one of the first steps towards autonomous driving, is automated lane keeping—as used in modern Tesla and Mercedes Benz cars, for example. Current automated driving systems cannot predict the behavior of other road users in all situations. Once the system fails to handle a traffic situation, the driver needs to take over the steering and keep the car on the road. These transitory situations are a safety concern if the driver fails to reclaim control properly. The median estimate among the general public is that autonomous driving will be widespread by 2030 [4], whereas some experts argue that autonomous driving will not be feasible before 2075 [5]. Until driving is autonomous, automated driving systems will require driver intervention at certain times.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_45

The aim of this research was to develop and perform a preliminary evaluation of a lane-keeping assistance system in which drivers are supported by auditory feedback as a function of the position on the road [6]. Auditory feedback may be beneficial for regaining steering control from automated driving, especially if the driver is visually overloaded. Auditory feedback may also be useful in regular manual driving when visual information is temporarily lacking, such as when driving in heavy fog or during a visual distraction.

Previous research concurs that the use of real-time feedback can enhance driving performance. Such performance gains were demonstrated for example by Powell et al. [7], who provided tonal cues to racing drivers based on the lateral G-force, and by Houtenbos et al. [8], who provided auditory beeps as a function of the speed and direction of another car approaching an intersection. Furthermore, lane departure and forward collision warning systems are already commercially available. In these systems, the driver is alerted by audio when a problem occurs.

We tested auditory feedback concepts while drivers did not receive any visual information regarding the road environment. This 'blind driving' paradigm can be regarded as the ultimate condition for testing human-machine interfaces: if drivers are able to steer a vehicle by means of sound only, this may provide evidence that the sound cues are effective if visual information is compromised.

This paper presents a design iteration of a previous concept of blind driving by means of auditory feedback (hereafter called 'Blind Driving 1'; BD1 [9]). The focus in BD1 was on investigating how far into the future the system has to 'look' to determine the predicted position of the car—the prediction time. In BD1, auditory feedback was based on the predicted location of the car 0 s, 1 s, 2 s or 3 s into the future. When the predicted location of the car deviated more than 0.5 m from the center of the lane, audio feedback (i.e., a tone) was issued to alert the drive to correct their trajectory. The tone became louder the farther the predicted position was from the lane center. Results of this previous project showed that without predictor feedback (i.e., 0 s prediction), participants were more likely to depart the road compared to with predictor feedback. In this paper, the algorithm presented in BD1 is enhanced with the aim to improve the accuracy of the prediction path.

2 Method

2.1 Apparatus

A fixed-base driving simulator was used (Fig. 1; Green Dino, The Netherlands). An interface was programmed in MATLAB/Simulink r2015a to retrieve location, speed, and steering data from the simulator and to generate audio output via Sennheiser CX-200 headphones. When wearing the headphones, the participants were still able to hear engine and tire sounds via loudspeakers mounted in the simulator. Similar to BD1, the participants had to steer away from the sound: sound on the left was produced when the predicted lateral error was left of the center of the right lane, and vice versa. During the experiment, the screens of the simulator were turned off.

Fig. 1. The driving simulator used in this research. In all trials, the screens were turned off and participants wore headphones.

2.2 Track

The track was a two-lane 7.5-km road without intersections and without other road users (Fig. 2). It contained straight segments and mostly 90-degree sharp curves with a radius of about 20 m (for research using the same track, see [9–12]). The lane width was 5 m. The width of the simulated car was 1.76 m, and its length was 4.22 m. In each trial, participants drove 5 min. The driven distance per trial varied between 2,069 m and 4,206 m, depending on the number of times the car left the road and was reset on the center of the right lane with zero speed. If driving the full 4,206 m, participants encountered twelve 90-degree curves and one 180-degree curve.

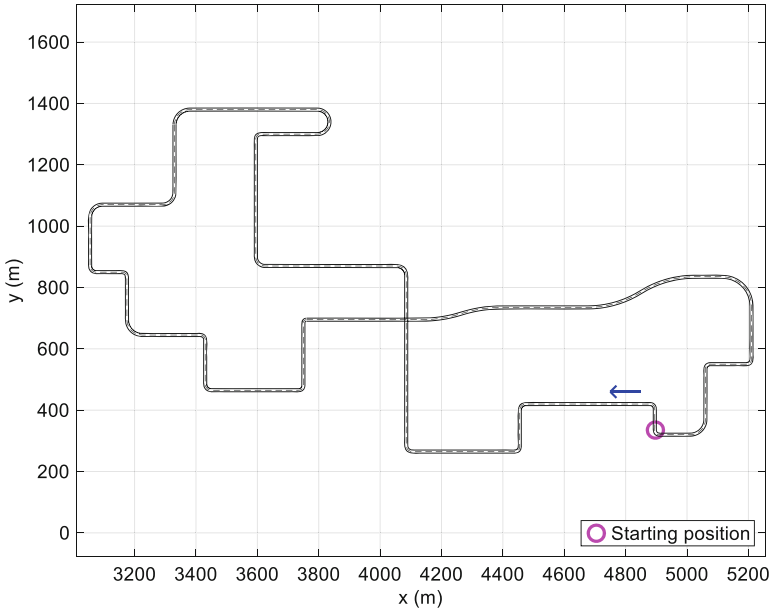


Fig. 2. The test track.

2.3 Participants

In total, five males (mean age = 26.6 years, SD = 6.3 years) participated in the study. None of the participants had hearing impairments. Participants 1–4 (authors 2–5 of this paper) had been involved in the design, and therefore had detailed knowledge of the feedback concepts. Participant 5 was an expert racing driver who was new to the feedback designs and was not informed about their working mechanisms in any way. Participant 5 was invited in order to investigate how well a competent driver, who is naïve to the auditory systems, is able to keep the car on the road.

2.4 Speed and Gearbox Settings

An automatic gearbox was used. The speed of the car was predetermined; the participants did not use the pedals. The car automatically accelerated to a speed of about around 80 km/h on straights, and decelerated to 20–40 km/h for the curves, depending on curve radius.

2.5 New Algorithm for Issuing Feedback

In the BD1 concept, when driving through a curve, the predicted location of the car was mostly outside of the road boundaries because the prediction was based on the momentary velocity vector of the car. In the present study, the steering angle was used in the prediction, making it possible to create a more accurate prediction of the future position

of the car (see also [13, 14]). Figure 3 illustrates the predicted path in the BD1 and BD2 algorithms, both for a 2 s prediction.

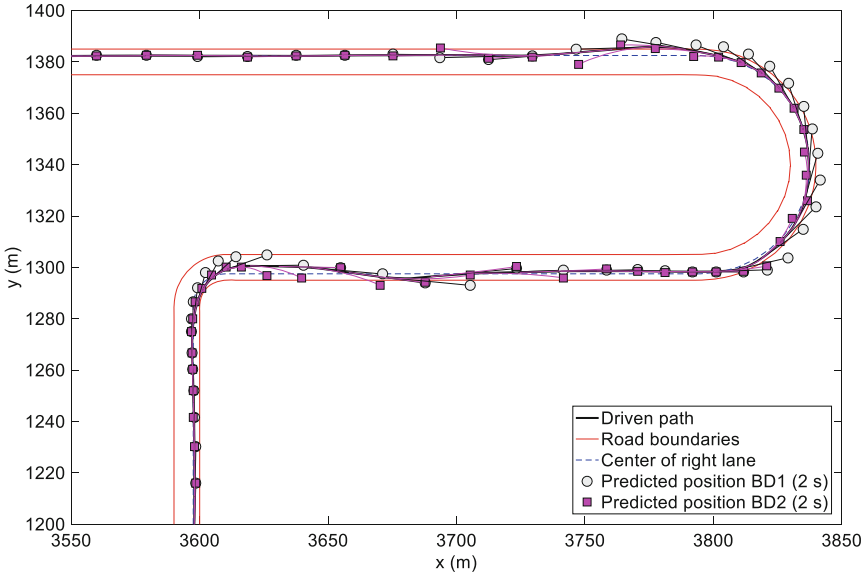


Fig. 3. Working mechanism of the BD1 versus BD2 predictor feedback, both with a 2 s prediction. The figure shows the path driven by a participant through two curves. The circular markers represent the predicted position with the BD1 system (a straight line from the current location), whereas the square markers represent the predicted position with the BD2 system (a curved path from the current location). The markers are shown with 1 s intervals.

It was observed in preliminary tests that a shorter prediction time yielded better driving performance at high speeds and on straights, whereas a longer prediction time yielded better driving at lower speeds and in curves. Long prediction time on straights may lead to oscillatory steering behavior, because a small error is amplified by a long prediction path. A variable prediction time may solve these problems. This study included a condition with a prediction time that varied, from 3 s at 20 km/h (in curves) to 2 s at 80 km/h (on straights).

An overview of the tested concepts is provided in Table 1. In summary, there were two different kinds of feedback: feedback from iteration 1 (BD1), which linearly predicts the vehicle location, and feedback based on the algorithm presented in the current iteration (BD2), which takes steering into account.

In all concepts, volume feedback was provided when the predicted lateral position with respect to the center of the right lane exceeded 0.5 m. The larger the distance from the lane center, the louder the volume became. Further details about the pitch and volume are provided in [9]. The decision to select a 1 m wide tolerance zone was based on our earlier research in which the same threshold was used [9, 15]. De Groot et al. [15] also indicated that off-target feedback (i.e., augmented feedback provided when deviating

more than 0.5 m from the lane center) yielded better lane keeping performance than on-target feedback (i.e., augmented feedback provided when deviating less than 0.5 from the lane center).

Table 1. The four blind driving concepts that were tested by the five participants.

Feedback name	Prediction type
BD1 (2 s)	2 s prediction based on current car speed
BD2 (2 s)	2 s prediction based on current car speed and current steering angle
BD2 (3 s)	3 s prediction based on current car speed and current steering angle
BD2 (3 s)	Variable 2–3 s prediction (3 s when driving at 20 km/h, 2 s when driving at 80 km/h) based on current car speed and steering angle

2.6 Experiment Design

Participants 1–4 tested each of the four algorithms three times, in counterbalanced order. Participant 5 tested each of the four algorithms two times, and also drove four times with visual information (twice with and twice without the BD2 variable algorithm), in counterbalanced order. In summary, each participant performed a total of 12 trials of 5 min each (i.e., 1 h of driving per participant).

2.7 Dependent Variables

Driving performance was assessed by means of the number of resets. A reset occurred when the car drove outside of the road boundaries with all its four corners. Secondly, the on-target percentage (OTP) was used as a measure of lane-keeping accuracy. OTP was defined as the percentage of time that the current absolute lateral position was less than 0.5 m. Data from 3 s prior to 10 s after each reset were excluded from the calculation of OTP.

3 Results

Figure 4 shows the number of resets that the five participants experienced with the four tested systems. The variable prediction time and the 3 s prediction time led to better performance than the other two algorithms. It is noteworthy that for the BD1 system, Participant 5 performed better than Participants 1–4. Participants 1–4 performed better with the BD2 systems (which they designed themselves) than with the BD1 system.

Most resets occurred around curves (see Fig. 5 for an illustration). The prediction time in curves for the variable model was mostly around 3 s (as there are few curves where the car reaches a speed higher than 20 km/h). This explains that the resets for the 3 s prediction time and variable prediction times are similar (Fig. 4). There was one trial without a single reset, for the BD2 system with a 3 s prediction time.

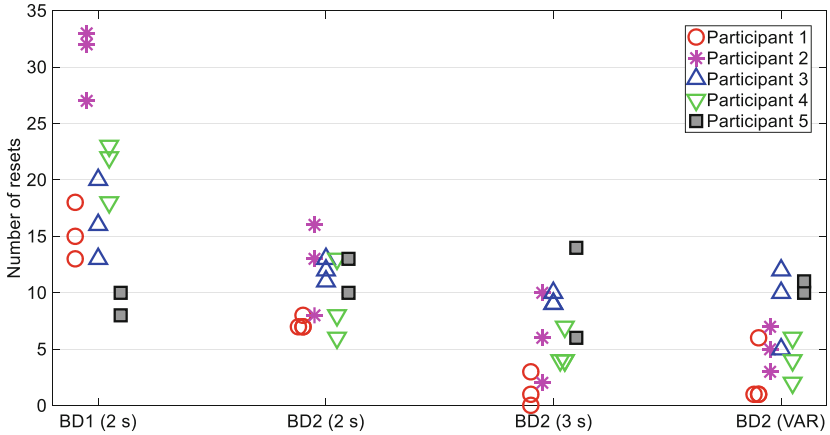


Fig. 4. Number of resets.

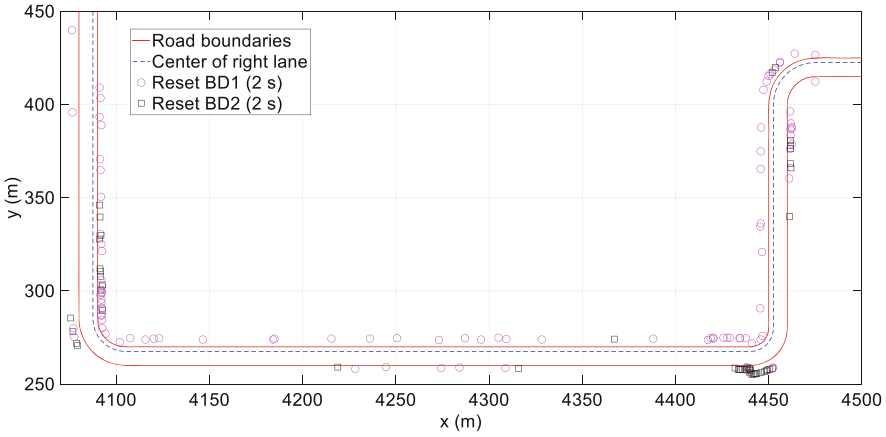


Fig. 5. Locations of resets in part of the course for two of the four algorithms. It can be seen that participants crashed on distinct locations depending on the curve and the concept used.

The results for OTP (Fig. 6) mirror the results of the number of resets (Fig. 4), with the BD1 system yielding lower OTP values than the three BD2 systems, for Participants 1–4. It can also be seen that there were substantial individual differences, with some participants performing substantially better than others. There was no trial that had an OTP greater than 50%. As a reference, Participant 5 attained an OTP of 98% when driving with visual feedback in one of his trials.

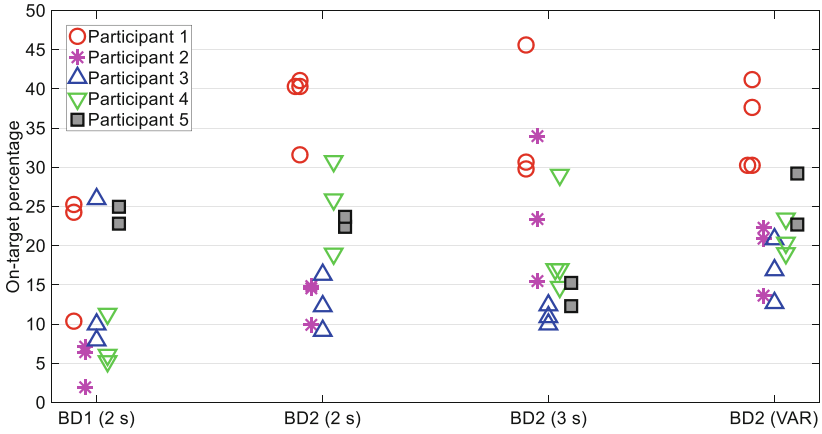


Fig. 6. On-target percentage.

4 Discussion

In this design study, we implemented a predicted position based on momentary steering angle, for providing real-time auditory feedback to the driver.

The results of test drives with human subjects suggest that the proposed concept of using steering angle in the prediction of future position yielded a substantial improvement in driving performance as compared to a velocity-based predictor (Fig. 4), although the proposed concept still yielded resets during driving. The prediction time of 3 s featured one trial without a single reset. In other words, with sufficient practice and knowledge of the workings of the system, it is possible to drive a course for 5 min without leaving the road, purely based on auditory information.

The experiment was conducted with four young male engineering students as participants, who were also working on the project. Hence, these participants are not representative of the general population. Furthermore, because the sample size was small, we did not apply null hypothesis significance testing of any sort. Instead, the goal of this work was to examine whether it is possible to drive blindly by means of sound only.

Participant 5 drove without prior knowledge of the algorithms and performed better with the BD1 algorithm than Participants 1–4 did. Participant 5 commented afterwards that the auditory feedback was very hard to understand even after several trials of practice and after having driven with visual feedback and auditory feedback combined. The fact that the BD1 and BD2 concepts were not based on current lateral position, but on future lateral position, and the fact that BD1 and BD2 required different steering actions for a given audio input, may have been confusing. Participant 5 mentioned afterwards that he had not realized that the way the audio feedback has to be used is fundamentally different for these two algorithms. In BD1, no auditory feedback implies that the driver has to drive straight ahead; therefore, the steering wheel should be in the centered position. In the BD2 algorithms, feedback promotes a change of steering wheel position: no sound means the steering wheel is at the correct angle, and if there is audio feedback,

the driver has to keep turning the steering wheel until the sound stops. These findings suggest that future research should be directed towards the intuitiveness and stimulus-response compatibility of the auditory stimuli.

In the future studies we propose to investigate the effectiveness of multimodal feedback (e.g., a combination of vibrotactile and auditory feedback) in blind driving. A more realistic scenario in which participants have to control the speed of the car themselves may also yield insightful insights. Further development of techniques for issuing auditory feedback based on the type of road and type of curve may also be required. Additionally, future research should apply larger sample sizes as well as female participants. Finally, we point out that the idea of blind driving using headphones is not practical for real-life applications. Our experiment should be seen as a paradigm for investigating the value of auditory feedback under conditions where visual feedback is compromised. Future research could investigate spatial auditory feedback (e.g., via the car's speakers) in naturalistic conditions, such as driving when being visually distracted, when driving in rain or fog, or when visual information is otherwise unavailable or occluded.

Acknowledgements. Pavlo Bazilinskyy and Joost de Winter are involved in the Marie Curie ITN: HFAuto (PITN-GA-2013-605817).

References

1. World Health Organization: Global Status Report on Road Safety 2015. World Health Organization, Geneva, Switzerland (2015)
2. Dingus, T.A., Guo, F., Lee, S., Antin, J.F., Perez, M., Buchanan-King, M., Hankey, J.: Driver crash risk factors and prevalence evaluation using naturalistic driving data. PNAS **113**, 2636–2641 (2016)
3. Jamson, H., Merat, N., Carsten, O., Lai, F.: Fully-automated driving: The road to future vehicles. In: Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 2–9 (2011)
4. Kyriakidis, M., Happee, R., De Winter, J.C.F.: Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. Transport. Res. F **32**, 127–140 (2015)
5. Shladover, S.: The truth about "self-driving" cars. Sci. Am. **314**, 52–57 (2016)
6. Verbist, K., Boer, E.R., Mulder, M., Van Paassen, M.M.: Car lane-keeping using auditory feedback. Delft University of Technology, The Netherlands (2009)
7. Powell, N., Lumsden, J.: Exploring Novel Auditory Displays for Supporting Accelerated Skills Acquisition and Enhanced Performance in Motorsport. In: Proceedings of the 21st International Conference on Auditory Display, pp. 173–180 (2015)
8. Houtenbos, M., De Winter, J.C.F., Hale, A.R., Wieringa, P.A., Hagenzieker, M.P.: Concurrent audio-visual feedback for supporting drivers at intersections: A study using two linked driving simulators. Appl. Ergon. **60**, 30–42 (2016)
9. Bazilinskyy, P., Van der Geest, L., Van Leeuwen, S., Numan, B., Pijnacker, J., De Winter, J.C.F.: Blind driving by means of auditory feedback. In: 13th IFAC Symposium on Analysis, Design, and Evaluation of Human–Machine Systems, pp. 525–530 (2016)
10. De Groot, S., Centeno Ricote, F., De Winter, J.C.F.: The effect of tire grip on learning driving skill and driving style: A driving simulator study. Transport Res F **15**, 413–426 (2012)

11. Van Leeuwen, P.M., Subils, C.G., Jimenez, A.R., Happee, R., De Winter, J.C.F.: Effects of visual fidelity on curve negotiation. Gaze behaviour and simulator discomfort. Ergon. **58**, 1347–1364 (2015)
12. Van Leeuwen, P.M., Happee, R., De Winter, J.C.F.: Vertical field of view restriction in driver training: a simulator-based evaluation. Transport. Res. F. **24**, 169–182 (2014)
13. Godthelp, H., Milgram, P., Blaauw, G.J.: The development of a time-related measure to describe driving strategy. Hum. Factors **26**, 257–268 (1984)
14. Van Winsum, W., Brookhuis, K.A., Waard, D.: A comparison of different ways to approximate time-to-line crossing (TLC) during car driving. Acc. Anal Prev. **32**, 47–56 (2000)
15. De Groot, S., De Winter, J.C.F., López-García, J.M., Mulder, M., Wieringa, P.A.: The effect of concurrent bandwidth feedback on learning the lane keeping task in a driving simulator. Human Factors **53**, 50–62 (2011)

Using Interior Design Principles to Improve the User's Perception of Vehicle Interiors: A Study on Visual Parameters

Susan Shaw(✉), Claudia Escobar, and Shawn Lee

Hyundai America Technical Center, Inc., Ann Arbor, USA
{sshaw, cescobar}@hatci.com, wagen@hyundai.com

Abstract. Consumers purchase a vehicle due to the comfort and spaciousness they feel upon entering the driver's seat. Since perceptions are highly subjective and related to visual effects, this study aimed to identify the visual parameters that improve such perceptions. Visual parameters were identified, and then analyzed using Kansei Engineering to determine which effects would have an impact on the perception of spaciousness and comfort. Consumer clinics subjectively analyzed the visual parameters. Size alone does not determine spaciousness in a vehicle; altering the shape, color, and angle of components make a space feel larger and more comfortable. This study identified visual parameters for further analysis. The study recognized multiple easily modifiable components to increase the perception of spaciousness, feel of comfort, and convenience to the driver. Further testing will ascertain the visual trends that have the most influence on spaciousness and comfort.

Keywords: Visual effects · Vehicle interior design · Subjective evaluation

1 Introduction

Despite spending an average of 2.5 h a day in their cars, many people make their purchase after just a few minutes inside. Most buyers choose a vehicle based on the comfort and spaciousness felt upon entering the driver's seat. Some vehicles, even though designed to be as spacious and comfortable as the market trends demand, will not provide that perception during first impressions. Despite a vehicle interior being larger dimensionally, users may state it feels small or cramped, while dimensionally smaller vehicles may feel/look larger and roomier. Because perceptions are highly subjective and related to visual effects, this study aimed to identify the visual parameters that improve such perceptions, improving the feeling of spaciousness and comfort.

2 Method

2.1 Test Vehicles

JD Powers and IQS Appeals scores of various mid-size vehicles along with consumer feedback and rating determined vehicle selection determined five test vehicles.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_46

2.2 Participants

Experts. Fifteen experts participated in the study. Fields of expertise included architecture, arts, interior design and engineering.

Customers. Twenty-two participants ranging from 35 to 65 years old whom drive daily, have a high school diploma or bachelor's degree and been an American citizen for a minimum of the past 10 years. The participant mix was 45% male and 55% female.

2.3 Visual Parameters

Figure 1 shows the zones used to group the identified visual parameters. Zone 1 encompassed the crash pad, zone 2 covered the center console, zone 3 involved the steering wheel, zone 4 included the door trim interior, and zone 5 covered the frame, including the A-pillar and contacts with the headliner and windshield.

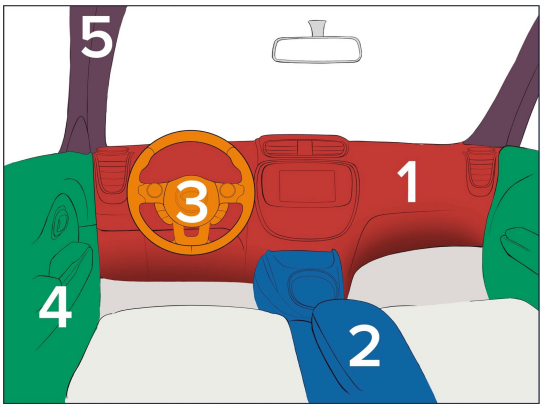


Fig. 1. Vehicle zones for visual parameter evaluation.

Zones also assisted in gathering unaided feedback from participants when answering questions regarding certain zones. If a user mentioned a visual parameter in their response, it more strongly validated the effectiveness of the visual parameter. Kansei Engineering analyzed the visual parameters to determine which effects would have an impact on the perception of spaciousness and comfort in the vehicle.

2.4 Kansei Engineering

Kansei Engineering allows translating technology of a consumer's feeling and image for a product into design elements, originally designed as a consumer-oriented technology for new product development [1]. This study used Kansei best practices such as using affirmative-negative pairs (robust and not robust) rather than antonym pairs

(robust and weak). Users rated each zone on a scale from 1 to 7 where the affirmative was a 7 and the negative a 1.

Table 1 shows the list of Kansei words from a previous research study in Korea [2], used as reference for this project. The revised list in Table 2 is a re-evaluation of Table 1 based on word definitions best understood by an American audience.

Table 1. Kansei engineering pairs.

Factor	Item	Definition	Cronbackh's α
Space	Perfect	Entirely without any flaws, defects, or shortcomings	0.91
Completeness	Harmonious	Forming a pleasingly consistent whole	
	Enough	Sufficient for the purpose or to satisfy desire	
	Unified	Making or becoming a single unit	
	Impressive	Having the ability to impress the mind	
	Delicate	Fine in texture, quality, construction	
	Elegant	Tastfully fine or luxurious in style, design	
	Well-balanced	Rightly balanced, adjusted, or regulated	
	Luxurious	Ministering of conducive to luxury	
Openness	Big	Large, as in size, height width, or amount	0.80
	High	Having a great or considerable extent or reach upward or	
	Vast	Of very great area or extent	
	Refreshing	Having the power to restore freshness, vitality	
	Pleasant	Pleasing, agreeable, or enjoyable	
	Cleared	Free from darkness, obscurity, or cloudiness	
	Cool	Feeling comfortably or moderately cold	
	Flatten	Making flat	
	Cozy	Snugly warm and comfortable	
Dullness	Blunt	Having an obtuse, thick or dull edge or point	0.78
	Lumpy	Full of lumps	
	Voluminous	Of ample size, extent or fullness	
	Stumpy	Short and thick	
	Robust	Strongly or stoutly built	
	Hard	Solid and firm to the touch	

(*continued*)

Table 1. (*continued*)

Factor	Item	Definition	Cronbackh's α
Narrowness	Cramped	Confined or severely limited in space	0.74
	Somber	Gloomily dark	
	Stuffy	lacking fresh air or ventilation	
Perceived roominess	–	How spacious the in-vehicle space is	–

2.5 Visual Parameter Evaluation

Statistical Factor Scoring. Scoring each statistical factor consisted of averaging the response of several Kansei words that make up the specific factors. The previous study had between 3 and 9 Kansei words per factor. This revised study had 3 factors with 2 to 4 words per factor.

Table 2. Revised Kansei engineering pairs.

	OPENNESS	COHESIVENESS	COMFORT	Factor4	Factor5
Big	0.82	-0.04	0.22	0.08	0.02
Vast	0.81	-0.06	0.13	0.15	0.10
Robust	0.73	-0.24	-0.03	0.14	-0.02
Spacious	0.70	-0.13	0.48	0.12	-0.01
Impressive	0.50	-0.24	0.14	0.42	0.33
Unified	0.17	-0.87	0.08	0.03	0.06
Well-Balanced	0.09	-0.78	0.25	0.08	0.19
Pleasant	0.28	-0.49	0.32	0.38	0.30
Not Stuffy	-0.09	0.15	-0.80	-0.17	-0.21
Not Cramped	-0.35	0.25	-0.74	0.01	-0.07
Hard	-0.20	-0.04	-0.01	-0.83	-0.05
Gloomy	-0.10	0.41	-0.35	-0.57	0.07
Lumpy	-0.03	0.18	-0.17	-0.04	-0.92

Grouping into factors occurred based on a statistical method called Factor Analysis. In Kansei engineering, these factors are a 'psychological space' that defines how people think about automotive interiors (Fig. 2).

Expert Evaluation. Space-design related experts participated in a clinic offering insight into the application of visual parameter design factors from their industry experience. This insight applied to the interior design of each test vehicle. They initially

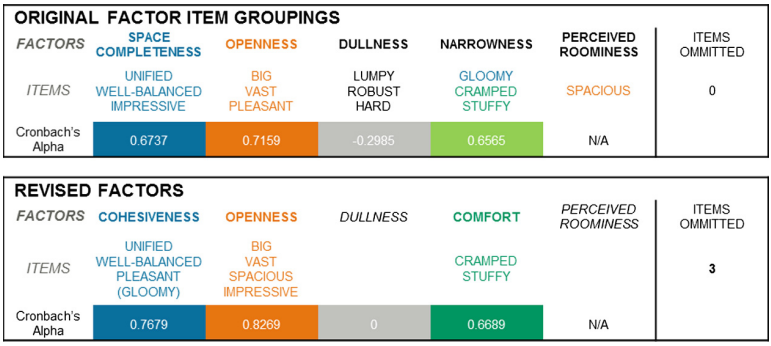


Fig. 2. Kansei engineering factors.

presented on various visual parameters from their portfolio and personal industry experience prior to looking at vehicles and analyzing the effect of the various design factors.

Customer Evaluation. A second evaluation conducted with the participants gathered their perception of spaciousness and comfort. Participants ranked the vehicles by feel of comfort and spaciousness/openness. Each participant received a custom randomized vehicle rotation schedule for the clinic.

Focus Groups. Participants joined focus groups after the Customer Evaluation. They share their preferences and feedback, and gave unbiased ranks from most to least spacious and comfortable.

3 Results

3.1 Identified Visual Parameters

Size alone does not determine spaciousness in a vehicle. Altering the shape, color, and angle of common components can make a space feel larger and more comfortable. Table 3 shows the thirteen identified visual parameters including a dashboard composed of a straight versus a curved line. A curved line is longer than a straight line but the added material may make the space feel smaller. Analyzing colors on the wheel, dash, and door additionally identified color and material matching/continuity as a visual parameter. Other parameters include wide versus narrow angles and vanishing points as well as the ascending and descending slopes of vehicle components. Figure 3 graphically describes each visual parameter.

3.2 Comfort, Cohesiveness and Spaciousness

Comparing the identified visual parameters by zone estimated comfort, cohesiveness and spaciousness (feeling of openness) for the five test vehicles. Figure 4 shows comparisons of the tested vehicles.

Table 3. Identified visual parameters by zone

Zone	Visual parameter
Zone 1	1. Crash pad forced perspective vanishing point
	2. Center stack vertical angle
	3. Crash pad color balance
	4. Crash pad top surface contour
	14. Dash to door continuity
	5. Center stack surface contour
Zone 2	9. Center console slope
Zone 3	10. Steering wheel spoke number
	11. Steering wheel color unity
Zone 4	7. Door profile deletion/addition
	8. Door arm rest slope
Zone 5	13. Windshield angle
	6. A-pillar contour

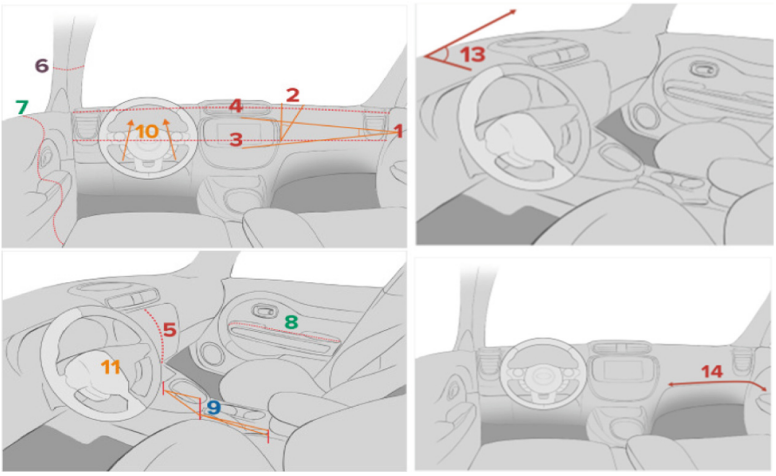


Fig. 3. Identified visual parameters: see Table 3 for names of effects.

Each vehicle had a combination of visual parameters included. No one vehicle stood out from the rest of the vehicles. Each zone identified multiple easily modifiable components to increase the perception of spaciousness, feel of comfort, and convenience to the driver. From the three design goals of spaciousness, comfort, and convenience, spaciousness was the least desired in all zones. The design goal that is most influential on the user preference of each zone varies by functionality. Spaciousness was preferred in zones 2, 3 and 5 whereas convenience was preferred in zone 1 and comfort in zone 2. Convenience due to the frequent use of items in zone 1 while driving.

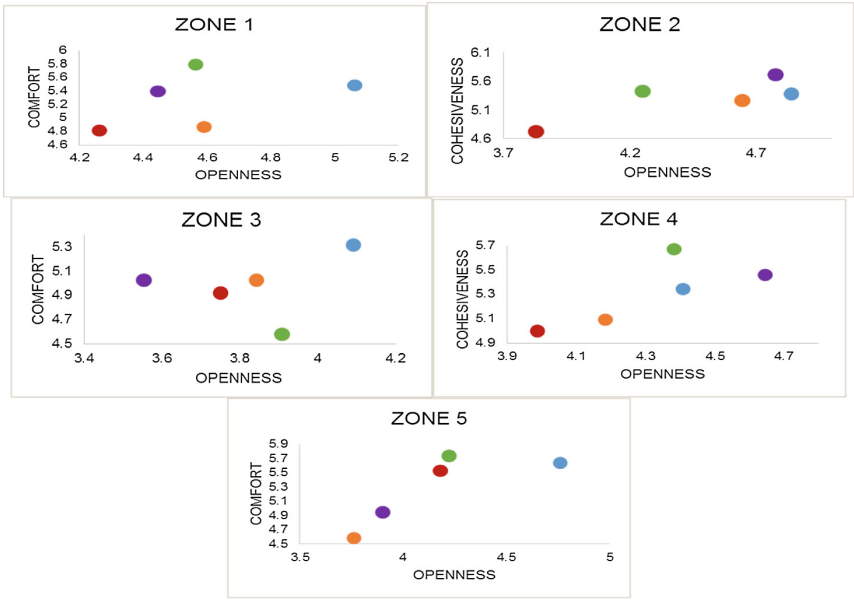


Fig. 4. Comparisons for five test vehicles by zone.

4 Conclusions

This study was able to serve in identifying several visual parameters related to vehicle interior design. Simple modifications in shape, color, and angle increase the perception of spaciousness and comfort.

Some noted limitations include that this study used only mid-size vehicles. It would be beneficial to conduct the study again across other segments, mainly premium vehicles, known for their roominess and comfort.

Next steps involve testing the identified effects to detect visual trends that have the most influence on spaciousness and comfort from the end user perspective, then combining the top effects into packages for further validation using virtual reality and prototyping.

References

1. Jindo, T., Hirasago, K.: Application studies to car interior of Kansei engineering. Int. J. Ind. Ergonom. **19**, 105–114 (1997)
2. Yang, E., Ahn, H.-J., Kim, N.-H., Jung, H.-S., Kim, K.-R., Hwang, W.: Perceived interior space of motor vehicles based on illusory design elements. Hum. Factors Man. **25**, 573–584 (2015)

Possibility of Automobile Seat Evaluation with Seat Fidgets and Movements

Junya Tatsuno[1(✉)], Setsuo Maeda[2], George Sammonds[3], and Neil Mansfield[3]

[1] Faculty of Engineering, Kindai University, 1 Takaya-umenobe, Higashi-hiroshima City, Hiroshima 739-2116, Japan
tatsuno@hiro.kindai.ac.jp
[2] Faculty of Applied Sociology, Kindai University, 3-4-1 Kowakae, Higashi-osaka City, Osaka 577-8502, Japan
[3] Dyson School of Design Engineering, Imperial College of London, South Kensington Campus, London SW7 2AZ, UK

Abstract. In the previous study, we suggested the seat fidgets and movements (SFMs) to evaluate driver discomfort which increases with time. As the next stage, we investigated application possibility of SFMs to the evaluation of the automobile seats. In the experiment, three types of automobile seat were provided. The participants were asked to travel on the road with some vibration section for 60 min. During the experiment, they executed reading task instead of a driving task to simulate front seat passenger. Besides, the participant movement was video-recorded to analyze the frequency of SFMs off-line. As a result, the acceleration values varied according to the type of seat and were significantly different. The participants' discomfort ratings had a different tendency from the acceleration values. The SFMs frequency occasionally indicated a similar tendency to the discomfort ratings. These results exhibit that SFMs will be one of the useful ways to evaluate ride comfort objectively.

Keywords: Whole-body vibration · ISO2631-1 · Unconscious action · Vehicle ergonomics

1 Introduction

The occupants of an automobile are exposed to whole-body vibration during traveling. Many technologies such as suspension system [1–3] and tire [4, 5] have been developed to decrease whole-body vibration exposure. The automotive seat must be one of the essential elements for reducing whole-body vibration because such vibration is finally transmitted through the seat. Conventionally, automotive seats have been mainly evaluated using the SEAT value, which is defined in ISO 10326-1:1992 [6]. The SEAT value is the transmissibility calculated from the acceleration on the seat and the acceleration at the base of the seat. Since the SEAT value is derived from data obtained in laboratory experiments, several problems have been pointed out. For example, Yu and Khameneh [7] reported that subjects were unable to distinguish between three seats with different transmissibility characteristics. Besides, the methods for measuring and

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_47

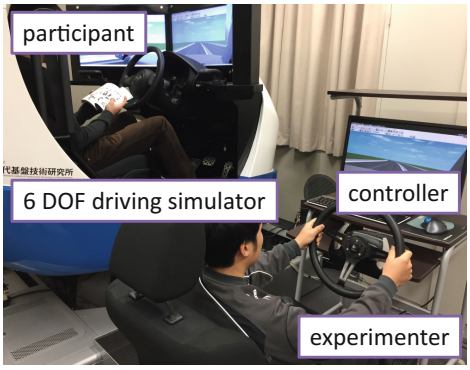
evaluating the whole-body vibration of vehicles are defined in ISO2631-1 [8]. In this standard, the degree of discomfort is defined using the magnitude of physical vibrations. Several investigating studies were also conducted using the measurements described in ISO2631-1 to evaluate whole-body vibration in automobiles [9–11].

On the other hands, it was reported that the duration of whole-body vibration exposure has an influence on the discomfort of occupants in the automotive seat [12]. When a constant vibration is exposed to occupants, the evaluation result with the scales of ISO 2631-1 never gradually increases. Therefore, several studies have focused on quantification of long-term driver discomfort. Sammonds et al. proposed an objective measure of discomfort via the analysis of fidgets and movements in a car seat [13]. Their experimental result showed that the frequency of participants' seat fidgets and movements (SFMs) increased with duration of driving and frequency of volunteers' seat fidgets and movements (SFMs) had a high correlation with the subjective discomfort rating. Then, as the next stage of the research, we performed driving simulator experiments to investigate application possibility of SFMs to the evaluation of the automobile seats.

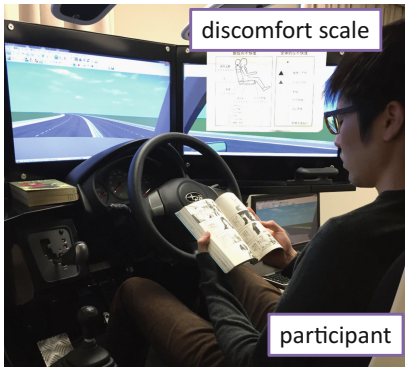
2 Material and Method

2.1 Experimental Outline

Figure 1 shows the configuration of the experimental apparatus. For whole-body vibration exposure to participants, we used a 6 DOF driving simulator (Fuji Heavy Industries, Japan). In this experiment, the participants were exposed to whole-body vibration as passengers and were asked to execute a reading task instead of a driving task. Then, the experimenter operated the driving simulator at the speed of 60 km/h with the external controller. We used the software (UC-win/Road, Forum8 Co., Ltd, Japan) for VR environment. We created the test course of approximate 7 km length and fixed fifteen bumpy sections randomly. By connecting the end point and the initial point



(a) Whole experimental apparatus



(b) Participants' task

Fig. 1. Photos of experimental apparatus.

virtually, we could provide the semi-permanent traveling. Each bumpy section consisted of 10-speed bumps. We reported that varying the height of the speed bumps could control the magnitude of the vibration to which the subjects were exposed [14]. In the experiment, the height of the speed bump was configured as 1 cm, 2 cm and 3 cm.

2.2 Test Seats

We prepared three types of automobile seat (A, B and C) for the experiment. Both Seat A and B were installed to commercially available C-segment cars. Seat C was used for a commercially available luxury car. Before starting traveling, we replaced the seat according to the order that was randomly determined for each participant (Fig. 2).

(a) Seat A (b) Seat B (c) Seat C

Fig. 2. Photos of test seats.

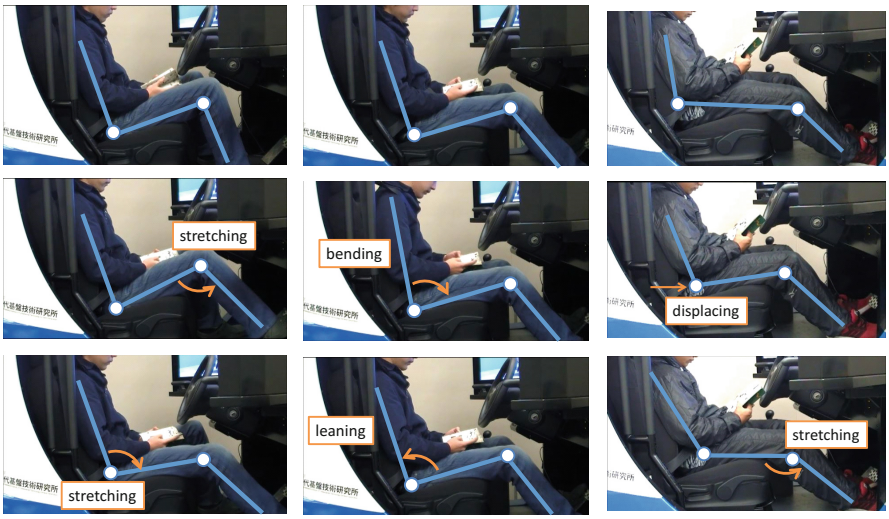
2.3 Evaluation Method

Whole-Body Vibration Exposure. Since we had concerns about the effects of fixing accelerometer on participants' feeling, we carried out another traveling after first traveling to measure the whole-body vibration exposure. On the course for measuring vibration acceleration, there are nine vibration sections (3 bump heights × 3 repetitions). The weighted root-mean-square (rms) accelerations on the seat surface were measured by ISO2631-1. In this measurement, an accelerometer (Brüel & Kjær, Type 4515-B) and a vibration meter (Rion, VM-54) were used.

Subjective Rating. The participants were required to rate their level of discomfort every 10 min. This evaluation method was also used in the previous study. The evaluation was the discomfort scale developed based by the Borg scale. The participants rated their overall discomfort by choosing a number between 0 and 100.

Seat Fidgets and Movement. We video-recorded the participant movement during the experiment with the handy video recorder (GZ-MG275, Victor, Japan). After each trial

had finished, we investigated the appearance timing of participants' unconscious movement. Figure 3 shows images captured from sample movie. We recognized that there are three types of unconscious movements as follows:



(a) movement of lower limb (b) movement of torso (c) movement of whole-body

Fig. 3. Captured image from sample video.

- Type 1: any movement of the limbs (excluding any rubbing/scratching of the face etc.)
- Type 2: any movement of the torso
- Type 3: any movement of the whole body

Since the correction between the magnitude of SFMs and the duration of traveling was not evident, we enumerated SFMs without any distinction in the analysis.

2.4 Participants

The trial participants included five male students with driving licenses. Before the experiment, permission was obtained from the bioethics committee of the Faculty of Engineering, Kindai University. Furthermore, the informed consent was obtained from all participants (Table 1).

Table 1. Participants' profiles.

Participant	Age (years)	Height (cm)	Weight (kg)	Order
P1	22	171	53	B-A-C
P2	22	172	53	C-B-A
P3	21	158	67	C-A-B
P4	21	165	53	A-B-C
P5	21	167	60	B-C-A

3 Results and Discussion

3.1 Vibration Acceleration

Table 2 shows the vibration total value by using measurement results of frequency-weighted acceleration. We calculated the mean of the vibration total value for each seat, as shown in Fig. 4. From this result, it was found that the vibration total value became greater when the bump of vibration section was set higher. By using the scale defined in ISO 2631-1, the mean values were classified as "Not uncomfortable" when the bump height was 1 cm. Moreover, the mean values of 2 cm and 3 cm were rated as "A little uncomfortable" and "fairly uncomfortable", respectively. To verify the differences of the seat, we conducted one-way ANOVA. Consequently, the results for 2 cm ($F(2, 8) = 5.50$, $p < 0.05$) and 3 cm ($F(2, 8) = 4.08$, $p < 0.10$) were significant. If we focused on the acceleration values for the comparison of the seat, Seat A and C relatively become better than Seat B.

Table 2. Vibration total values measured under each seat condition.

Seat	A			B			C		
bump	1 cm	2 cm	3 cm	1 cm	2 cm	3 cm	1 cm	2 cm	3 cm
P1	0.2005	0.3761	0.5930	0.2032	0.4092	0.6439	0.2114	0.3803	0.5928
P2	0.1885	0.3449	0.5227	0.1827	0.3927	0.6392	0.2021	0.3735	0.6863
P3	0.2132	0.4601	0.7221	0.2743	0.5198	0.7696	0.2521	0.4519	0.6949
P4	0.2434	0.4424	0.6903	0.2689	0.5418	0.8121	0.1821	0.3477	0.5833
P5	0.1722	0.3455	0.5997	0.2110	0.4880	0.7857	0.1977	0.4174	0.6267

3.2 Discomfort Rating

As mentioned above, the participants answered their discomfort with value be-tween 0 to 100 every 10 min. Tables 3, 4 and 5 are the participants' rating result for Seat A, B, and C, respectively. By using these data, we calculated increments from first discomfort rating and made a bar chart as shown in Fig. 5. It was found that the discomfort rating for Seat B was greater than Seat A and C. This tendency is similar to the total vibration value. On the other hand, the discomfort ratings for Seat A were relatively getting bigger after 30 min passed. This result suggests that ride comfort of Seat A may substantially change with the progress of traveling time. However, as a result of one-way ANOVA, there were no significant differences among seat types.

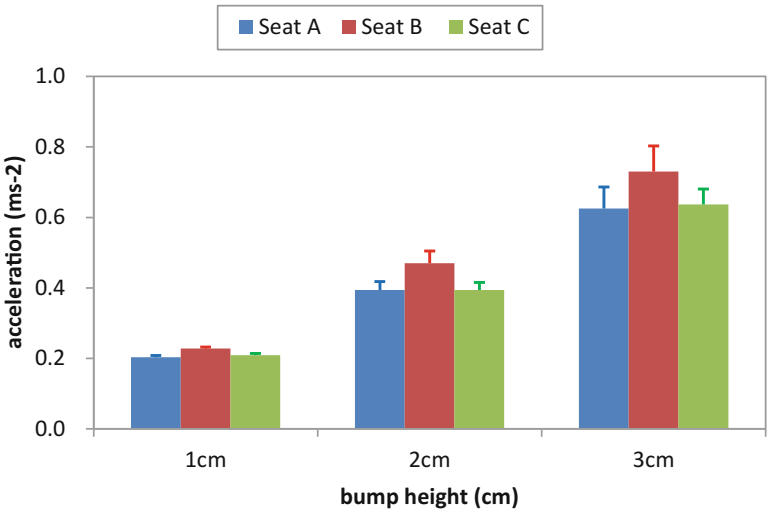


Fig. 4. Mean of the vibration total value measured on each bumpy section.

3.3 SFMs

Tables 6, 7 and 8 show the frequency of SFMs when each participant sat on Seat A, B, and C, respectively. Also, Fig. 6 illustrates the mean of SFMs frequency of each seat. When we compared Seat B with Seat C, the SFMs frequencies of Seat B became greater than Seat C mostly. As for these two seats, the similar tendency to the vibration values could be observed. The vibration total value of Seat A was smaller than Seat B as shown in Fig. 4. Conversely, the subjective discomfort of Seat A was greater than Seat B in the last half of the experiment as shown in Fig. 5. Although there are large variations in the SFMs frequency because of limited sample size, the SFMs frequencies of Seat A were occasionally greater than Seat B. Thus, these results exhibit that SFMs will be one of the useful ways to evaluate ride comfort objectively.

Table 3. Participants' discomfort rating for Seat A.

Time (min)	0	10	20	30	40	50	60
P1	0	5	15	60	65	70	80
P2	1	3	7	15	35	75	75
P3	1	2	7	10	13	17	20
P4	0	15	20	25	25	30	20
P5	1	2	3	8	12	15	17

Table 4. Participants' discomfort rating for Seat B.

Time (min)	0	10	20	30	40	50	60
P1	0	10	50	60	65	70	80
P2	2	5	17	27	35	60	65
P3	1	1	5	8	3	7	7
P4	1	2.5	10	17	30	30	35
P5	1	2	4	5	8	10	12

Table 5. Participants' discomfort rating for Seat C.

Time (min)	0	10	20	30	40	50	60
P1	0	5	15	20	35	50	65
P2	10	15	15	30	40	55	60
P3	8	10	17	20	20	23	23
P4	1	13	20	23	25	27	27
P5	1	2	3	8	12	15	20

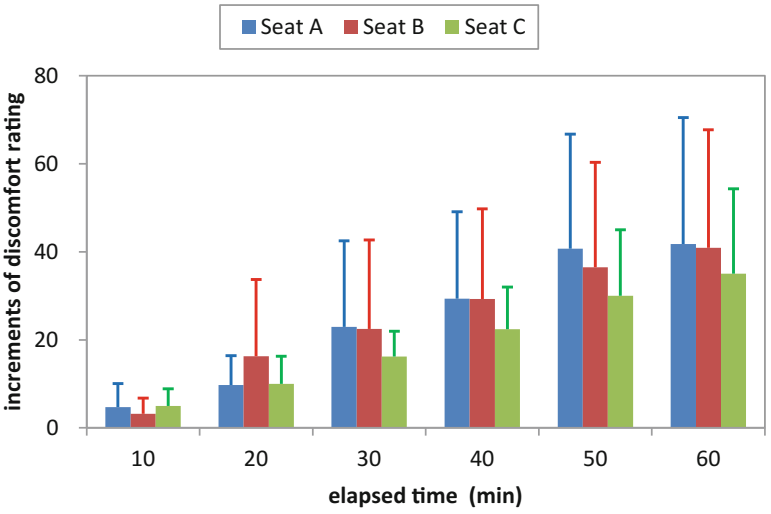


Fig. 5. Mean of the increments from first discomfort rating.

Table 6. Frequency of SFMs when participants sat on Seat A.

Time (min)	0 to 10	10 to 20	20 to 30	30 to 40	40 to 50	50 to 60
P1	0	0	1	0	0	2
P2	2	2	3	4	2	3
P3	1	1	1	4	1	3
P4	0	2	5	4	5	8
P5	2	1	2	0	1	1

Table 7. Frequency of SFMs when participants sat on Seat B.

Time (min)	0 to 10	10 to 20	20 to 30	30 to 40	40 to 50	50 to 60
P1	2	0	1	4	1	4
P2	5	4	7	4	6	2
P3	1	5	1	2	2	3
P4	2	1	2	5	3	4
P5	4	2	1	4	3	3

Table 8. Frequency of SFMs when participants sat on Seat C.

Time (min)	0 to 10	10 to 20	20 to 30	30 to 40	40 to 50	50 to 60
P1	2	0	3	1	2	2
P2	2	3	0	1	4	3
P3	0	0	1	0	2	0
P4	5	3	5	6	6	4
P5	5	2	2	2	1	1

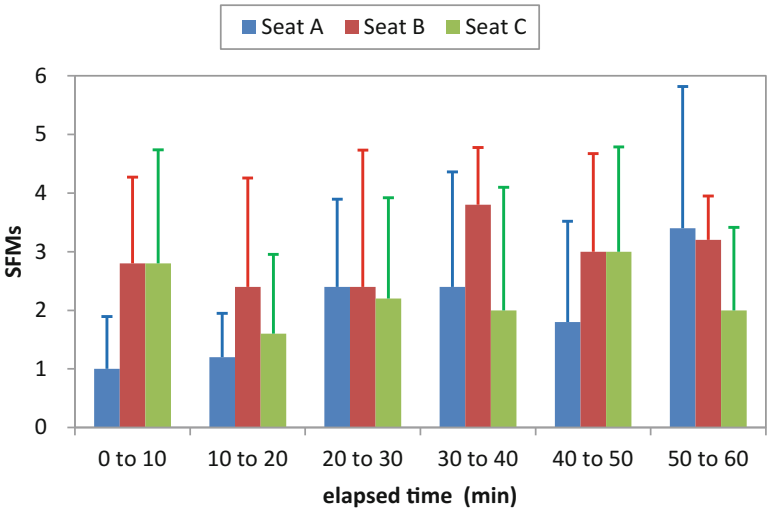


Fig. 6. Mean of the SFMs frequency.

4 Conclusion

Automobile seat is one of the important elements for reducing whole-body vibration, and many methods have been proposed to evaluate automobile seat comfort subjectively or objectively. The international standards ISO 2631-1 are defined for measuring

and evaluating the whole-body vibration of vehicles. This standard employs the magnitude of physical vibration to evaluate the degree of discomfort. Meanwhile, it has been recognized that discomfort in the vehicle may increase as traveling duration extends. Therefore, several studies have focused on quantification of long-term driver discomfort. In the previous study, it was suggested that frequency of participants' seat fidgets and movements (SFMs) had a high correlation with the subjective discomfort rating. Then, as the next stage of the research, we performed driving simulator experiments to investigate application possibility of SFMs to the evaluation of the automobile seats.

We provided three types of the automobile seat in the experiment. The five participants were asked to ride the driving simulator three times under the condition installed each seat to the driving simulator. We utilized three kinds of evaluation in the experiment; vibration acceleration, subjective discomfort rating, and SFMs analysis. The vibration total values varied according to the type of seat and were significantly different. When the mean values of the participants' discomfort ratings to each seat every sampling time were calculated, the different tendency from the acceleration values appeared in last half of test duration. Besides, the mean values of the participants' SFMs frequency occasionally indicated a similar tendency to the participants' discomfort ratings. Thus, these results exhibit that SFMs will be one of the useful ways to evaluate ride comfort objectively.

However, there are some problems to be solved. First, we have to increase the sample size to lower variations. We are going to carry out this experiment for ten participants. As the next phase of this research, we are planning another experiment where participants ride the driving simulator for a longer duration than this experiment. In the experiment, since counting of the SFMs was manually performed, the workload was massive. Thus, improving working efficiency is one of the important issues. Moreover, autonomous driving technology has been rapidly evolving in recent days. Human vibration perception, while performing the second task instead of driving, is a major research issue. This study will contribute to developing automobile seat on which occupants can perform the second task comfortably.

References

1. Yokota, Y., Asami, K., Hamajima, T., Nakajima, N.: Toyota Electronic Modulated Suspension (TEMS) System for the 1983 Soarer, SAE Technical Paper, No. 840341 (1984)
2. Tanaka, T., Tani, M., Yuasa, H., Mimuro, T., Yoshida, H.: Active Control Technology For Passenger Car, SAE Technical Paper, No. 885105 (1988)
3. Satoh, M., Fukushima, N., Akatsu, Y., Fujimura, I., Fukuyama, K.: An active suspension employing an electrohydraulic pressure control system. In: Proceedings of the 29th IEEE Conference on Decision and Control, pp. 2226–2231 (1990)
4. Gipser, M.: FTire: a new fast tire model for ride comfort simulations. In: International ADAMS User's Conference, Berlin, Germany (1999)
5. Mancosu, F., Matrascia, G., CheliPirelli, F.: Techniques for determining the parameters of a two-dimensional tire model for the study of ride comfort. Tire Sci. Technol. **25**(3), 187–213 (1997)

6. International Organization for Standardization: mechanical vibration - Laboratory method for evaluating vehicle seat vibration, ISO 10326-1 (1992)
7. Yu, L., Khameneh, K.: Automotive Seating Foam: Subjective Dynamic Comfort Study, SAE Technical Paper 1999-01-0588 (1999)
8. International Organization for Standardization: Mechanical vibration and shock evaluation of human exposure to whole-body vibration. International Standard, ISO 2631-1 (1997)
9. Nishiyama, S., Uesugi, N., Teshima, T., Kano, Y., Togii, H.: Research on vibration characteristics between human body and seat, steering wheel, and pedals (effects of seat position on ride comfort). J. Sound Vib. **236**(1), 1–21 (2000)
10. Maeda, S.: Necessary research for standardization of subjective scaling of whole-body vibration. Ind. Health **43**, 390–401 (2005)
11. El Falou, W., Duchene, J., Grabisch, M., Hewson, D., Langeron, Y., Lino, F.: Evaluation of driver discomfort during long-duration car driving. Appl. Ergon. **34**(3), 249–255 (2003)
12. Mansfield, N.: Human Response to Vibration. CRC Press, London (2005)
13. Sammonds, G.M., Fray, M., Mansfield, N.J.: Effect of long term driving on driver discomfort and its relationship with seat fidgets and movements (SFMs). Appl. Ergon. **58**, 119–127 (2017)
14. Tatsuno, J., Nakamura, H., Yonehara, M., Takehara, S., Ishimatsu, K., Maeda, S.: Driving simulator for subjective scaling validation of the ISO 2631-1 standard. In: Proceedings of the 40th International Congress and Exposition on Noise Control Engineering (inter-noise 2011), DVD-ROM (2011)

A Study on the Relationship Between Passengers' Visual Searching Efficiency in Bus-Waiting Process and the Elements of Bus Front Face Design

Tao Xiong(✉), Jianxin Cheng, and Tengye Li

School of Art Design and Media, East China University of Science & Technology, M. BOX 286, NO. 130 Meilong Road, Xuhui District, Shanghai, 200237, China
472353860@qq.com, 13901633292@163.com, lty900821@qq.com

Abstract. Visual search is an important cognitive method for people to obtain information. Through eye movement, people can obtain and process visual information. We can identify eye movement patterns of individuals or multiple individuals by analyzing experiment data of eye movement. When people take public transportation, bus-waiting is an important stage. During the process of waiting, it is the key to get the vehicles' identity information accurately and quickly. Early studies have found that the bus face is an essential element for recognition. However, there are still some problems such as the ambiguity of human-computer interaction in this area. Based on the previous researches, this paper studies the layout, color and other factors of the information of bus face by using eye movement equipment from the aspects of design and psychology, which can also provide principles for the optimization of bus face and enhance passengers' information acquisition experience.

Keywords: Industrial design · Visual search · Eye tracking study · Bus

1 Introduction

With the rapid development of urban public transportation system, cities in China that start using bus earlier have completed the standardization of bus design [1]. Most of the information on urban buses is mainly located outside the body, especially on the front face, usually carrying important information such as bus line number, start station and terminal station and so on, which is very important for passenger's waiting experience and information recognition efficiency. However, the preliminary study found that as an important information zone of the bus, there are still many problems on the front part because many passengers can't timely and accurately identify the bus information to complete the recognition process.

At present, researches on bus at home and overseas mainly focus on the field of man-machine relationship inside the bus cabin, such as the dimensions of the seat and the height of the handrail, but researches outside the bus based on the visual searching features are much fewer. Therefore, based on the early analysis of the bus appearance system, this paper will redesign the face of the bus, and make it an archetype. And then

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_48

the eye tracking technology was used to record subjects' visual searching trajectory in the process of observation, analyze their visual data of the bus face and summarize the relationship between the passengers' visual searching features and the bus face, which will be used as the guidelines for optimizing the front face of buses [2].

2 Researching Methods

Based on the previous research, this paper has redesigned the face of the bus body. And then we made a simulation of passenger waiting scenario, using the eye movement equipment to record the subjects' visual search trajectory during the waiting process. Through the comparative analysis of the subjects' eye movement data, we summarized the basic relationship between the visual search of waiting passengers and the face of the bus body, optimizing the design of the bus body.

2.1 Optimization of Bus Face Design

In the previous study, the main conclusions about the front part of the bus body are as follows. (1) In the process of waiting, the front part of the bus is the hot zone that passengers have been paying most attention to, and the order of attention in this zone clearly has a spatial order, which follows the order of "up, down, left and right". (2) Passengers in the waiting process will give priority to digital information, and then the other information. (3) It is urgent to improve the information visualization on the bus face. The result shows that the front face is an important area recognized by passengers, and the upper part of it carries the most important information, so layout, colors, and typeface of this part should be redesigned to reduce the difficulty of cognition. (4) There is also a huge margin for improvement as to the lower part, due to its flexibility.

Based on the conclusions above, and subsequent information collection and questionnaire surveys, the redesign plan of the bus face is shown in Fig. 1.

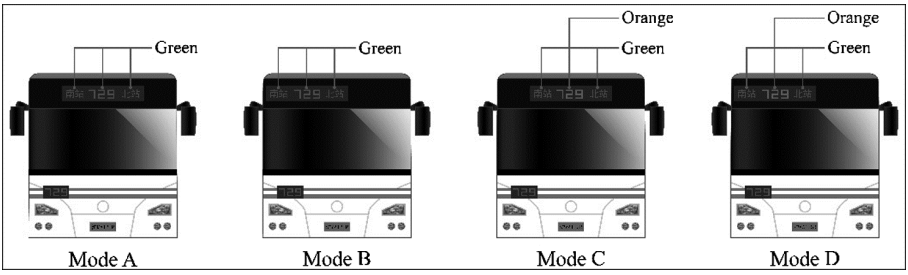


Fig. 1. The redesign plan of the bus face.

Next we will do experiments and data analysis on Mode A and Mode B respectively at first, and then select the combination of "Mode A + C" or "Mode B + D" to do the color test according to the analysis result.

2.2 The Design of Experimental Scene

The experiment was carried out indoors, simulating the real-life bus-waiting scenario. First we selected three different bus stops and recorded bus videos according to the direction, the position and the range of viewing angles of waiting passengers, and then replaced the front part of the bus in the videos with our new design of bus faces by video software as the experimental material.

2.3 Experimental Subjects

The experimental subjects are postgraduate students, including 10 males and 10 females, a total of 20 people. Aged 22–25 years old, with eyesight all above 0.8 and no color blindness or color weakness.

2.4 Experimental Equipment

Eye tracking device, digital video, video software, timer, and electronic computer. The eye movement data is output by professional eye movement data software, and analyzed by BeGaze and SPSS data analysis software.

2.5 Experimental Procedure

The participants were instructed to stand in position in accordance with the real-life bus-waiting scenario in the simulation studio, eyes at the corresponding height and appropriate angle with the "bus face" and a distance of 50 ± 2 cm with the display. Then the experimental staff will help the subjects wear eye tracking device and do the angle adjustment and calibration.

After the preparation process is completed, the experimental instructions will be issued by a staff and the experimental procedure will be recorded. After the experiment, the data of the subjects will be preserved for further analysis.

3 Data Analysis

3.1 Layout Analysis

The timeline of the video is divided into two parts according to the time when the subjects identify the target bus. The period from the bus body appearance to the success of subject's identification is identified as "pre-recognition" while the period after subject's recognition is called "post-recognition". Then record the eye movement data of the subject's visual search trajectory under Mode A and Mode B respectively, and the statistical results are shown in Fig. 2 and Table 1.

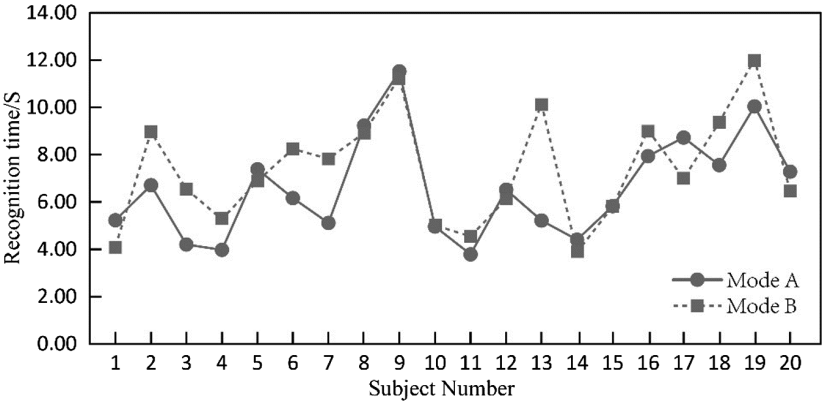


Fig. 2. Comparison line chart of recognition time between Mode A and Mode B.

Table 1. Velocity and amplitude of saccade of Mode A and Mode B.

Group	Velocity of saccade/((°) s^{-1})		Amplitude of saccade/(°)	
	Average number	Standard deviation	Average number	Standard deviation
Mode A	203.70	10.59	10.79	1.66
Mode B	202.90	9.29	10.52	1.56

Comparing the length of recognition time (The period from the bus's appearance to the success of subject's recognition) of Mode A and Mode B (Fig. 2), we find out that the recognition time of Mode B is higher than that of Mode A, and its volatility is greater. The average recognition time of Mode A is 6.59 s while that of Mode B is 7.37 s, which is 11.84% higher than that of Mode A. Although some studies have pointed out that people's perception of visual information has the characteristics of series processing, which means, it follows certain orders both chronically and spatially [3]. However, most bus information currently is located on the top-middle part of the bus face rather than the top-left part, which leads passengers to have developed an observation inclination in the long-term bus waiting experience, that is, habitually searching the bus information located in the top-middle part. At the same time, people need to browse continually to obtain their target information [4]. So when the bus of Mode B occurs, passengers will take some time to adapt the layout change as the information is on the left now, which explains why the recognition time of Mode B is longer. And the record of eye movement and later interview survey also have verified this point.

As we can see from the Table 1, the average number and standard deviation of velocity of saccade in Mode A is higher than that of Mode B by 0.39% and 13.99% respectively. Also the average number and standard deviation of amplitude of saccade in Mode A is higher than that of Mode B by 2.57% and 6.02%. According to the theory of cognitive load, the greater is the cognitive load, the smaller are the velocity and amplitude of saccade. Similarly, the smaller is the cognitive load, the greater are the velocity and amplitude of saccade. How to reduce the cognitive load is a key issue in

effective cognition and learning [5]. And the uniqueness of visual information content will facilitate the occurrence and change of saccade activities [6]. The velocity and amplitude of saccade in Table 1 can reflect the size of cognitive load of the waiting passengers in two different modes. The cognitive load of Mode A is smaller because there is less modification in the layout of bus information. So passengers can get the bus information successfully with their previous habit. But the cognitive load of Mode B is bigger as the target information is distributed in the left part, whose location is against passengers' inclination. And in this way, passengers in the initial observation will appear short-term inadaptation and start to re-search, so the cognitive load will improve.

In Mode B state, the recognition time is longer than Mode A's, and the cognitive load also improves. These reflect the impact of changes in the layout of bus face to passengers. From the above analysis, we can get the conclusion that the layout design of Mode A is more suitable to passengers' visual search habits.

3.2 Text Color Analysis

The text on the bus is an important visual element for information communication. In general, the text mainly includes bus line number, start station and terminal station and so on. In particular, line number on the bus face is usually the first information captured by passengers' eyesight, and it is very important for passengers whether they can get on the target bus successfully. In addition to the layout design, color of line number also demands careful consideration. From Sect. 3.1, it can be seen that the layout of Mode A is better than that of Mode B as it can help passengers to identify the target bus quickly. Therefore, we choose Mode A and Mode C as the experimental analysis objects in this section because their layout designs of the bus face are the same. We analyze and compare the eye movement data, and the result is shown in Fig. 3.

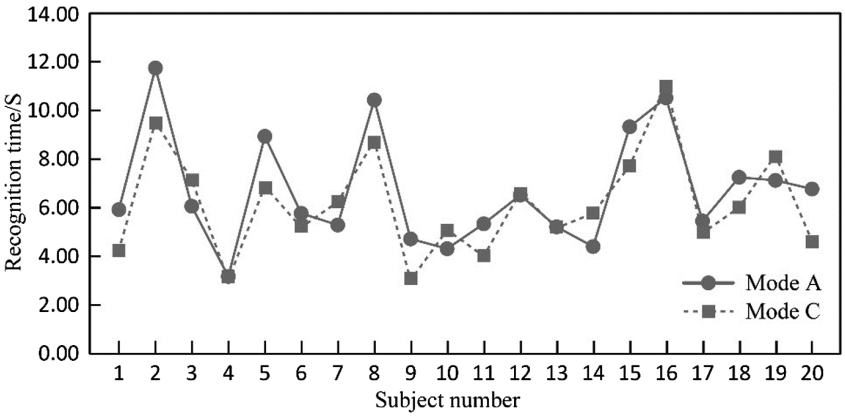


Fig. 3. Comparison line chart of recognition time between Mode A and Mode C

As we can see from Fig. 3, the recognition time of Mode A is higher than that of Mode C, and its volatility is bigger. In the state of Mode C, subjects' recognition time

is reduced by 8.91% compared with Mode A's. In general, brightness, purity and location are the main factors that affect whether a certain color can be effectively identified. High-brightness color has a sense of prominence that is conducive to identification while low-brightness color is not easy to be identified with a sense of obscurity [7]. The color of the line number in Mode C is orange that has high wave length and recognition rate, which is a kind of high-brightness color. However, the color of the bus line information such as start station on both sides of the line number is green, of which brightness and purity are relatively low. In this way, the sense of prominence of the line number increases as orange and green create a contrast, attracting passengers' attention more effectively. So the recognition time of Mode C is lower. And the colors of the line number, start station, terminal station in Mode A are all green, so the recognition time of it is higher than Mode C as the sense of prominence and the contrast are not strong.

4 Conclusion

Based on the previous research, we redesign the bus face, then use the eye movement equipment and data analysis software to arrange and analyze the subjects' visual search trajectory in different modes. And the research results are as follows: Although it is conducive to passengers' observation to put the line information on the upper left side of the bus face when there are more than one bus enter the station, passengers in the waiting process still tend to focus on the upper middle part at first; The addition of the line number on the lower part of the bus face is of help to passengers' recognition; Passengers in the visual searching process will focus on the part with high-brightness and high-purity at first, and it can help passengers to identify the target bus quickly by increasing the sharp visual contrast of the line information. With the results of this investigation, further research will focus on further optimization of the bus face design.

References

1. Liu, M., Xiong, T., Guo, J., Xiao, M., Bei, D.: Eye saccade research for waiting visual search and the characteristic of bus body design in Chinese. J. Univ. Shanghai Sci Technol. **5**, 472–478 (2016)
2. Yan, H.: The study of interpreting mental state based eye movement trajectory using SVM. Dissertation, Taiyuan University of Technology, Taiyuan (2010)
3. Han, Y.: The order of eye movement when observing different shapes and colors. J. Psychol. Sci. **20**(40–43), 95–96 (1997)
4. Sternberg, S.: Memory-scanning : mental processes revealed by reaction-time experiments. J. Am. Sci. **57**(4), 421–457 (1969)
5. Pi, Z.: The presenting formats of video podcasts effect on learning effectiveness and mechanism: an eye tracking study in Chinese. Dissertation, Central China Normal University, Wuhan (2014)
6. Yang, S.N., Mcconkie, G.W.: Eye movements during reading: a theory of saccade initiation times. J. Vis. Res. **41**(25/26), 3567–3585 (2001)
7. Zhang, H.: Study of feasible program about decoration design of buses' appearance in Changchun city in Chinese. Dissertation, Northeast Normal University, Changchun (2012)

Vision and Driving Support for Shielded Vehicles – Implementation and Test of an Electronic Vision Replacement System with Augmented Reality

Daniel López[1(✉)], Marcel C.A. Baltzer[1], Christian Lassen[1], and Frank Flemisch[1,2]

[1] Fraunhofer FKIE, Fraunhoferstrasse 20, 53343 Wachtberg, Germany
{daniel.lopez.hernandez,marcel.baltzer,christian.lassen,
frank.flemisch}@fkie.fraunhofer.de
[2] Institute Für Arbeitswissenschaft, RWTH Aachen University, Aachen, Germany

Abstract. The compromises encountered in the design of shielded and armored vehicles regarding the trade-offs between outside visibility and crew protection often result in the decrease of situational awareness. Drivers using busy roads face the challenge of maneuvering large vehicles, which, due to their safety requirements, have large blind spots that make it difficult to detect elements that could pose a threat to the safety of the crew and the vehicle. This paper presents the implementation of a camera-monitor system used as an electronic vision replacement and its evaluation results from a test trial with soldiers from the German Bundeswehr. Additionally, elements of Augmented Reality were incorporated to further assist the driver while navigating through the simulated urban scenario.

Keywords: Vision replacement · Shielded vehicles · Augmented reality

1 Introduction

1.1 Background

One of the important aspects of safety is that of Situational Awareness. A significant advantage and reduced danger can be gained by offering a better perspective of the surroundings to the occupants of a military vehicle. This is of particular importance when considering an urban scenario that requires maneuvering an oversized vehicle through narrow roads filled with vehicles, rubble or other obstacles. The driver can be easily overwhelmed by performing the normal driving tasks, as well as looking for potential threats in the surroundings. Furthermore, the protective armor that is desired for increasing occupant protection hampers the crew visibility by reducing the size of windows and hatches. An option for increasing the situational awareness of the occupants, as well as helping with the maneuvering of the vehicle, is using a vision replacement system that can offer the driver an unobstructed view of the outside. Additionally, such system can be enhanced by using augmentation for highlighting the vehicle dimensions and its position on the road, as well as the marking of obstacles.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_49

1.2 Vision Replacement

Several solutions on the market help drivers to view beyond their line of sight. Most of which rely on a monitor that displays the different views from cameras mounted on the outside. As such, these systems are intended as driving aids since drivers still need to use the small windows around them.

Vision support and vision replacement systems aim to improve situation awareness by presenting the driver with more information from what would normally be available to them. In the case of vision support, mirrors and camera-monitor systems give extra views of what drivers can already see from their position as well as those areas that are out of the line of sight, considerably reducing blind spots. Vision support systems have been available in the market for some years in the form of rear parking cameras, night vision or more recently, bird view systems in some automobiles.

Vision replacement systems, on the other hand, do not require the driver to see directly through a window making complete armor coverage possible. Instead, cameras mounted on the outside display an unobstructed view on a monitor or a Head Mounted Display (HMD). Beside the higher level of protection, vision replacement systems make the use of Augmented Reality possible to further enhance perception.

2 Camera-Monitor System Implementation

In order to investigate the usability of a camera-monitor system as a vision replacement system for driving, an experiment was conducted together with the German Bundeswehr. In this experiment, the mentioned camera-monitor system was installed on a real vehicle and tested on a closed track. This experiment aims to expand the findings of a previous project that focused on the comparison of different solutions for vision replacement where the camera-monitor system was evaluated against a Head Mounted Display for Virtual Reality [1]. Here, the camera-monitor system showed that while it is a simpler implementation of a vision replacement system, it avoids some of the most common problems found in HMD's; far from eye displays makes it possible to use lower resolutions which in turn increases framerate. In addition, since the system is directly implemented into the cabin, drivers are not required to wear any devices that can interfere with their tasks. Furthermore, and HMD as the one tested, blocks completely the vision of the other elements inside the cabin such as dials and controls potentially disorientating the drivers.

2.1 Panorama Stitching

Since the objective is to offer the driver an unobstructed view of the front of the vehicle, a 180° panoramic view needs to be generated. A large field of view can be achieved using fish-eye lenses [2] or with omnidirectional cameras [3, 4] but the large distortion makes them unfavorable. In order to avoid such lens distortions, multiple-wide angle cameras are used and the camera feeds are stitched together. The number of cameras and their field of view depend predominantly on the required field of view of the panoramic view and the computation speed. As we add more cameras to the system, the

number of images to be stitched together increase, which in turn elevates the computational cost. A system with two cameras has the least computational cost in a multi-camera system. On the other hand, both the cameras are required to have higher field of view to compose a panoramic view of 180°. This leads us to opt for a three cameras system, each having a horizontal field of view of ~90° with overlapping field of view illustrated in Fig. 1.

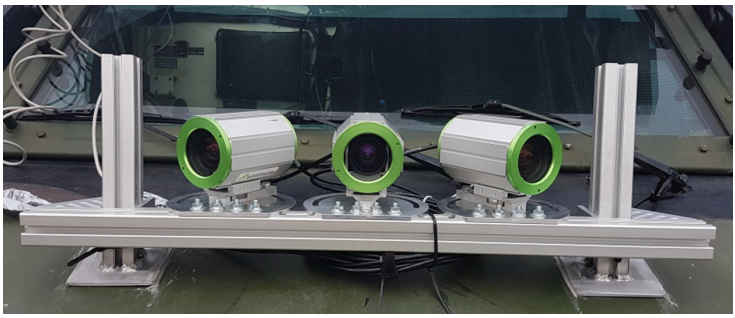


Fig. 1. The cameras were fitted with lenses with ~90° of field of view (FOV). The mounting angle was 45° between each camera. In this arrangement, the cameras covered a FOV of ~180°.

The camera in the middle points towards the front, whereas the other two cameras are directed towards the two sides. The image from the middle camera is considered as the reference frame and the images from left and right cameras are stitched to the left and right sides of the reference frame to create the final panorama. The required overlap between the field of view of the adjacent cameras has been empirically determined as 40°. The final panoramic view is presented on a five-monitor system in the driver's cabin.

The entire process can be divided into two stages: offline and online stages. Offline phase performs the system calibration and computes the lookup tables (LUTs) required for the online stage. The online stage is responsible for capturing the images and composes the panoramic view of the vehicle surrounding in real-time.

During the offline phase, the camera parameters are determined. These parameters are used in the camera calibration step to remove the lens distortions. The key points are extracted from the overlapping regions of the undistorted adjacent images. The extracted key points from the adjacent images are matched to obtain image correspondence. The latter is used in the next step to estimate the homography mapping of one image to the other. Using this homography, one image can be warped so that both images are on the same coordinate system. This warping results in large distortion near the edges of the warped image, which is reduced by a perspective warping. Finally, the pixel mapping required for the image transformation to the panoramic view are computed and stored in LUTs. The cameras are fixed firmly to avoid the displacement caused due to the vehicle movement. Recalculation of the offline stage is only necessary when the positions of the cameras or the camera parameters change.

On the other hand, the online phase captures the camera frames and applies the pixel mappings from the LUTs to compose the panoramic view. The mapping encoded by the LUTs includes the image undistortion and warping. In the above stitched panoramic

view, some undesired artifacts can appear around the border between the adjacent images. Blending techniques are used to remove the visible seam in the panorama.

In addition to the panoramic view, rear cameras are incorporated in form of a Picture-in-Picture solution in the front left and front right screen (Fig. 2).

Fig. 2. Rear view as Picture-in-Picture in the front right screen

2.2 Augmentation

Besides the increased range of view provided by vision replacement systems, the use of Augmented Reality gives the opportunity to overlay information relevant to the situation directly on screen. This way, obstacles, road signs, pedestrians or other vehicles that haven't been seen by the driver can be highlighted in a non-obstructive way. Furthermore, relevant information about the mission such as route, objectives or targets can be displayed. Augmented reality is not new in the military domain; it has been used for already many decades in fighter airplanes, where other aircraft are tracked and highlighted on the HUD and as an aid for refueling or targeting systems. On ground vehicles, on the other hand, the integration is more difficult due to the high cost, complexity and high computational demands.

As a first step towards more sophisticated Augmented Reality (AR), a pattern detection of objects was conducted. Those patterns focused on vehicles and pedestrians (Fig. 3).

Fig. 3. Left: Marking of a detected inflatable "car". Right: Marking of a detected pedestrian [5].

For vehicle detection, the procedure of boosted cascade of simple features [6] was used. For pedestrian detection, the procedure of Histograms of oriented gradients (HOG) [7] was applied.

To improve situation awareness of the driver concerning the position of the vehicle in its environment, a virtual hood was displayed in the 5-monitor setup (Fig. 4).

Fig. 4. Virtual hood (in blue) in field trial [5].

3 Vehicle Platform and Installation

As a test platform for the field evaluation of the system, a shielded personnel carrier TPz Fuchs of the German Bundeswehr was used. The relatively spacious cabin facilitated the mounting of the monitor array besides offering good options for powering the needed equipment.

The monitor array consists of five 13″ LCD 720p monitors that cover 160° of the driver's horizontal field of view (Fig. 5). The monitors displayed the panoramic stitched image of the front of the vehicle captured by three Full HD GigE cameras fitted with wide-angle lenses capable of capturing a field of view of nearly 90°. The wide-angle lenses were required in order to compensate for the image lost as part of the panoramic stitching process.

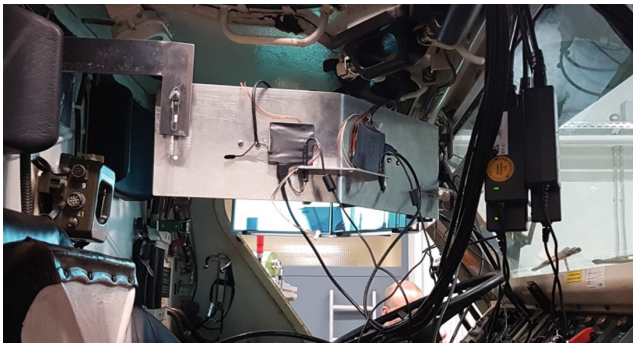


Fig. 5. The five monitors surround the driver covering 160° of field of view.

Figure 6 Left shows the placement of the cameras on the hood of the vehicle. When mounted on the center of the hood, the cameras had good coverage of both sides. As for the mounting high, a compromise had to be made so that the resulting view offered a good perspective of the surroundings and of the road right in front.

Fig. 6. Left: Mounting of the front cameras on the test vehicle. Right: Mounting of the camera for the right rear view mirror.

Since the installed monitors prevented all visibility of the driver to the outside, two extra cameras were used to replace the outside mirrors. The mirror views were incorporated into the middle left and middle right monitors. The cameras for the mirrors were installed on each side just behind the door.

These cameras and lenses are the same type as those used for the front view. Since the image from the camera was directly displayed without any processing, the wide-angle image was retained which gave the driver a better view than that offered by the normal mirrors (Fig. 6 right shows the mirror cameras as they were installed).

Two computers located in the back compartment of the vehicle controlled the cameras and monitors.

This ensured that only low voltage wiring reached the driver cabin as a safety measure. The video from the cameras was captured by the main computer, this was then processed, stitched and displayed to the monitors. A second computer was used for the augmentation elements. Here the outline of the vehicle limits was generated as well as a prototype for object recognition. This information was then transmitted to the Main PC for its display.

Figure 7 shows the communication lines between the elements of the camera-monitor system as well as their location in the vehicle. The frontal and side arcs represent the field of view achieved by the camera combination.

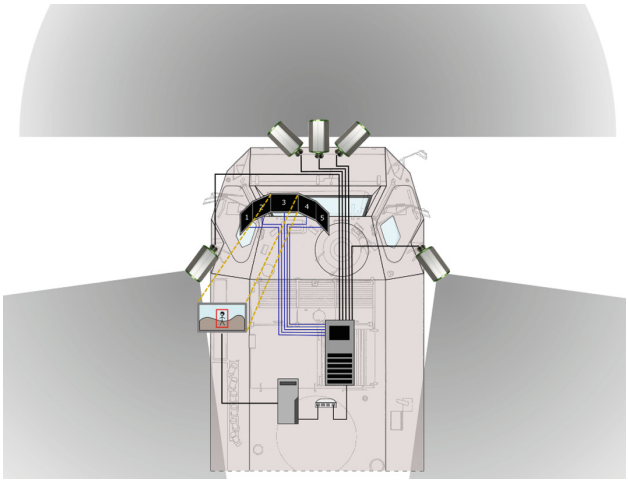


Fig. 7. Representation of the camera-monitor system components and the location in the test vehicle.

4 Evaluation Process

For the evaluation of the camera-monitor system, the equipped vehicle was driven on a close area that simulated an urban scenario. The testing area contained several buildings separated by narrow roads. Additional objects were used to simulate common scenarios found in normal driving through a city. These included a traffic light, a stop sign and a roundabout. A fixed route was determined so that all participants would encounter the same situations. Each participant tried two different modalities of the system, a version with only the image from the cameras and a version with the augmentation.

After an explanation about the procedure of the evaluation, the test participants were given different questionnaires, which gather demographical details and the physical condition. This was followed by a short instruction on the vehicle about the technical details of the experiment setup and safety regulations. In addition, the brightness level of the cameras was adjusted to the current weather conditions. Afterwards two trial runs were completed. Every participant drove with both versions of the vision replacement system. The order of the condition was randomized. For each run, a training phase of approximately 5 min each was performed for the familiarization with the setup. Each version of the system was evaluated in regards to its usability, intuitive interaction, physical influences and perceived workload. Eye tracking data was also recorded. The used questionnaires were NASA Task Load Index (TLX), INTUI, simulator sickness questionnaire (SSQ) and system usability scale (SUS). In a conclusive interview, the strengths and weaknesses of the systems together with preferences and needs were discussed with the participants.

5 Results

5.1 Sample

Twelve participants drove through the fixed route. All were males and members of the German Bundeswehr. The average age was 39 (Standard Deviation (SD): 9.4). All participants had a license for car, truck and armored vehicles. Ten participants described their driving behavior as confident/experienced, two as cautious.

5.2 Results from Questionnaires and Interviews

The perceived workload, which was measured with the NASA-TLX, was evaluated for significant differences with a t-test for related samples. The analysis showed slight tendencies, which were not significant because of the small sample size (Figs. 8, 9, 10, 11, and 12).

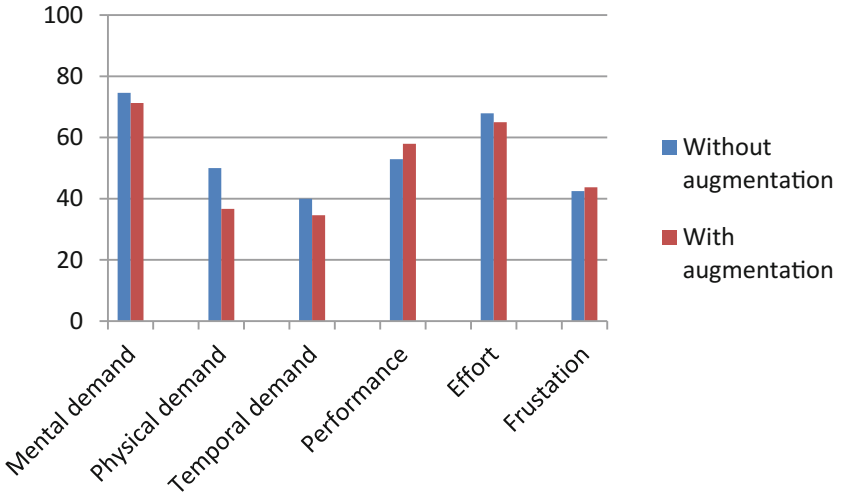


Fig. 8. Perceived workload comparison between the two system versions.

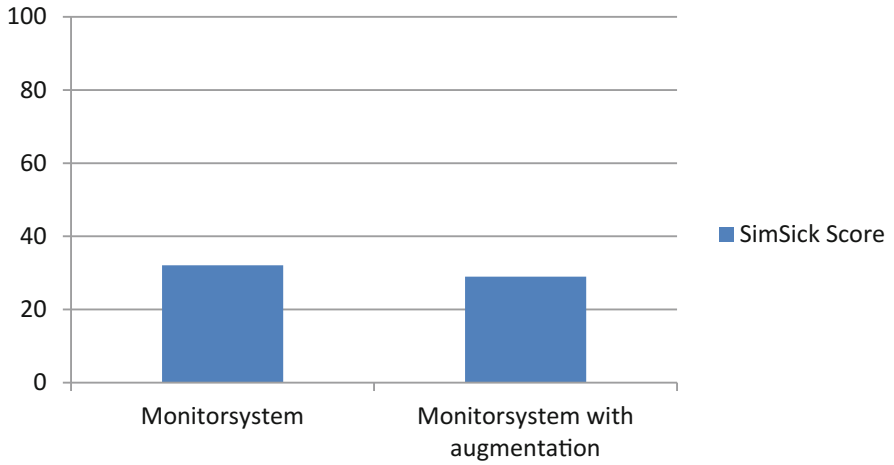


Fig. 9. Simulation sickness results show no significant difference between both system versions.

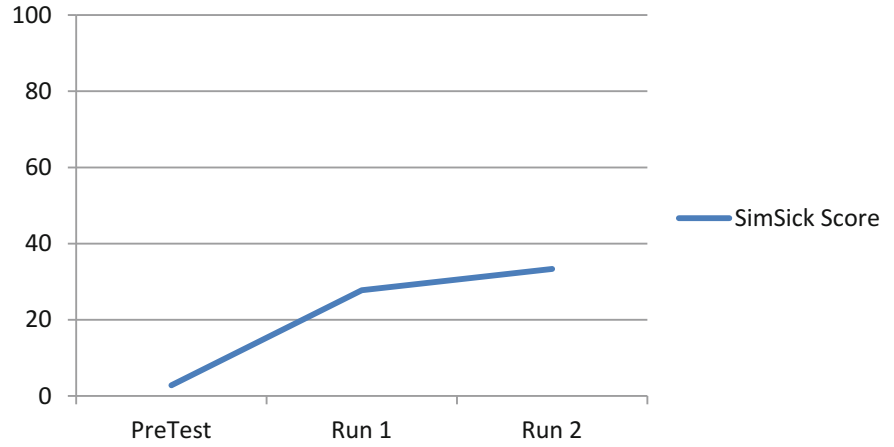


Fig. 10. Simulation sickness had a tendency of increasing over time.

The analysis of the SSQ showed no significant effect regarding the use of the different system versions. The sickness was perceived as low in general. The temporal process had an influence on the sickness. The ratings before and after the first run differ clearly. A significant effect was discovered with $F(1.272, 22) = 8.709$, $p = 0.07$. The ratings stayed nearly even after the second run (Tables 1 and 2).

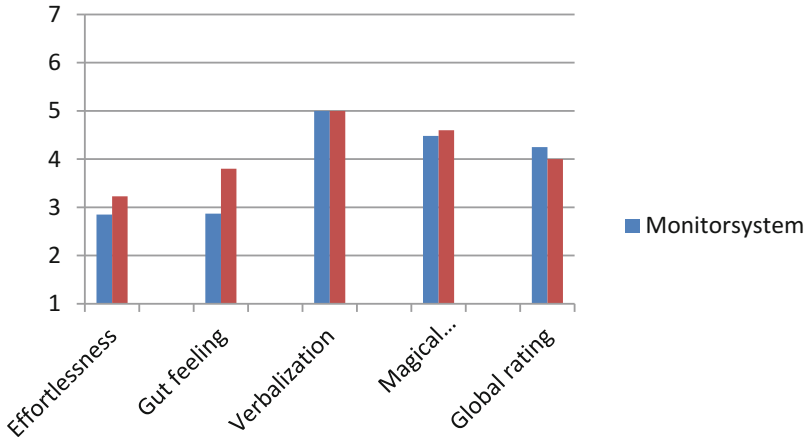


Fig. 11. Usability comparison between the two system versions.

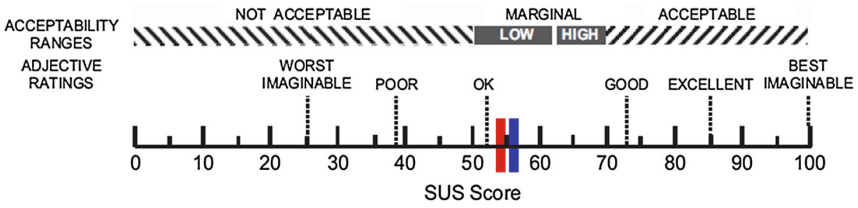


Fig. 12. Both systems received a marginal acceptability score with an OK rating.

Table 1. Results from the NASA-TLX analysis.

	Without augmentation M (SD)	With augmentation M (SD)
Mental demand	74.58 (14.72)	71.25 (22.49)
Physical demand	50.00 (28.92)	36.67 (24.72)
Temporal demand	40.00 (28.44)	34.58 (30.33)
Performance	52.92 (27.26)	57.92 (23.93)
Effort	67.92 (21.04)	65.00 (23.93)
Frustration	42.50 (25.98)	43.75 (29.09)

Table 2. Results from the usability questionnaire.

	Without augmentation M (SD)	With augmentation M (SD)
Effortlessness	2.85 (1.23)	3.23 (1.09)
Gut feeling	2.87 (1.07)	3.80 (0.83)
Verbalization	5.00 (1.43)	5.00 (1.43)
Magical experience	4.48 (0.98)	4.60 (0.99)
Global rating	4.25 (1.14)	4.00 (1.28)

The interaction was rarely evaluated as effortless, which means the task required a high demand of attention. The process of intercation was logical for the participants. In general the interaction was perceived slightly as fascinating. Both versions were evaluated as moderately intuitive. No significant effect could be found.

A score was calculated from the results of the SUS Questionnaire, which can be bewteeen 0 (worst application) and 100 best application. The monitor system with augmentation attained a value of $M = 56.36$ $SD = 17.97$. The score of the monitor system without augmentation was $M = 54.32$ SD = 15.04. There is no significant effect between the two systems, which means that they were perceived as equal in their usability.

In the final interview, strenghts and weaknesses of the systems were discussed with the participants. The camera-monitor system offers a better all-round visibility and provides no dead angles. Also no excessive head movements are necessary. The displayed mirrors were evaluated as very good and helpful for the driving. The disadvantage of the monitors were the reduced perception of distances. Also it was mentioned, that the monitors are getting more fuzzy to the side. The participants had the impression of a blurred and distorted vision. Some participants had difficulties to distinguish the road surface, for example, paved road from grass. For some participants, it was confusing that the seating position was on the left side of the vehicle, but the monitor system showed the centred view. The augmentation was helpful, especially for driving around corners; it was described as a good help for orientation. Some participants criticized that the representation of the hood was too inaccurate and for that reason confusing. Different solutions were mentioned to solve the shortcomings of the system. It would be helpful to display a distance recognition like the visualization of a parking aid. Also an auto brightness adjustment should be integrated. The measurement of the vehicle and trajectory could be indicated with an augmentation. The angle of the mirrors should be adjusted manually. Additionally, the vehicle speed and engine RPMs could be integrated in the monitors.

5.3 Eye Tracking (DL)

In addition to the results from the questionnaires and interviews, eye tracking was used to expand the understanding of using a surrounding monitor arrangement for driving. Tobii Pro 2 Glasses were used since their light construction and wireless connectivity had little interference with the driving tasks.

Figure 13 shows a heat map of the combined data of the entire runs of all participants. The heat map represents the distribution of the total fixation points. In this case, a high count of the combined fixation points is on the center of the middle monitor, area that is marked with a more intense color. As the count of fixation points decreases to the edges of the images, the color also fades. It is important to note that the images on the screens do not represent any particular driving situation; they are used for the mapping of the fixation points only.

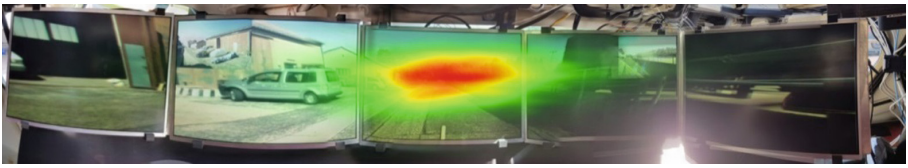


Fig. 13. Eye tracking heat map of combined data from all participants.

6 Discussion

The performed experiment confirms some of the benefits of an electronic vision replacement system. The camera-monitor system made possible driving without a direct view of the outside. The possibility to have a panoramic and unobstructed view was well received by the participants who also welcomed the addition of the mirrors on the screens. Regarding the quality of the displayed view, the image was compact and clear in the center monitors, but as the image gets closer to the outside edges, the distortion and blurriness makes outside monitors less usable if objects are closer. In our experiment, this effect did not have a major impact on the participants while driving since most of the time, the wide-angle lenses on the cameras made possible to display the total width of the road in the center monitor. This is beneficial since the drivers are required to perform less head movements. This was also confirmed by the data obtained from the eye tracking glasses where heavy use of the center monitor was evident. An issue that arises when displaying images capture by cameras onto a monitor is the loss of depth. This conversion of a 3D to a 2D perspective made difficult for the participants to correctly judge the distance of objects ahead and the position of the vehicle. This was partially compensated by the use of the augmentation that highlighted the limits of the vehicle, particularly when cornering. Other forms of augmentation such as the explicit indication of the distance to certain objects can minimize this negative effect. Regarding the augmentation for object recognition, this was not used in this experiment during driving since it was a prototype in early stages. Nevertheless, it was demonstrated to the participants who mentioned that they would welcome its addition. Although the overall score in acceptability and usability is positive, some areas need attention. As previously mentioned, the side monitors need improvement in sharpness and less distortion. Similarly, manually adjusting the brightness was not ideal, in this case, a camera and lens combination with automatic brightness adjustment is recommended.

Overall, the outcome of the experiment shows that vision replacement systems offer good potential for increasing the situational awareness while keeping a high level of armor protection. As for technical limitations such as image quality and framerate, they can only be expected to improve, as better technology is available.

References

1. Baltzer, M.C., Krasni, A., Boehmsdorff, P., Linder, T., Heesen, M., Flemisch, F.: Towards virtually transparent vehicles: first results of a simulator study and a field trial. In: AVT-256 Specialists' Meeting on Augmented Reality for Improved Situational Awareness and Survivability of Combat Vehicles, NATO STO, Prague (2015)
2. Corp, A.: Generating Panoramic Views by Stitching Multiple Fisheye Images. Tech. Rep WP-01107-1.0, San Jose, CA (2009)
3. Huang, K.S., Trivedi, M.M., Gandhi, T.: Drivers view and vehicle surround estimation using omnidirectional video stream. In: IEEE Intelligent Vehicles Symposium (2003)
4. Gandhi, T., Trivedi, M.M.: Vehicle surround capture: survey of techniques and a novel omni-video-based approach for dynamic panoramic surround maps. IEEE Trans. Intell. Transp. Syst. **7**(3), 293–308 (2006)
5. Baltzer, M., Lassen, C., López, D., Madhogaria, S., Gonzalez, J., Bender, D., Flemisch, F.: Sicht- und Fahrunterstützung für Geschützte Fahrzeuge 2, Final Report (2017)
6. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: IEEE International Conference on Image Processing (2001)
7. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), (pp. 886–893) (2005)

Measuring the Cognitive Demands of In-vehicle Dashboard and Centre Console Tasks

Madeleine Conaghan[1(✉)], Ian Colwill[1], Edward Elton[2], Periklis Charchalakis[1], and Elias Stipidis[1]

[1] Vetronics Research Centre, University of Brighton, Moulsecoombe, Brighton BN2 4GJ, UK
M.Conaghan1@uni.brighton.ac.uk,
{i.colwill,P.Charchalakis,E.Stipidis}@vetronics.org
[2] Brighton Product Lab, University of Brighton, Moulsecoombe, Brighton BN2 4GJ, UK
E.Elton@brighton.ac.uk

Abstract. The cognitively demanding dashboard and centre console tasks that drivers face while operating a vehicle were assessed through a lab based study. Thirty-four participants drove a simulated course for approximately 20 minutes whilst completing a series of secondary tasks. During the simulation, continuous assessment of driving performance, reaction time and indicators of cognitive burden were collected. Following the driving exercise, participants were asked to complete a survey for self-assessment of secondary task difficulty. The survey results reveal that drivers find using the satellite navigation system as the most challenging and most distracting task.

Keywords: In-vehicle tasks · Cognitive demands · Vehicle ergonomics

1 Introduction

There has been a 257% increase in in-vehicle system features over the past 40 years, and with implementation of new technological advancements this number is predicted to rise to 329% over the next 15 years [1]. When that data is combined with accident statistics, it shows a potential correlation with the 11% increase in the amount of people involved in crashes from 1995–2010 [2]. The highest reported reason for traffic accidents in 2013 was due to driver error [2]. 'Driver error' can be categorised under skill based errors, decision errors or perceptual errors [3]; with each error category using descriptions such as wrong response, exceeded ability and misjudged [4]. These all point to cognitive workload as the main cause of error. As cognitive workload increases, it creates a diversion of mental resources while driving thus increasing the crash risk [5].

Although the amount of information a driver is expected to process has increased, the capability of the human brain has not. Miller (1956) suggested that there is a constant capacity hypothesis of 7 ± 2 chunk of information at any one given time [6]. Since 1956, many theorists have questioned Miller's theory, but have concluded humans can only process between 5–9 chunks of information through their working memory [7]. Therefore, human capability is still at a maximum of nine chunks of information, but the amount of information a person is expected to comprehend from a car whilst driving is

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_50

evolving and exceeding this limit theory. A research study was conducted in a laboratory environment to fully understand the extent of this problem. This study looked to establish how demanding a driver perceives the secondary tasks undertaken while driving are.

2 Method

The driver controlled the driving simulation through a real-world car, and was then asked to complete a series of tasks which the participants then rated how demanding and challenging they found the tasks were. Drivers participated in the study for 20 min to allow them to immerse themselves in the environment. This methodological approach offered the greatest insight into the driver's perception of secondary task difficulty while driving, and allowed subjective qualitative survey data to be collected and analysed quantitatively. Thirty-four participants took part in the research study and provided insight into what Dashboard and Centre Console tasks they found demanding.

2.1 Driving Simulation

The participants drove around a 3D simulated track using a 2016 ford fiesta passenger vehicle for control input. The virtual simulated environment was projected on large screens situated around the vehicle as seen in Fig. 1. The vehicle simulation allowed the driver to experience a realistic driving scenario and become submerged in the experiment.

Fig. 1. This image shows a participant driving the vehicle simulator through the Ford fiesta.

The simulator delivered audio instructions at specific locations around the course that asked the participants to complete a range of secondary driving tasks detailed in Table 1. The drivers completed nine secondary tasks while operating the simulator that they would assess for comparative difficulty during a post drive subjective review. Each driver completed each of the tasks three times.

Table 1. This table shows the tasks completed in the simulator in bold, all other tasks included in the review and the category they each hold.

Task	Task category
Responding to vehicle alerts	Dashboard glance
Looking at the fuel level	Dashboard glance
Looking at the speed dial	Dashboard glance
Changing the air conditioning	Multimedia interaction
Making a call through the car	Multimedia interaction
Changing the music	Multimedia interaction
Changing the volume	Multimedia interaction
Using the multimedia centre	Multimedia interaction
Using reversing camera	Multimedia glance
Using satellite navigation	Multimedia glance
Looking in the mirrors	Mirror glance
Using steering wheel buttons	Side of wheel interaction
Putting the windows down	Side of wheel interaction
Turning on/off the headlights	Side of wheel interaction
Putting window wipers on/off	Side of wheel interaction
Using indicators	Side of wheel interaction

2.2 Subjective Review

After completing the driving tasks, participants completed a subjective review on the relative difficulty of secondary in vehicle tasks. Sixteen individual secondary tasks were presented to the participants and they were asked to subjectively rate them. Nine out of the sixteen tasks were tested in the simulator and the participants were asked to provide review on the other seven tasks based on their own experience. The tasks that were completed within the vehicle are shown in bold in Table 1, they included Responding to vehicle alerts, Looking at the fuel level, Looking at the speed dial, Changing the air conditioning, Changing the music, Changing the volume, Using the multimedia centre, Looking in the mirrors and Using indicators.

The participants were asked to rate the tasks listed in Table 1 on a 5-point Likert scale. Each task was rated twice; the first rating focused on how challenging the participant found the task (rated highly challenging to not challenging) and the second rating focused on the level of distraction a task caused (rated from highly distracting to not distracting). The subjective evaluation allowed data to be gathered on drivers' perception of how demanding these in-vehicle interactions were. The subjective evaluation was

undertaken immediately after the participants drove the simulator to allow the experience to still be clear in their minds.

The tasks were split into five vehicle task categories; dashboard glance, multimedia interaction, multimedia glance, mirror glance and side of wheel interaction. These categories were based on previous research outcomes that indicated tasks in these locations were the most frequently used within the vehicle. Certain task categories had more technology and functionality within their location and therefore have more tasks in those categories. The tasks that they were asked to rate within each of these categories included:

2.3 Participants

Participants were selected based on two variables; (1) had a full driving licence, and (2) had driven within the past year. All drivers were considered for the study and a random sampling method was used to include a varied sample of participants. Participants were recruited from within the University of Brighton (UK). Thirty-four participants, aged between 18–59 years old with a skewed mean age (mean = 26) took part in the study. The sample included 12 female and 22 male participants from different ethnic groups and abilities to reflect the diversity of the wider population.

3 Analysis Method

The survey results were analysed by collating the answers using the same Likert scale as the data was collected against. The likert scale included the ratings from 'highly challenging'- 'challenging'- 'ok'- 'not very challenging'- 'not at all challenging'. The total of the results for highly challenging, challenging and ok ratings within the Likert scale were combined to indicate where the highest demanding tasks were. The same scale was used for the distracting tasks, from highly distracting, distracting- ok- not very distracting- not at all distracting.

4 Results

The results show that the participants rated 'using the satellite navigation system' as the overall most demanding task. 90% of participants found using the satellite navigation system distracting and 87% of participants found using the satellite navigation system challenging. The participants exclaimed multiple reasons for this outcome such as keeping an eye on the road while trying to use the navigation system as difficult while trying to concentrate on the road. Another reason given was difficulty in concentrating on getting to the right destination whilst using satellite navigation and having to drive at the same time.

Making a telephone call through the car was found to be the second most distracting task as 88% of participants reported this as distracting. This was followed by 79% of participants reporting changing the music input distracting and 74% of participants reported using the multimedia centre as distracting.

The top four most distracting tasks all fall under the multimedia interaction category; suggesting drivers found this to be the most demanding location within the car. Further results can be seen in Fig. 2 that highlights other highly distracting tasks, which are also located in the multimedia centre locality. The participants reported taking their eyes off the road in order to use the somewhat complex multimedia centre functions as one of the reasons for this outcome. They also reported that they find it challenging to operate functions that require high level of precision and focus, while maintaining their attention on the primary driving tasks.

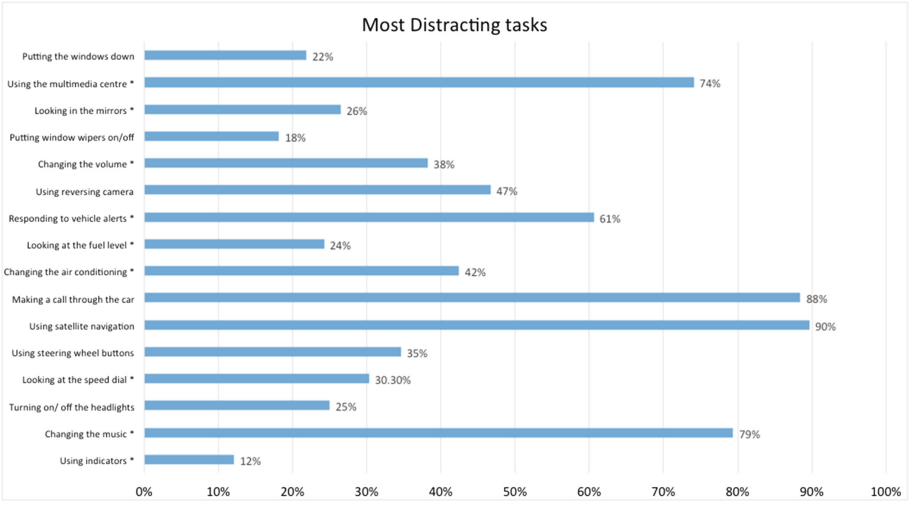


Fig. 2. This figure shows the tasks that the participants found distracting with an asterisk (*) next to the tasks that were completed within the simulator.

The participants rated the tasks from their own experience as more demanding than those that were conducted within the vehicle simulation. Of the tasks that were tested in the vehicle, 79% of participants found changing the music distracting, 74% of the participants found the multimedia centre distracting and 61% of participants found responding to vehicle alerts as distracting.

The research findings similarly revealed that 85% of participants found making a telephone call through the car as challenging, 71% of participants found changing the music as challenging and 51% of participants found responding to vehicle alerts and using the multimedia centre as challenging. Further results in Fig. 3 indicate other highly challenging tasks, which are also located in the multimedia centre locality. This shows that participants also found the multimedia interaction to be the most challenging as well as the most distracting.

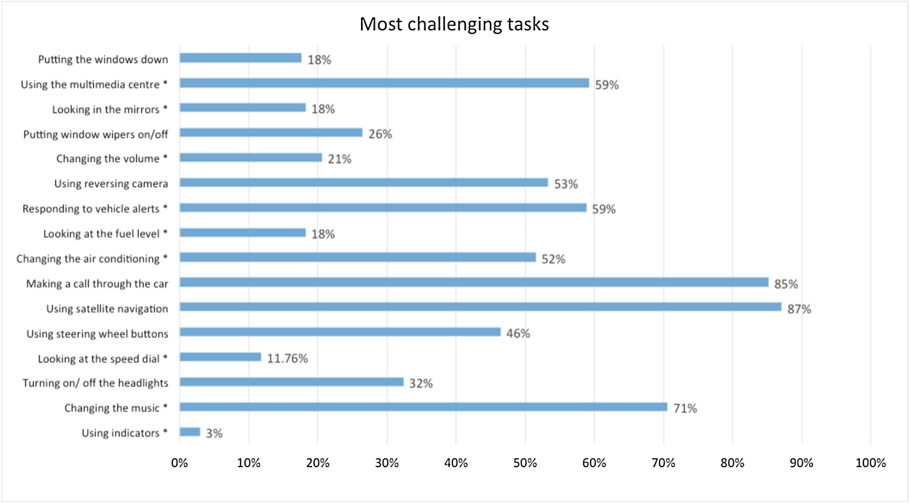


Fig. 3. This figure shows the tasks that the participants found challenging with an asterisk (*) next to the tasks that were completed within the simulator.

The participants also rated the tasks from their own experience as more challenging than the tasks rated based on those conducted within the vehicle simulator. The tasks that were conducted in the simulator that the participants found significantly challenging included; changing the music (71%), using the multimedia centre (59%), responding to vehicle alerts (59%) and changing the air conditioning (52%). All of these tasks take place in multimedia centre locality that suggests that based on both personal experience and the vehicle simulation experience, that the multimedia centre locality is the most challenging as well as the most distracting.

When comparing the results for the challenging tasks and the results of the distracting tasks, several differences were prominent. Interestingly, 19% of participants found looking at the speed dial more distracting than challenging, 18% found changing the volume more distracting than challenging and 15% found using the multimedia centre more distracting than challenging. There was a small difference between distracting and challenging for many of the other tasks as shown in Fig. 4.

After rating the task demands, the participants gave reasons for their answers. Within their stated reasons, there are five main trends, including; 25% of participants felt that the details of the task were too difficult; this included looking at precise details within the dashboard and centre console and using small buttons. Other reasons given for why they found the tasks challenging and demanding included; the location of the task (14%), the task being out of their direct eye line of sight (14%), having to divert their attention off the road (11%) and the task being out of direct reach (11%). Within each of these trends, answers such as struggling to find certain controls, dials that weren't centered or in reach and confusion around multiple button options.

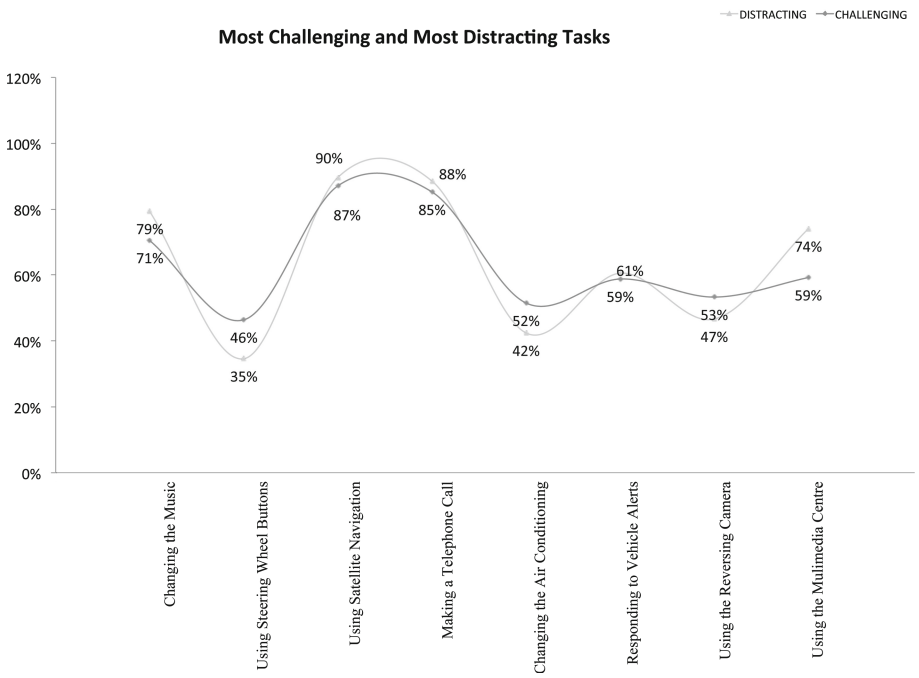


Fig. 4. This figure shows the difference between the challenging and distracting results.

5 Conclusion

The results provide an indicator of tasks that drivers perceive as demanding within a vehicle. The stated reasons given for perceived distraction or challenge provide insight into the problems that drivers face daily. It is clear that drivers may not find a task difficult but they are aware that it distracts them while driving. There is a clear opportunity for Designers to limit the distractions involved in completing a task within the car, especially as there are now more functions than ever [1].

Only 11% of participants found looking at the sped dial challenging where as 30% of participants found this task demanding. This 19% difference between drivers that found looking at the speed dial more distracting than challenging indicates that the task alone may not be difficult, but it is considered a distraction when combined with a complex primary task, such as driving. At the point of design, the way drivers use, comprehend and act on a task should be considered because conducting a secondary task concurrently with a complex primary task may be distracting.

An issue with the data may be because of the age range of the participants being disproportionately skewed (mean age = 26). This may be a confounding variable between the outcome and the participants own experiences with using technology. Amendments of this study would be to have a proportionate mean age of participants. Future work may include comparing the results of younger participants with older participants within the same study.

The study established that 90% of participants faced distraction from the satellite navigation system. Although this task was not analysed in the simulator prior to the review, this result shows drivers own experience with this tasks is particularly distracting. The significance of this data suggests that conducting this task while driving may have potentially detrimental effects. Interestingly, over the past 15 years, the increase in features, particularly satellite navigation systems have been more commonly included in vehicles [1], which may have a correlation with the 11% increase in human error related deaths over a 15-year period [2].

Overall, the tasks that participants found the most distracting were not undertaken in the vehicle simulator, however the tasks that were tested within the simulator were still considered a distraction. The outcome of tasks rated based on the participant's prior experience as opposed to rated on immediate experience from the simulator may be different, however the participants found both sets of tasks demanding. Future work may be undertaken after the driver has completed a wide range of vehicle tasks within a simulator (or on the road) prior to the subjective review of all of these same vehicle tasks.

References

1. M. Conaghan.: Analysis and comparison of Volkwagen Polo's from 1975–2015 and 2030 nils concept cars (2015)
2. Department for transport: National Travel Survey **2012**(30), 2013 (2013)
3. Reason, J., Manstead, A., Stradling, S., Blaxter, J., Campbell, K.: Errors and violations on theroad: a real distinction? Ergonomics 33(10.22), 1315–1332 (1990)
4. Shappell, S.A.: System–HFACS, The Human Factors Analysis and Classification, I.O. Aviation, Ed., University of Illinois at Urbana-Champaign, 2000
5. AAA Foundation for Traffic Safety: Measuring cognitive distraction in the automobile. Washington, June 2013
6. Miller, G.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. **63**(2), 81–97 (1956)
7. Cowan, N., Chen, Z., Rouder, J.N.: Constant capacity in an immediate serial-recall task a logical sequel to Miller (1956). Psychol. Sci. **15**(9), 634–640 (2004)

Cognitive Load Assessment of Tractor Driver Based on Cognitive Task Workload Model

Yeqing Pei, Xiaoping Jin(✉), Zhenghe Song, Haoyang Li, Ling Luo, and Bowen Zheng

Beijing Key Laboratory of Optimized Design for Modern Agricultural Equipment College of Engineering, China Agricultural University, Beijing 100083, People's Republic of China
664693979@qq.com, jinxp@cau.edu.cn

Abstract. Based on cognitive characteristics, this research aimed at quantifying a tractor driver's cognitive load (CL) during field work. First of all, the operation tasks under different operating states were analyzed hierarchically utilizing multiple resources theory. Then the tasks were further classified into 4 processing resources, i.e. visual, auditory, cognition, and psychomotor. The cognitive load during the tasks was further quantified by time occupied (TO), action-set switching (ASS) and level of action load (LAL). Finally, this method was used for evaluating the driver's cognitive load during pulping operation and seeding operation. The effectiveness of this method was validated in the driving simulation platform with a SMI eye tracker, while repeating the process during field work. The diameter of pupils and blink rate (BR) are used to characterize tractor driver's cognitive load.

Keywords: Cognitive load · Multiple resources theory · Tractor driver

1 Introduction

Cognitive load is a multi-dimensional concept describing worker's mental pressure or capability of information processing. It involves working requirement, time pressure, task complexity, operators' capability, operators' effort, etc. [1]. The most widely-used measurement methods of cognitive load are subjective evaluation, performance measurements and physiological assessment [2]. Optimal cognitive load is beneficial for performance and reliability of human-machine systems, and driver's health and comfort as well. Hence, it is important to study the evaluation of driver's cognitive load. However, most research on human-machine interface of agricultural machine cab focused on physiological load [3–5], neglected cognitive load. In addition, most research of cognitive load devoted to qualitative research [6]. This paper references cognitive task load model proposed by Neerincx [7], improves the unquantifiable elements. The improved model can be used to evaluate workers' cognitive load quantitatively. In the end, this method is used to measure the tractor driver's cognitive load during pulping operation.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_51

2 Method

Multiple resources theory indicates that people's processing resources are usually described by four components: visual (V), auditory (A), cognitive (C) and psychomotor (P) and are commonly referred to by the acronym VACP. These can be used to describe each individual behavior [8]. McCracken, Aldrich and others proposed rating scales for each VACP components [9–11]. It provides a rating of the degree to which each processing resource is used.

Our research combines multiple resources theory with time factor, which is a three-dimensional model. It includes three dimensions influencing cognitive load. The first dimension is time occupied (TO), which estimates the time to accomplish the task. The second dimension is level of action load (LAL), which describes driver's workload. The last dimension is action-set switching (ASS), which emphasizes on cognitive resources requirement of attention diversion. Driver's cognitive load state in a certain stage can be defined when the above factors are confirmed.

3 Investigation

To analyze the tractor driver's operation processes, we reference videos recorded by Luoyang Tractor Institute Co., Ltd. They install a driving recorder in the tractor cab during field work. Work scenario is selected as pulping operation and seeding operation, the driver is selected as a local peasant. The model of the tractor during seeding operation is Tractor A, and the model of the tractor during pulping operation is Tractor B.

Driving recorder A and Driving recorder B are installed in tractor cab during pulping operation. A is located at the left-rear corner of the cab to record driver's action with the left hand and left foot. B is located at right-rear of the cab to record driver's action with the right hand and right foot.

Driving recorder A, B, C and D are installed in the tractor cab during seeding operation. A is located at the left of the front windshield to record driver's action with the left hand and left foot. B is located at the right of the front windshield to record driver's action with the right hand and right foot. C is located at the rear windshield to record operation on right operation panel. D is located at the middle of the rear windshield to record the state of the farm implement.

4 Analysis

4.1 Task Analysis

Driver's operation processes obtained from video are hierarchically divided into goal level, task level and action level. Goal level demonstrates an entire goal in a specific time duration, consisting of several tasks. The task level indicates the tasks to accomplish a specific goal. Action level is defined as the operation behaviors in order to realize a specific task. Actions of action level are categorized and rated on a scale of zero to seven, as shown in Table 1 [12].

Table 1. Behavior parameters rating form

Visual(V)		Auditory(A)		Cognitive(C)		Psychomotor(P)	
No activity	0.0	No activity	0.0	No Activity	0.0	No activity	0.0
Visually register/detect	1.0	Detect/register sound	1.0	Automatic	1.0	Speech	1.0
Visually discriminate	3.7	Orient to sound	2.0	Alternative selection	1.2	Discrete actuation	2.2
Visually inspect/check	4.0	Orient to sound	4.2	Sign/signal recognition	3.7	Continuous adjustive	2.6
Visually locate/align	5.0	Verify auditory feedback	4.3	Evaluation/judgment	4.6	Manipulative	4.6
Visually track/follow	5.4	Interpret semantic content	4.9	Encoding/decoding, recall	5.3	Discrete adjustive	5.8
Visually read	5.9	Discriminate sound characteristics	6.6	Evaluation/judgment	6.8	Symbolic production	6.5
Visually scan/search/monitor	7.0	Interpret sound patterns	7.0	Estimation, calculation, conversion	7.0	Serial discrete manipulation	7.0

Pulping operation is divided into smashing cole straw task and scraping in the task level. The operation route is spiral. Driver drives tractor from the outer race of paddy fields to the inner race of paddy fields. When the driver arrives at the corner of paddy fields, he makes a left turn and reverses the tractor to pulp at the corner area. Once the edge of paddy fields is reached, the driver shifts the tractor into forward gear. Repeating again and again, until all field is pulped. Figure 1 shows task analysis of pulping operation specifically.

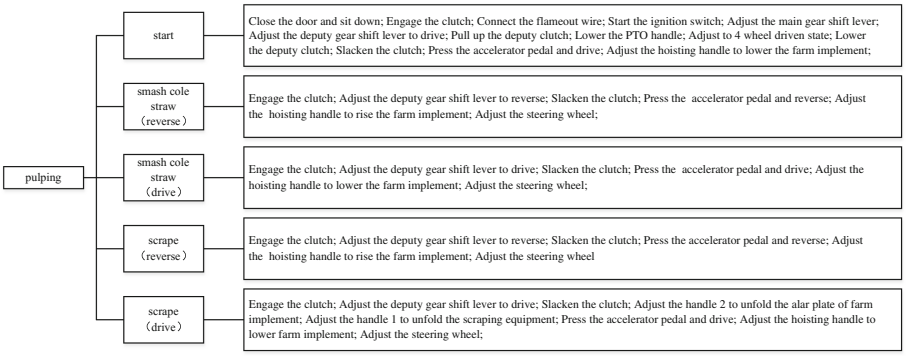


Fig. 1. Task analysis of tractor driver during pulping operation

Seeding operation's route is straight. Figure 2 shows task analysis of seeding operation specifically

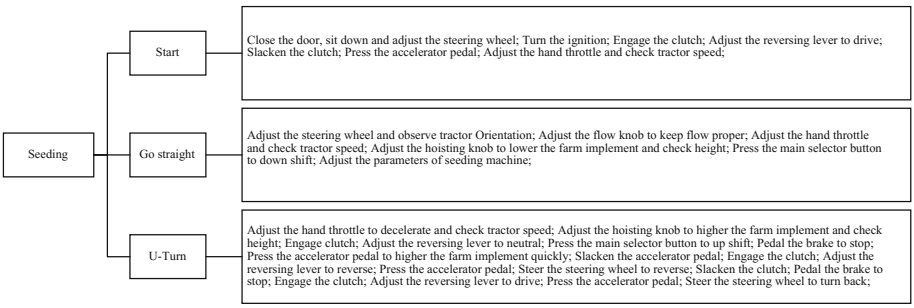


Fig. 2. Task analysis of tractor driver during seeding operation

4.2 Analysis of Time Occupied

The first dimension, time occupied, is described by behavior time of operator. Behavior time is divided into reaction time, movement time and operation time. Reaction time is categorized into simple reaction time and choice reaction time. When the operator is stimulated by only one kind of stimulus, the time spent is simple reaction time. When the operator receives more than one kind of stimulus, the time spent is choice reaction time. Simple reaction time via different sensory modality are obtained from empirical data [13]. Tractor driver mainly obtains information from the visual channel, so the type of reaction time is simple reaction time.

Movement time refers to the duration from action start to operation start, which is evaluated by Fitts' law [14]. Movement time is given by:

$$T_{\mathrm{m}} = I_{\mathrm{m}} \log_2(D/S + 1), \tag{1}$$

where

I_m is the movement control iteration time ranging from 70 to 120 ms, D the hand moving distance, S the diameter of switch. The value of the movement control iteration time is set at 100 ms.

Operation time denotes the duration during which the tractor driver operates buttons or levers, which is read from the video.

4.3 Analysis of Level of Action Load and Action-Set Switching

The load of each action is obtained from the behavior parameters rating form in Table 2. The third factor ASS is obtained from the video. When the type of the tractor

Table 2. ANOVA of pupil diameter

		Sum of squares	df	Mean square	F	Sig.
Horizontal pupil diameter	Between groups	1901.969	16	118.873	4.425	0.000
	Within groups	886.413	33	26.861		
	Total	2788.382	49			
Vertical pupil diameter	Between groups	1868.599	16	116.787	3.739	0.001
	Within groups	1030.780	33	31.236		
	Total	2899.379	49			

driver's action changes once, the number of ASS is increased by 1. To quantify the influences on cognitive load from time occupied, level of action load and action-set switching, the cognitive load yields:

$$\mathrm{CL} = \sqrt{(\mathrm{TO})^2 + (\mathrm{LAL})^2 + (\mathrm{ASS})^2}. \tag{2}$$

5 Experiment

5.1 Experiment Equipment

Analysis on eye movement is adopted to verify the evaluation method described above. Pulping operation and seeding operation are simulated in the lab. Subject wears an SMI HED eye-tracker during simulation. The iView X software is run, and one-eye image is moved and the focus is adjusted to clearly show the pupil at the center of video. Adjust scenario camera to the right ahead of the eye. A five point calibration is conducted before simulation environment to setup the measurement parameters. The gaze position, pupil diameter, pupil position and blink rate are recorded during operation, and data is analyzed by BeGaze3.5.

5.2 Pulping Operation

Change of pupil diameter can indicate variation of cognitive load, and blink rate is generally considered to be related to visual load. So this research adopts blink rate and pupil diameter to indicate tractor driver's cognitive load.

Pupil diameter varies individually. In order to improve accuracy, self-contrasted method is adopted. The pupil diameter is firstly recorded when subject is under resting condition. Then, the pupil diameter is recorded again during operation. Note that the two sets of data are obtained under the same condition of illumination. Horizontal pupil diameter and vertical pupil diameter at work are both shown in Fig. 3.

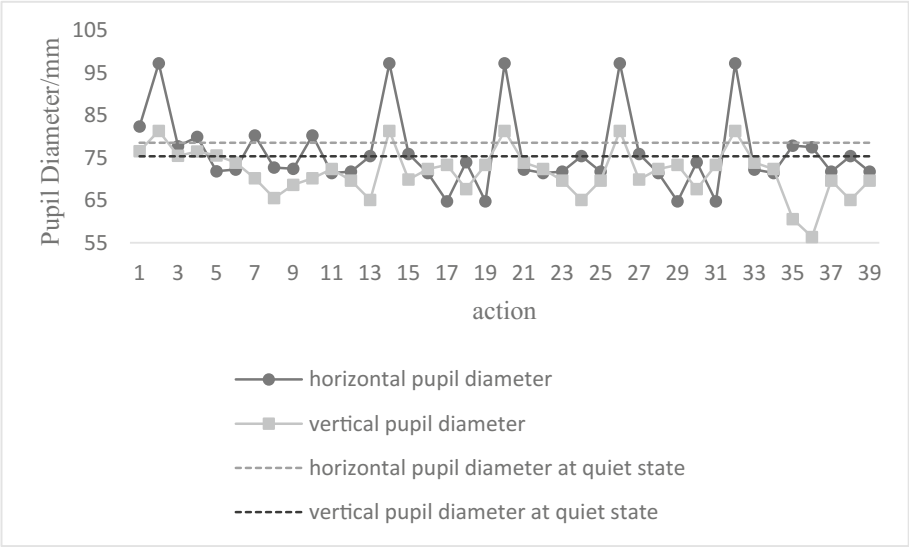


Fig. 3. Pupil diameter during pulping operation

SPSS software is used for ANOVA on the pupil diameter, the value of α is set at 0.1. Table 2 shows the analysis result. The Sig. value is less than 0.1, meaning that there is a significant difference in pupil diameter at different action, i.e. at different cognitive load, based on ANOVA result.

The number of blinks, i.e. blink rate, is recorded by the iVIEW X software. The blink rate is recorded respectively under resting condition and during operation. Figure 4 shows the blink rate during operation, in contrast to the blink rate during resting condition.

Similar ANOVA are also conducted on blink rate with α set at 0.1. Table 3 shows the analysis result, which means that there is a significant difference in blink rate at different cognitive load.

The load value obtained from the cognitive load model is plotted with experimental results of pupil diameter, as shown in Fig. 5.

Correlation analysis is conducted by SPSS, and the result is shown in Table 4.

As can be seen from Table 4 (a), significance equals 0.727, which is greater than 0.05. Therefore, the correlation coefficient between horizontal pupil diameter and the load is not statistically significant. On the contrary, significance in Table 4 (b) is smaller than 0.05, so the vertical pupil diameter is negatively correlated with cognitive load. Correlation coefficient is −0.524, Fig. 6 shows the scatter diagram between vertical pupil diameter and cognitive load.

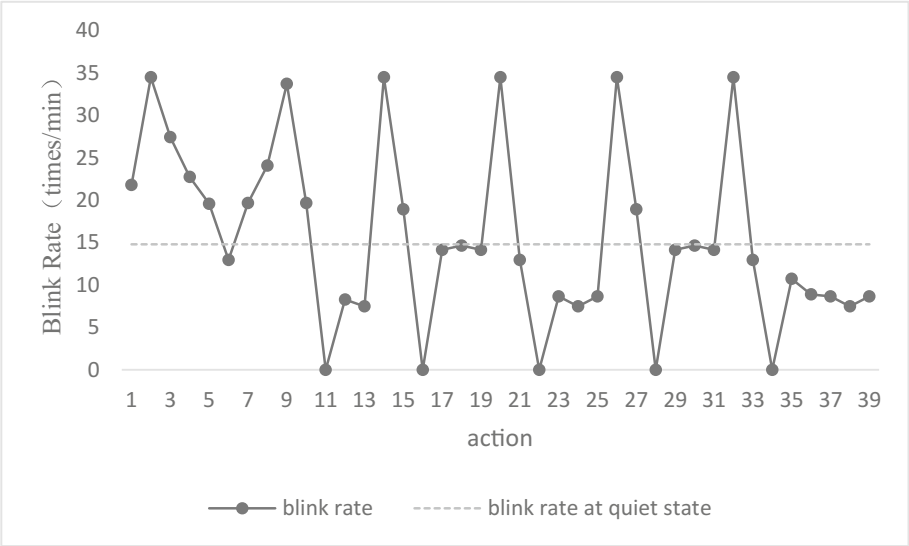


Fig. 4. Blink rate level during pulping operation

Table 3. ANOVA of blink rate

	Sum of squares	df	Mean square	F	Sig.
Between groups	3591.968	16	224.498	1.800	.075
Within groups	4114.823	33	124.692		
Total	7706.791	49			

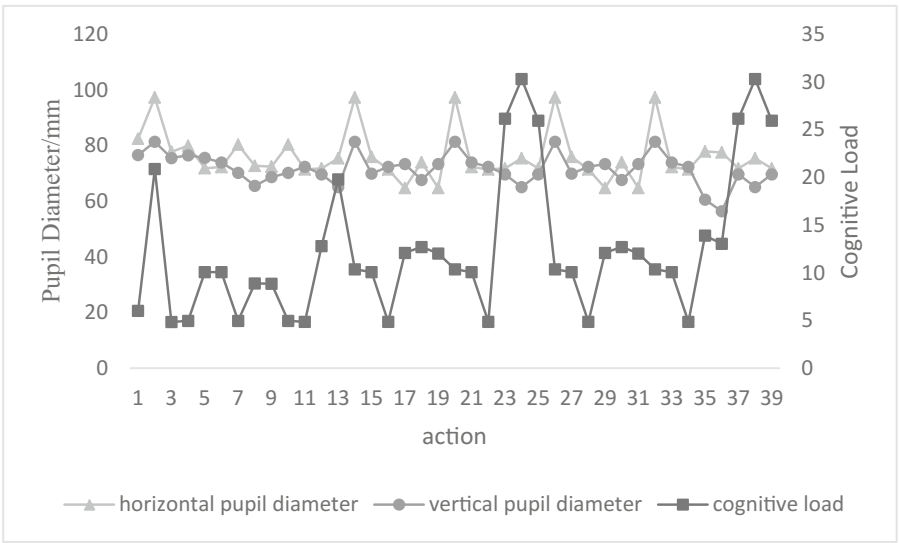


Fig. 5. Pupil diameter and cognitive load during pulping operation

Table 4. (a) Correlation analysis of horizontal pupil diameter and cognitive load (b) Correlation analysis of vetical pupil diameter and cognitive load

		Cognitive load	Horizontal pupil diameter
cognitive load	Pearson correlation	1	−0.058
	Sig.(2 tailed)		0.727
	N	39	39
Horizontal pupil diameter	Pearson correlation	−0.058	1
	Sig.(2 tailed)	0.727	
	N	39	39
		Cognitive load	Horizontal Pupil Diameter
cognitive load	Pearson correlation	1	−0.524**
	Sig.(2 tailed)		0.002
	N	33	33
Vertical pupil diameter	Pearson correlation	−0.524**	1
	Sig.(2 tailed)	0.002	
	N	33	33

**Correlation is significant at the 0.01 level (2 tailed)

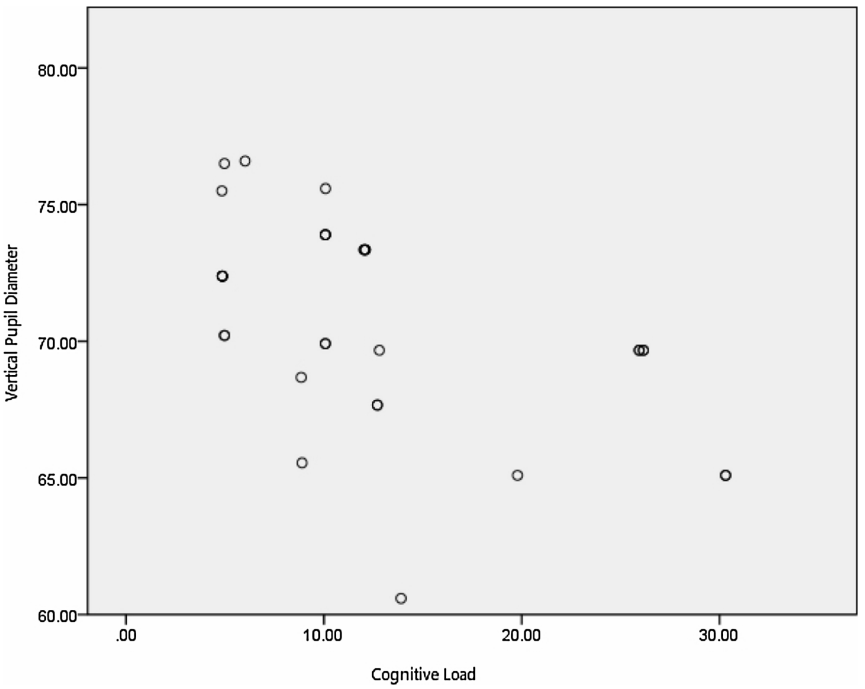


Fig. 6. Correlation analysis scatter diagram of cognitive load and vetical pupil diameter

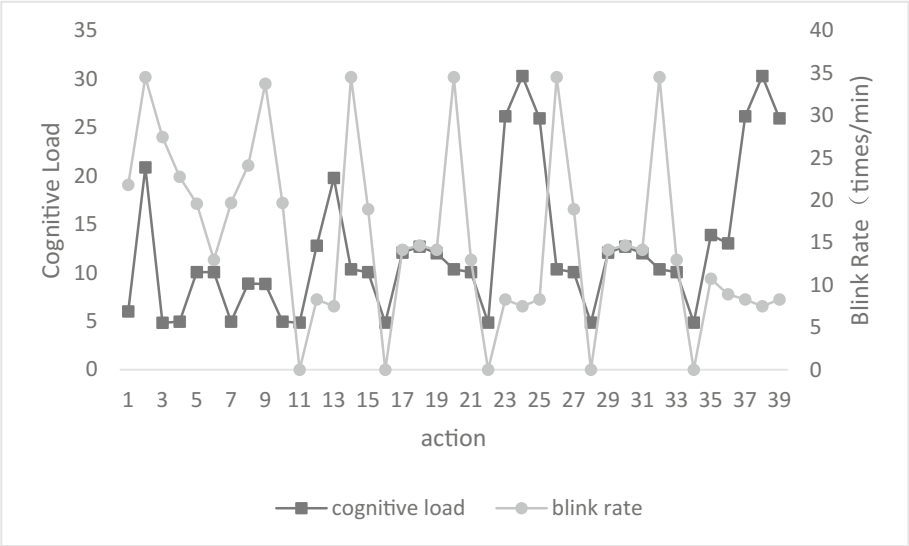


Fig. 7. Cognitive load and blink rate during pulping operation

The comparison between the load value and the blink rate is shown in Fig. 7. Correlation analysis is also conducted, and the result is shown in Table 5.

Table 5. Correlation analysis of blink rate and cognitive load

		Blink rate	Cognitive load
Blink rate	Pearson correlation	1	−0.163
	Sig.(2 tailed)		0.321
	N	39	39
Cognitive load	Pearson correlation	−0.163	1
	Sig.(2 tailed)	0.321	
	N	39	39

As can be seen from Table 5, significance equals 0.321, which is greater than 0.05. So the correlation coefficient between blink rate and the load is not statistically significant.

From the above experimental result, we can draw a conclusion that the tractor driver's pupil diameter decreases with the increase of operation complexity. Pupil diameter can reflect brain's activity. In the experiment, subject's pupil diameter changes constantly, which can reflect the transient state of the cognitive load.

5.3 Pulping Operation vs. Seeding Operation

The actions during similar functions are compared between pulping operation and seeding operation. The results are shown in Figs. 8 and 9.

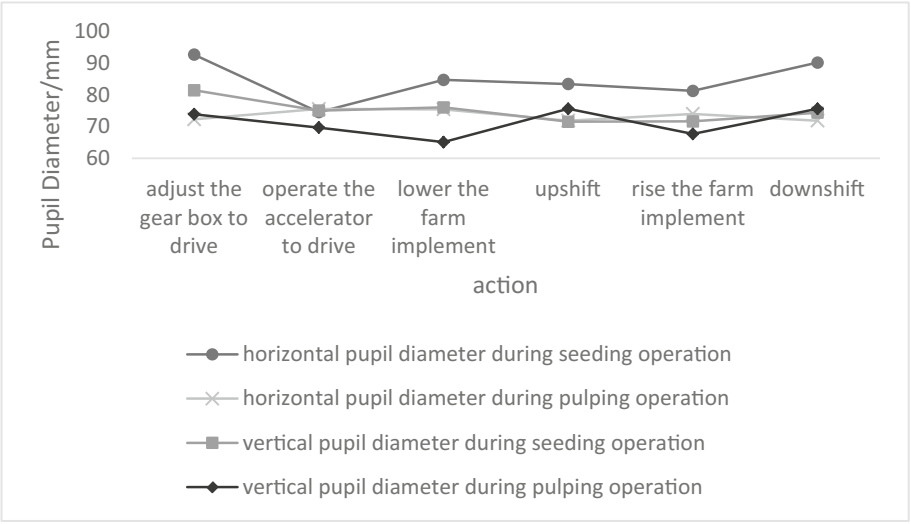


Fig. 8. Pupil diameter during seeding operation and pulping operation

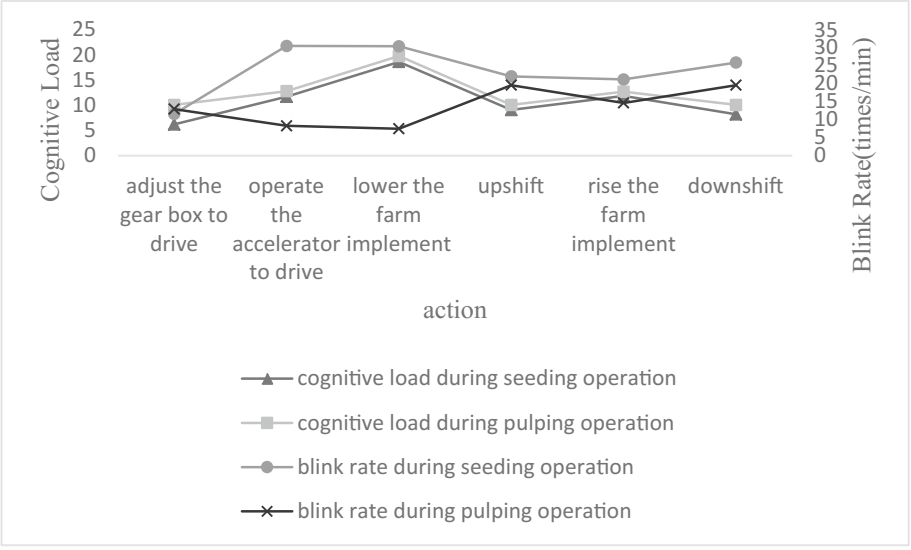


Fig. 9. Comparison of blink rate and cognitive load between seeding operation and pulping operation

Cognitive load in pulping operation is higher than that in seeding operation, which may indicate that the cognitive load in electronic control platform is lower than that in mechanical platform. The curve of average cognitive load during an entire operation over blink rate/pupil diameter may be U type instead of linear type, but more operation is needed to prove this hypothesis.

6 Conclusion

This research improves previous cognitive load model with the reference of multiple resources theory, and puts forward a method of quantitative evaluation of cognitive load during field work. Experiment result with an SMI HED eye-tracker proves the validity of the model and shows that cognitive load negatively correlates with vertical pupil diameter. This method is nonintrusive to normal operation and can reflect the tractor driver's cognitive load objectively and accurately. More operation such as ploughing will be analyzed in further research to investigate the average cognitive load during an entire operation over blink rate/pupil diameter.

Acknowledgments. This research is funded by project named Experimental Design and Data Processing of Occupant Workload Verification (F20150574). The authors also thank Luoyang Tractor Institute Co., Ltd for the help in experiments.

References

1. Wang, J., Fang, W., Guangyan, L.I.: Mental workload evaluation method based on multi-resource theory model. J. Beijing Jiaotong Univ. **34**(6), 107–110 (2010). (in Chinese)
2. Wickens, C.D., Gordon, S.E., Liu, Y., Lee, J.: Introduction to Human Factors Engineering, 2nd edn., pp. 327–330. East China Normal University Press, Shanghai (2007)
3. Wang, J., Masakazu, K., Moriizumi. S.: Influences of the driving intensity of tractor drivers on the drivers' physiological load IEA (2009). (in Chinese)
4. Wang, Y.: Physiological characteristics of tractor driver and traffic safety. China Agric. Mach. Saf. Superv. **04**, 43 (2003). (in Chinese)
5. Song, C.: Studies of Driver's Physiological Loads Based on HRV. Inner Mongolia Agricultural University (2007). (in Chinese)
6. Jin-Bo, L.I., Bai-Hua, X.U., Tian, X.H.: Construction of prediction models of cognitive load in human-machine interaction process. Acta Psychol. Sin. **42**(42), 559–568 (2010). (in Chinese)
7. Neerincx, M.A.: Cognitive task load design: model, methods and examples. In: Hollnagel, E. (ed.) Handbook of Cognitive Task Design, pp. 283–305. Lawrence Erlbaum Associates, Mahwah (2003)
8. Cullen, L.: Validation of A Methodology for Predicting Performance and Workload. Eurocontrol Experimental Centre (1999)
9. Mccracken, J.H., Aldrich, T.B.: Analyses of Selected LHX Mission Functions: Implications for Operator Workload and System Automation Goals (1984)

10. Szabo, S.M., Bierbaum, C.R.: A Comprehensive Task Analysis of the AH-64 Mission with Crew Workload Estimates and Preliminary Decision Rules for Developing an AH-64 Workload Prediction Model, vol. 2. Appendixes A-E (1986)
11. Bierbaum, C.R., Szabo, S.M., Aldrich, T.B.: Task Analysis of the UH-60 Mission and Decision Rules for Developing a UH-60 Workload Prediction Model, vol. 2. Mission Analysis Appendixes A-E (1987)
12. Keller, J.: Human performance modeling for discrete-event simulation: workload Winter Simulation Conference: Exploring New Frontiers, San Diego, California, USA, December. IEEE Xplore, pp. 157–162 (2002)
13. Sun, L., Cui, K., Sun, L.: Human Factors, pp. 99–100. Science Press, Beijing (2011)
14. Ding, Y.: Usability design of e-commerce website from fitts' law. Mark. Modernization **34**, 101–102 (2006). (in Chinese)

Road and Rail: Automation

A Review of Non-driving-related Tasks Used in Studies on Automated Driving

Frederik Naujoks[1(✉)], Dennis Befelein[1], Katharina Wiedemann[2], and Alexandra Neukum[1]

[1] Center for Traffic Sciences, University of Würzburg, Röntgenring 10, 97070 Würzburg, Germany
{naujoks, befelein, neukum}@psychologie.uni-wuerzburg.de
[2] Würzburg Institute for Traffic Sciences, Robert-Bosch-Straße 4, 97209 Veitshöchheim, Germany
wiedemann@wivw.de

Abstract. Conditionally automated driving (CAD) relieves the driver from monitoring current traffic conditions. This type of driving inherently enables the driver to execute different non-driving-related tasks (NDRTs). However, the driver still must be available as a backup option. With this in mind, the classification and evaluation of various NDRTs concerning their impact on driver performance in takeover scenarios represent an important contribution toward the creation of safe CAD functions. We reviewed various NDRTs that were used in studies on automated driving. The focus was on assigning aspects of these activities (e.g., ability to visually monitor traffic, necessity of sustained attention to NDRT, etc.) to various steps of the takeover process (e.g., noticing and interpreting takeover requests), which could be impaired by the execution of the respective NDRT. This, in turn, would increase the demands on the driver with respect to managing the takeover situation.

Keywords: Automated driving · Takeover situations · Non-driving-related tasks

1 Introduction

Conditionally automated driving (CAD, SAE Level 3) relieves the driver from manually controlling the vehicle and continuously monitoring the driving situation. Driving time can thus be used for non-driving-related tasks (NDRTs); however, the driver still should be ready to take over manual vehicle control occasionally (e.g., in case of system limits). Therefore, driving performance during takeover situations has received considerable attention and a large variety of NDRTs has been used in empirical studies. While there have been extensive efforts to categorize different secondary tasks according to the workload imposed on the driver during manual driving, there is no such taxonomy yet in the context of automated driving. Based on a review of NDRTs used in automated driving research, we propose a theoretical model that explains the mechanisms involved in disengaging from the NDRT and taking back manual vehicle control.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_52

2 Overview of Non-driving-related Tasks

Existing classifications of NDRTs generally follow two distinct approaches: **Phenomenological approaches** predominantly target the categorisation of various NDRTs as well as their prevalence. [1–4] and [5] presented lists of tasks, which divide NDRTs into a total of nine major categories, such as physical activities or operation of accessories installed inside the vehicle. The focus of this approach is on activities that can be assumed to be prevalent in a practically relevant manner, based on observational or survey studies.

Demand-based approaches, on the other hand, aim to determine the level of demands on the driver associated with various NDRTs. For example, in [6], accidents and near-accidents during the so-called 100-Car Naturalistic Driving Study were reviewed and various NDRTs were grouped by level of demand (e.g., "easy": a single glance/operational step necessary for task completion; "complex": more than three glances/operational steps). Together with other NDS studies [7], these findings resulted in the drafting of NHTSA guidelines on the use of electronic devices with visual-manual interfaces [8], which consider visual distraction as a main criterion for the compatibility with manual vehicle operation. For partially automated driving, [9] presented a list of activities, which attributed limitations of the ability to react to system limits and malfunctions to the following demand dimensions: (1) *primary modality* of the task (visual vs. auditory), (2) *interaction* (active: controlled by the driver vs. passive: controlled by the task), (3) *interruptibility* (easy vs. difficult) and (4) *coding of information* (verbal vs. spatial). In this model, the assumed degree of the driver's limitation of the ability to take over manual vehicle control results from the proximity to the driving task, which is seen as a visual, active and not easily interruptible task with spatial coding of information. In line with the multiple resource model [10], it is assumed that the manual takeover of vehicle operations is particularly complicated if the same resources are required by the NDRTs and the driving task.

Existing categorisations of NDRTs have so far been based on a so-called dual-task paradigm. During the processing of each respective NDRT, it is always the driver's responsibility to monitor the current traffic conditions and to intervene with corrective measures in cases of system malfunctions or -limits. In contrast to this scenario, CAD permits ignoring driving-related tasks permanently and requires the driver to switch tasks from the NDRT to manual vehicle operation only when system limits are reached (so-called task-switching paradigm). This paradigm change necessitates significant expansions of the classification of NDRTs. The aim of this paper is thus to expand the procedure in [9] to CAD, starting with a review of NDRTs examined in previous studies on automated driving (SAE L1-3).

Tables 1 and 2 provide an overview on NDRTs used in empirical studies, categorised in standardised and everyday NDRTs. In line with the multiple resource model [10], standardised NDRTs can be expected to interfere mainly with noticing and interpreting takeover requests, for example, by presenting cognitive (e.g., n-back task) or visual demands (e.g., SuRT, RSVP task). However, CAD is expected to facilitate a variety of more-complex NDRTs (cf. Table 1), the execution of which may be incompatible with driving under lower levels of automation, such as the utilisation of

the available driving time to write emails or consume entertainment programs. For this reason, it is desirable to extend existing classifications by including such natural NDRTs. On the other hand, it is entirely conceivable that, by releasing the driver from driving-related tasks, CAD could facilitate a state of passive fatigue [11], which primarily results from the absence of stimulating NDRTs and/or the execution of non-challenging NDRTs. Therefore, the classification provided in the next section also includes aspects previously given less attention such as the interruption or termination of NDRTs, as well as under- or overload caused by the execution of NDRTs during the automated drive.

Table 1. Overview of everyday NDRTs used in studies on automated driving.

NDRT	Description	References
In-vehicle information systems (IVIS)	Interaction with in-vehicle information systems	[12, 13]
Free choice of tasks	Free choice of NDRT by participant (e.g., putting on make-up, watching a DVD)	[14–16]
Reading	Text reading or looking at graphical material	[14, 15, 17–22]
Weather app	Retrieving weather-related information from a smartphone	[23]
Entering destination in navigation system	Entering a destination in the in-car navigation system	[24]
Quiz-Task	Answering questions with multiple answers (similar to "who wants to be a millionaire")	[25, 26]
Reading out text aloud	Reading a text that is displayed visually	[27, 28]
Answering questions after hearing texts	Listening to information (e.g., nearby restaurants), answering questions verbally	[29]
Visual information search	E.g., retrieving stock prices from a display and reading them out loud	[27]
Eating	Eating (e.g., sweets that are placed on the passenger seat)	[14, 15, 17]
Gaming	Playing computer games	[21, 30, 31]
Listening to music	Listening to music played by the radio	[32]
Speaking on the phone	Talking to the experimenter (e.g., solving arithmetic tasks)	[26, 33]
Writing/dialling on phone	Writing of text messages or dialling with cell phone	[15]
Text entry	Entering a text with or without instruction to write specific texts	[15, 34]
Watching a movie	Watching a movie with or without instruction to answer questions	[14, 32, 35–37]
Leaning out of the window	Leaning out of the window during automated drive	[38]
Sitting cross-legged	Sitting cross-legged during automated drive	[38]

Table 2. Overview of standardised NDRTs used in studies on automated driving.

NDRT	Description	References
Tracking task	Keeping tracking object within tolerance limits (e.g., with computer mouse)	[39, 40]
Surrogate Reference Task (SuRT)	Visual search of target stimulus among cluster of stimuli, selecting region where target is located	[41–51]
n-Back	Serial presentation of targets and distractors. Identifying when current stimulus matches the target n steps earlier in the sequence.	[45, 51]
20-Questions-Task	Asking 20 yes/no-questions to guess an item	[52, 53]
Addition	Adding numbers shown on a display	[54]
Rapid serial visual presentation task	Counting targets on a serial presentation of targets and distractors	[55]
Peripheral detection task	Identifying targets that appear in random intervals at different positions	[56]
Rotated Figures Task	Deciding whether two differently rotated figures match or not	[57]
Position of stimulus	Indicating the position of a target among distractors	[58]
Cloze-test	Filling in missing words in a cloze test	[59]
Oddball task	Reacting to rare stimuli within a series of auditory stimuli	[60]

3 Takeover Process Model

Taking back manual control over the vehicle consists of a task switch from the NDRT (or from inactivity) to the execution of the driving task. We propose that this affords time consuming re-configuration of the drivers' *sensory* (e.g., taking one's eyes off the NDRT and attending to relevant HMI displays), *motoric* (e.g., freeing one's hands and taking them back on the steering wheel) and *cognitive* state (e.g., re-configuration of mental task sets or response rules) to meet the demands of manually controlling the vehicle. Based on the review presented above, we assume that this transition process can essentially be divided into four steps, each of which corresponds to the following specific requirements placed upon the driver to successfully take over manual vehicle control (see Figs. 1 and 2):

(1) **Maintaining fitness to drive during automated driving** (step: "*CAD engaged*"): During the automated drive, both the lack of demands (*passive* fatigue) and excessive demands (*active* fatigue [11]) may have negative effects on the availability of the driver to take over manual vehicle control because both cases involve a shift in the driver's arousal level. Therefore, empirical assessments should identify such activities, which could lead to unsuitable demands on the driver in the medium or long term [60].

(2) **Noticing the necessity to take over manual vehicle control as soon as the takeover request is issued** (step: "*CAD degraded*"): The need to take over

Fig. 1. Takeover process model during conditionally automated driving.

manual vehicle control is usually indicated by a takeover request. This step of the takeover process is strongly interfered by NDRTs that decrease the drivers' capacity to recognise, process and interpret the takeover request [10]. For example, when the NDRT and the takeover request are presented in the same modality [20] or if the NDRT requires the driver to take his/her eyes off the display elements of the CAD function [35–37], this step of the transition process may be prolonged.

(3) **Termination or interruption of the executed NDRT and interfacing with the control elements** (step: "*CAD disengagement*"): As soon as the driver recognises the need to take over manual control, the termination of some NDRTs probably represents an additional demand on the driver. For example, necessary operational steps like stopping or pausing electronic devices or the need to put objects aside [17] may prolong the time needed to interface with the control elements.

(4) **Focusing on subsequent manual driving** (step: "*Manual Drive*"): Finally, it is possible that the NDRT was not entirely completed during the takeover phase, and that its continuation during the subsequent manual operation of the vehicle may impair quality aspects of the takeover process.

Based on these demands, various dimensions of NDRTs were defined (e.g., location, demand of visual attention, degree of effort involved in the interruption, etc.), which may define the degree of limitation of the driver's availability to take over manual driving. Figure 2 provides an overview of these dimensions and their assignment to the assumed steps of the takeover process. Table 3 describes the proposed dimensions in more detail.

To determine the specification of a NDRT within the respective dimension, the table presents several options of operationalising the individual dimensions. To achieve this, directly measurable characteristics of the activity (e.g., duration of operational steps, glance behaviour, etc.) and the measurement of latent constructs (e.g., estimates of cognitive demands through standardised questionnaires) are listed to the extent possible.

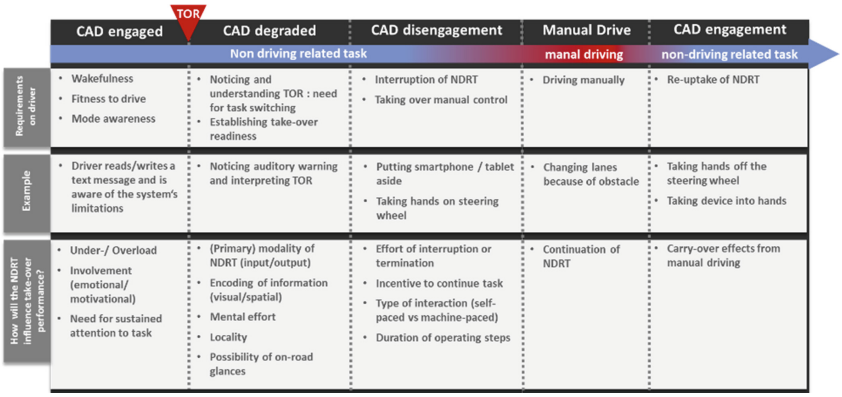


Fig. 2. Takeover process model during CAD and relation to NDRTs.

Table 3. Overview of NDRT dimensions and possible operationalisations.

Dimension	Operationalisations (examples)
Requirement: Maintaining fitness to drive during CAD (step: "CAD engaged")	
Under/Overload resulting in passive [11] or active fatigue [11]	Glance behaviour/physiology [61–63], observation [62, 64], self-reports [65]
Emotional *Involvement* leading to attentional tunnelling [66]	Glance behaviour/physiology [66, 67], self-reports [68, 69]
Need for sustained attention leading to vigilance decrement [70]	Vigilance tests [71, 72], task performance [70], physiology [73]
Requirement: Noticing the necessity to take over (step: "CAD degraded")	
Primary modality: use of visual or auditory channel for information retrieval from NDRT interferes with processing TOR [9, 10] *Encoding of information*: need of encoding spatial information related to the NDRT interferes with processing TOR [10] *Modality of input*: modality of input into NDRT [10] interferes with processing TOR	Performance decrement in dual task setting / switching costs in sequential task setting [74–76], self-reports [77–79]
Mental effort [80–82] imposed by NDRT limits cognitive capacity to process TOR	Glance behaviour [83, 84]; physiology [84–86], self-reports [87, 88]
Locality: Location of NDRT [89–91] requires turning away from the central road scene	Glance behaviour, body/head orientation [92, 93]
Low *possibility of on-road glances* while performing NDRT [94] limits monitoring of the traffic state	Frequency and length of glances on NDRT [95]

(*continued*)

Table 3. (*continued*)

Dimension	Operationalisations (examples)
Requirement: Termination/interruption of NDRT (step: "CAD disengagement")	
Effort of interruption or termination: high duration of actions needed to interrupt or end NDRT prolongs take-over action	Measurement of duration needed to interrupt or end the NDRT [96–98]
Incentive to continue task: extrinsic or intrinsic motivation to continue the NDRT [99] prolongs take-over action	Readiness to process NDRT with/without external incentives [99, 100]; self-report measures [99, 101, 102]
Type of interaction: machine-paced [9] NDRT prolongs take-over actions	Determined by nature of NDRT
Duration of operating steps: long duration of operating steps prolongs take-over action	Measurement of duration of operation steps of NDRT [96–98]

4 Outlook

This article illustrated an approach to prospectively evaluate the expected limitations caused by NDRTs on the ability to take over manual driving during CAD. We assume that the demands placed upon the driver by NDRTs correspond to specific limitations in his/her ability to take over manual driving that in turn can be assigned to specific NDRT dimensions. However, comprehensive, empiric assessments, which examine the relative contributions of the proposed demand dimensions, have not been conducted yet. Significantly more research is needed in this area so that the different aspects of NDRTs can be translated into a modelling framework to predict takeover time or quality when performing different NDRTs (for examples of modelling the demands imposed on the driver by secondary tasks during manual driving cf. [96–98]). Furthermore, evaluating whether the demands placed upon the driver exceed the driver's capacities also involves considering specifics of the take-over situation [103], for example by relating the available time budget to the expected takeover time when performing NDRTs [104].

Acknowledgements. This work results from the joint project Ko-HAF - Cooperative Highly Automated Driving and has been funded by the Federal Ministry for Economic Affairs and Energy based on a resolution of the German Bundestag.

References

1. Stutts, J.C., Reinfurt, D.W., Staplin, L., Rodgman, E.A.: The Role of Driver Distraction in Traffic Crashes. AAA Foundation for Traffic Safety, Washington (2001)
2. Gordon, C.: Driver distraction: an initial examination of the 'attention diverted by' contributory factor codes from crash reports and focus group research on perceived risks. 2005 IPENZ Conference, Auckland (2005)

3. Glaze, A.L., Ellis, J.M.: Pilot Study of Distracted Drivers. Transportation and Safety Training Center, Virginia Commonwealth University, Richmond (2003)
4. Huemer, A.K., Vollrath, M.: Ablenkung durch fahrfremde Tätigkeiten - Machbarkeitsstudie. Bundesanstalt für Strassenwesen, Bergisch Gladbach (2012)
5. Schoemig, N., Schoch, S., Neukum, A., Schumacher, M., Wandtner, B.: Simulatorstudien zur Ablenkungswirkung fahrfremder Tätigkeiten. Bundesanstalt für Strassenwesen, Bergisch Gladbach (2015)
6. Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J.: The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. National Highway Traffic Safety Administration, Washington (2006)
7. Olson, R.L., Hanowski, R.J., Hickman, J.S., Bocanegra, J.L.: Driver Distraction in Commercial Vehicle Operations. Virginia Tech Transportation Institute, Blacksburg (2009)
8. National Highway Traffic Safety Administration: Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. National Highway Traffic Safety Administration, Washington, DC (2012)
9. Spiessl, W., Hussmann, H.: Assessing error recognition in automated driving. IET Intell. Transp. Syst. **5**, 103–111 (2011)
10. Wickens, C.D.: The structure of attentional resources. In: Nickerson, R.S. (ed.) Attention and Performance VIII, pp. 239–257. Taylor & Francis, New York (1980)
11. May, J.F., Baldwin, C.L.: Driver fatigue: the importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Trans. Res. F.-Traffic **12**, 218–224 (2009)
12. Toffetti, A., Wilschut, E., Martens, M., Schieben, A., Rambaldini, A., Merat, N., Flemish, F.: CityMobil: human factor issues regarding highly automated vehicles on eLane. Transp. Res. Rec. **2110**, 1–8 (2009)
13. Flemisch, F., Griesche, S., Heesen, M., Kaussner, A., Niemann, J., Petermann, I., Schieben, A., Schoemig, N.: Highly automated vehicles for intelligent transport: Validation of preliminary design of HAVEit systems by simulator tests. Technical report, HAVEit Consortium (2010)
14. Carsten, O., Lai, F., Barnard, A., Merat, N., Jamson, A.: Control task substitution in semi-automated driving: does it matter what aspects are automated? Hum. Factors **54**, 747–761 (2012)
15. Llaneras, R.E., Salinger, J., Green, C.A.: Human factors issues associated with limited ability autonomous driving systems: drivers' allocation of visual attention to the forward roadway. In: 7th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 92–98. University of Iowa, Iowa City (2013)
16. Jamson, A.H., Merat, N., Carsten, O.M., Lai, F.C.: Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. Trans. Res. C.-Emerg. **30**, 116–125 (2013)
17. Helldin, T., Falkman, G., Riveiro, M., Davidsson, S.: Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. In: 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 210–217. ACM, New York (2013)
18. Dogan, E., Deborne, R., Delhomme, P., Kemeny, A., Jonville, P.: Evaluating the shift of control between driver and vehicle at high automation at low speed: the role of anticipation. In: 5th Transport Research Arena (TRA) Conference, Paris (2014)
19. Melcher, V., Rauh, S., Diederichs, F., Widlroither, H., Bauer, W.: Take-over requests for automated driving. Proced. Manuf. **3**, 2867–2873 (2015)

20. Naujoks, F., Mai, C., Neukum, A.: The effect of urgency of take-over requests during highly automated driving under distraction conditions. In: Ahram, T., Karowski, W., Marek, T. (eds.) In: 5th International Conference on Applied Human Factors and Ergonomics, vol. 7, pp. 2099–2106. AHFE Conference, Krakau (2014)
21. Schwalk, M., Kalogerakis, N., Maier, T.: Driver support by a vibrotactile seat matrix – Recognition, adequacy and workload of tactile patterns in take-over scenarios during automated driving. Proced. Manuf. **3**, 2466–2473 (2015)
22. Diederichs, F., Bischoff, S., Widlroither, H., Reilhac, P., Hottelart, K., Moizard, J.: Smartphone integration and SAE level 3 car automation: a new cockpit concept and its evaluation in a car simulator. In: VDI (ed.) Der Fahrer im 21. Jahrhundert, pp. 63–78. VDI-Verlag, Düsseldorf (2015)
23. Naujoks, F., Purucker, C., Neukum, A.: Secondary task engagement and vehicle automation–Comparing the effects of different automation levels in an on-road experiment. Trans. Res. F.-Traffic **38**, 67–82 (2016)
24. Tijerina, L., Johnston, S., Parmer, E., Pham, H., Winterbottom, M., Barickman, F.: Preliminary Studies in Haptic Displays for Rear-End Collision Avoidance System and Adaptive Cruise Control Applications. National Highway Traffic Safety Administration, Washington (2000)
25. Petermann-Stock, I., Hackenberg, L., Muhr, T., Mergl, C.: Wie lange braucht der Fahrer - eine Analyse zu Übernahmezeiten aus verschiedenen Nebentätigkeiten während einer hochautomatisierten Staufahrt. 6. Tagung Fahrerassistenzsysteme - der Weg zum automatischen Fahren. TU München, München (2013)
26. Neubauer, C., Matthews, G., Saxby, D.: Fatigue in the automated vehicle - do games and conversation distract or energize the driver? Proc. Hum. Fact. Ergon. Soc. Annu. Meet. **58**, 2053–2057 (2014)
27. Rudin-Brown, C.M., Parker, H.A.: Behavioural adaptation to adaptive cruise control (ACC): implications for preventive strategies. Trans. Res. F.-Traffic **7**, 59–76 (2004)
28. Louw, T., Merat, N., Jamson, H.: Engaging with highly automated driving: To be or not to be in the loop? In: 8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Salt Lake City (2015)
29. Seppelt, B.D., Lees, M.N., Lee, J.D.: Driver distraction and reliance: adaptive cruise control in the context of sensor reliability and algorithm limits. In: Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Rockport (2005)
30. van den Beukel, A.P., van der Voort, M.C.: The influence of time-criticality on situation Awareness when retrieving human control after automated driving. In: 16th International IEEE Conference on Intelligent Transportation Systems (ITSC), The Hague (2013)
31. Politis, I., Brewster, S., Pollick, F.: Language-based multimodal displays for the handover of control in autonomous cars. In: Burnett, G., Gabbard, J., Green, P., Osswald, S., Eren, A., Antrobus, V. (eds.) Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 3–10. ACM, New York (2015)
32. Blommer, M., Curry, R., Kochhar, D., Swaminathan, R., Talamonti, W., Tijerina, L.: The effects of a scheduled driver engagement strategy in automated driving. Proc. Hum. Fact. Ergon. Soc. Annu. Meet. **59**, 1681–1685 (2015)
33. Ma, R., Kaber, D.B.: Situation awareness and workload in driving while using adaptive cruise control and a cell phone. Int. J. Ind. Ergon. **35**, 939–953 (2005)
34. Zeeb, K., Buchner, A., Schrauf, M.: What determines the take-over time? An integrated model approach of driver take-over after automated driving. Accid. Anal. Prev. **78**, 212–221 (2015)

35. Jamson, H., Merat, N., Carsten, O., Lai, F.: Fully-automated driving: the road to future vehicles. In: Sixth International Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Olympic Valley (2011)
36. Mok, B., Johns, M., Lee, K. J., Miller, D., Sirkin, D., Ive, P., Ju, W.: Emergency, automation off: unstructured transition timing for distracted drivers of automated vehicles. In: 18th IEEE International Conference on Intelligent Transportation Systems (ITSC), Las Palmas (2015)
37. Walch, M., Lange, K., Baumann, M., Weber, M.: Autonomous driving: investigating the feasibility of car-driver handover assistance. In: Burnett, G., Gabbard, J., Green, P., Osswald, S., Eren, A., Antrobus, V. (eds.) Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 11–18. ACM, New York (2015)
38. Omae, M., Hashimoto, N., Sugamoto, T., Shimizu, H.: Measurement of driver's reaction time to failure of steering controller during automatic driving. SAE J.-Rev. Automot. Eng. **26**, 213–215 (2005)
39. Schaller, T., Schiehlen, J., Gradenegger, B.: Stauassistenz–Unterstützung des Fahrers in der Quer-und Längsführung: Systementwicklung und Kundenakzeptanz. 3th Conference on Driver Assistance, Munich (2008)
40. Damböck, D., Farid, M., Tönert, L., Bengler, K.: Übernahmezeiten beim hochautomatisierten Fahren. In: 5th Conference on Driver Assistance, Munich (2012)
41. Beggiato, M., Krems, J.F.: The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. Trans. Res. F.-Traffic **18**, 47–57 (2013)
42. Vollrath, M., Schleicher, S., Gelau, C.: The influence of cruise control and adaptive cruise control on driving behaviour–a driving simulator study. Accid. Anal. Prev. **43**, 1134–1139 (2011)
43. Stockert, S., Richardson, N.T., Lienkamp, M.: Driving in an increasingly automated world-approaches to improve the driver-automation interaction. Proced. Manuf. **3**, 2889–2896 (2015)
44. Gold, C., Lorenz, L., Damböck, D., Bengler, K.: Partially automated driving as a fallback level of high automation. In: 6th Conference on Driver Assistance, Munich (2013)
45. Lorenz, L., Hergeth, S.: Einfluss der Nebenaufgabe auf die Überwachungsleistung beim teilautomatisierten Fahren. In: VDI (ed.) Der Fahrer im 21. Jahrhundert, pp. 159–173. VDI-Verlag, Düsseldorf (2015)
46. Lorenz, L., Kerschbaum, P., Schumann, J.: Designing take over scenarios for automated driving: how does augmented reality support the driver to get back into the loop? Proc. Hum. Fact. Ergon. Soc. Annu. Meet. **58**, 1681–1685 (2014)
47. Kerschbaum, P., Lorenz, L., Bengler, K.: Highly automated driving with a decoupled steering wheel. Proc. Hum. Fact. Ergon. Soc. Annu. Meet. **58**, 1686–1690 (2014)
48. Beller, J., Heesen, M., Vollrath, M.: Improving the driver–automation interaction an approach using automation uncertainty. Hum. Factors **55**, 1130–1141 (2013)
49. Hergeth, S., Lorenz, L., Krems, J. F., Toenert, L.: Effects of take-over requests and cultural background on automation trust in highly automated driving. In: 8th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Salt Lake City (2015)
50. Körber, M., Weißgerber, T., Kalb, L., Blaschke, C., Farid, M.: Prediction of take-over time in highly automated driving by two psychometric tests. DYNA **82**, 195–201 (2015)
51. Radlmayr, J., Gold, C., Lorenz, L., Farid, M., Bengler, K.: How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. Proc. Hum. Fact. Ergon. Soc. Annu. Meet. **58**, 2063–2067 (2014)

52. Merat, N., Jamson, A.H., Lai, F.C., Carsten, O.: Highly automated driving, secondary task performance, and driver state. Hum. Factors **54**, 762–771 (2012)
53. Gold, C., Körber, M., Hohenberger, C., Lechner, D., Bengler, K.: Trust in automation – before and after the experience of take-over scenarios in a highly automated vehicle. Proced. Manuf. **3**, 3025–3032 (2015)
54. Lin, T.W., Hwang, S.L., Green, P.A.: Effects of time-gap settings of adaptive cruise control (ACC) on driving performance and subjective acceptance in a bus driving simulator. Saf. Sci. **47**, 620–625 (2009)
55. Lee, J., McGehee, D., Brown, T., Marshall, D.: Effects of adaptive cruise control and alert modality on driver performance. Transp. Res. Rec. **1980**, 49–56 (2006)
56. Petermeijer, S.M., Abbink, D.A., de Winter, J.C.: Should drivers be operating within an automation-free bandwidth? Evaluating haptic steering support systems with different levels of authority. Hum. Factors **57**, 5–20 (2015)
57. Stanton, N.A., Young, M., McCaulder, B.: Drive-by-wire: the case of driver workload and reclaiming control with adaptive cruise control. Saf. Sci. **27**, 149–159 (1997)
58. Ho, A.W., Cummings, M.L.: Multiple alarms and driving situational awareness (No. HAL2005-01). MIT, Cambridge (2005)
59. Beggiato, M., Hartwich, F., Schleinitz, K., Krems, J., Othersen, I., Petermann-Stock, I.: What would drivers like to know during automated driving? Information needs at different levels of automation. In: 7th conference driver assistance, Munich (2015)
60. Körber, M., Cingel, A., Zimmermann, M., Bengler, K.: Vigilance decrement and passive fatigue caused by monotony in automated driving. Proced. Manuf. **3**, 2403–2409 (2015)
61. Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., Othersen, I.: The interaction between highly automated driving and the development of drowsiness. Proced. Manuf. **3**, 6652–6659 (2015)
62. Barr, L.C., Yang, C.Y.D., Hanowski, R.J., Olson, R.: An Assessment of Driver Drowsiness, Distraction, and Performance in a Naturalistic Setting. Department of Transportation Research and Special Programs Administration, Washington, DC (2011)
63. Pal, N.R., Chuang, C.-Y., Ko, L.-W., Chao, C.-F., Jung, T.-P., Liang, S.-F., Lin, C.-T.: EEG-based subject-and session-independent drowsiness detection: an unsupervised approach. EURASIP J. Adv. Signal Process. **2008**, 1–11 (2008)
64. Kroes, S.: Detecting Boredom in Meetings. University of Twente, Enschede (2005)
65. Åkerstedt, T., Gillberg, M.: Subjective and objective sleepiness in the active individual. Int. J. Neurosci. **52**, 29–37 (1990)
66. Briggs, G.F., Hole, G.J., Land, M.F.: Emotionally involving telephone conversations lead to driver error and visual tunnelling. Trans. Res. F.-Traffic **14**, 313–323 (2011)
67. Ekman, P., Friesen, W.V.: Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto (1977)
68. Watson, D., Clark, L.A., Tellegen, A.: Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Personal. Soc. Psychol. **54**, 1063–1070 (1988)
69. Jackson, S.A., Marsh, H.W.: Development and validation of a scale to measure optimal experience: The Flow State Scale. J. Sport Exerc. Psychol. **18**, 17–35 (1996)
70. Mackworth, N.: The breakdown of vigilance during prolonged visual search. Q. J. Exp. Psychol. **1**, 6–21 (1948)
71. Robertson, I.H., Manly, T., Andrade, J., Baddeley, B.T., Yiend, J.: Oops!: performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia **35**, 747–758 (1997)
72. Lim, J., Dinges, D.F.: Sleep deprivation and vigilant attention. Ann. NY Acad. Sci. **1129**, 305–322 (2008)

73. Oken, B.S., Salinsky, M.C., Elsas, S.: Vigilance, alertness, or sustained attention: physiological basis and measurement. Clin. Neurophysiol. **117**, 1885–1901 (2006)
74. O'Donnell, R.D., Eggemeier, F.T.: Workload assessment methodology. In: Handbook of Perception and Human Performance, pp. 1–49. Wiley, New York (1986)
75. De Waard, D.: The Measurement of Drivers' Mental Workload. Traffic Research Center Groningen University, Groningen (1996)
76. Martens, M., Van Winsum, W.: Measuring Distraction: The Peripheral Detection Task. TNO, Soesterberg (2000)
77. Boles, D.B., Adair, L.P.: The multiple resources questionnaire (MRQ). Proc. Hum. Fact. Ergon. Soc. Annu. Meet. **45**, 1790–1794 (2001)
78. Pauzié, A.: A method to assess the driver mental workload: the driving activity load index (DALI). IET Intell. Transp. Sy. **2**, 315–322 (2008)
79. Pauzié, A.: Evaluating driver mental workload using the driving activity load index (DALI). In: C. Brusque (ed.) Proceedings of European Conference on Human Interface Design for Intelligent Transport Systems, pp. 67–79. HUMANIST publications, Lyon (2008)
80. Baddeley, A.D.: Working memory. Science **255**, 556–559 (1992)
81. Kahneman, D.: Attention and Effort. Prentice-Hall, Englewood Cliffs (1973)
82. Pashler, H.: Dual-task interference in simple tasks: data and theory. Psychol. Bull. **116**, 220–224 (1994)
83. May, J.G., Kennedy, R.S., Williams, M.C., Dunlap, W.P., Brannan, J.R.: Eye movement indices of mental workload. Acta Psychol. **75**, 75–89 (1990)
84. Wilson, G.F., Eggemeier, F.T.: Psychophysiological assessment of workload in multi-task environments. In: Damos, D.L. (ed.) Multiple-Task Performance, pp. 329–360. Taylor & Francis, London (1991)
85. Kramer, A.F.: Physiological metrics of mental workload: a review of recent progress. In: Damos, D.L. (ed.) Multiple-Task Performance, pp. 279–328. Taylor & Francis, London (1991)
86. Porges, S.W., Byrne, E.A.: Research methods for measurement of heart rate and respiration. Biol. Psychol. **34**, 93–130 (1992)
87. Zijlstra, F.R.H., Doorn, L.V.: The Construction of a Scale to Measure Subjective Effort. TU Delft, Delft (1985)
88. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (task load index): results of empirical and theoretical research. Adv. Psychol. **52**, 139–183 (1988)
89. Wittmann, M., Kiss, M., Gugg, P., Steffen, A., Fink, M., Pöppel, E., Kamiya, H.: Effects of display position of a visual in-vehicle task on simulated driving. Appl. Ergon. **37**, 187–199 (2006)
90. Burnett, G.: An evaluation of different positions for an in-vehicle navigation display. In: Proceedings of the International Conference of Traffic and Transport Psychology, Nottingham (2004)
91. Tretten, P., Normark, C.J., Gärling, A.: Warnings and placement positions in automobiles. In: Proceedings of the 17th Triennial Congress of the International Ergonomics Association, Beijing (2009)
92. Mehler, B., Reimer, B., Dobres, J., Coughlin, J.F.: Assessing the Demands of Voice Based In-Vehicle Interfaces - Phase II Experiment 3–2015 Toyota Corolla. MIT, Cambridge (2015)
93. Mehler, B., Reimer, B., McAnulty, H., Dobres, J., Lee, J., Coughlin, J.F.: Assessing the Demands of Voice Based In-Vehicle Interfaces - Phase II Experiment 2–2014 Mercedes CLA. MIT, Cambridge (2015)

94. Metz, B.: Worauf achtet der Fahrer?. Julius-Maximilians-Universität Würzburg, Würzburg (2009)
95. Dingus, T.A., Hulse, M.C., Antin, J.F., Wierwille, W.W.: Attentional demand requirements of an automobile moving-map navigation system. Transp. Res. A-Pol. **23**, 301–315 (1989)
96. Schneegaß, S., Pfleging, B., Kern, D., Schmidt, A.: Support for modeling interaction with automotive user interfaces. In: Tscheligi, M., Kranz, M., Weinberg, G., Meschtscherjakov, A., Murer, M., Wilfinger, D. (eds.) Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 71–78. ACM, New York (2011)
97. Pettitt, M., Burnett, G., Karbassioun, D.: Applying the keystroke level model in a driving context. In: Ergonomics Society Annual Conference, San Francisco (2006)
98. Purucker, C., Naujoks, F., Prill, A., Neukum, A.: Evaluating distraction of in-vehicle information systems while driving by predicting total eyes-off-road times with keystroke level modeling. Appl. Ergon. **68**, 543–554 (2017)
99. Ryan, R.M., Deci, E.L.: Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp. Educ. Psychol. **25**, 54–67 (2000)
100. Deci, E.L.: Effects of externally mediated rewards on intrinsic motivation. J. Personal. Soc. Psychol. **18**, 105–115 (1971)
101. Ryan, R.M.: Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. J. Personal. Soc. Psychol. **43**, 450–461 (1982)
102. Harackiewicz, J.M.: The effects of reward contingency and performance feedback on intrinsic motivation. J. Personal. Soc. Psychol. **37**, 1352–1363 (1979)
103. Gold, C., Naujoks, F., Radlmayr, J., Bellem, H., Jarosch, O..: Testing-scenarios for human factors research in highly automated vehicles. In: 8th Conference on Applied Human Factors and Ergonomics, Los Angeles, USA (2017)
104. Marberger, C., Mielenz, H., Naujoks, F., Radlmayr, J., Bengler, K., Wandtner, B.: Understanding and applying the concept of “driver availability” in automated driving. In: 8th Conference on Applied Human Factors and Ergonomics, Los Angeles, USA (2017)

Driverless Pods: From Technology Demonstrators to Desirable Mobility Solutions

Joscha Wasser[1,2(✉)], Cyriel Diels[1], Anthony Baxendale[2], and Michael Tovey[1]

[1] Centre for Mobility & Transport, Coventry University, Priory Street, Coventry CV1 5FB, UK
wasserj@uni.coventry.ac.uk
[2] HORIBA MIRA Ltd, Watling Street, Nuneaton, Warwickshire CV10 0TU, UK

Abstract. The advancement in automated vehicle technology allows for the application in new fields such as last mile mobility solutions. Whilst still in an early stage, some prototypes have been built with trials taking place on public roads. The main focus for this kind of vehicle however lies with proving the feasibility and subsequently the reliability of the technology. Relatively little attention has thus far been paid to the passenger comfort experience. However, it is argued here that this will become the most critical user requirement and hence the aim of this paper is to introduce a conceptual comfort framework for the design of such future vehicles and analyse and compare the current concepts in the context of the passenger comfort experience.

Keywords: Human factors · Driverless pods · Transport design · Passenger comfort

1 Introduction

With the steady advancement of automated vehicle technologies, various applications are being investigated and trialed. Last mile mobility concepts, which provide movement to and from transport hubs, are the earliest real world applications of highly automated vehicles, also referred to as SAE level 4 vehicles [1]. These vehicles do not require human intervention and can be used in confined areas with closely defined parameters, such as technology parks, university campuses, and large sites.

For the successful public acceptance and uptake of such mobility concepts, it is important that such vehicles are being perceived as comfortable, seamless, and intuitive [2, 3]. This can be expected to be especially relevant during the introductory period of such vehicles in which the general public may be hypercritical. To date, the development and evaluation of such vehicles has largely focused on the underlying technology, its performance, and other road users' perceptions [4]. In contrast, the design of the passenger experience has received relatively little attention. We here argue that the user experience will become the most critical requirement for such vehicles [5] and we here provide an initial evaluation of current concepts and framework to guide the design and evaluation of first mile and last mile mobility concepts.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_53

The basis for the research lies within two analytical tools, which both have a long history of being used in automotive design work. The Kano Model of Consumer Satisfaction will, in the context of this paper, aid to illustrate why this paper focuses on the interior design features and human factors as well as the HMI of driverless last mile mobility vehicles.

Secondly, vehicle assessment checklists were used to evaluate current vehicles. These checklists permit experts to complete an initial subjective evaluation during and after the design process. The checklists combined with the comfort models, which are discussed below provided the basis for the vehicle evaluation.

2 Kano Model

The Kano Model [6] considers product satisfaction as a function of its attributes and customers' expectations of these attributes. Applying his tool the product features are categorized in five groups based on the customer satisfaction with regard to the individual feature. The relation between all five categories can be seen in the form of a graph as illustrated in Fig. 1 [7].

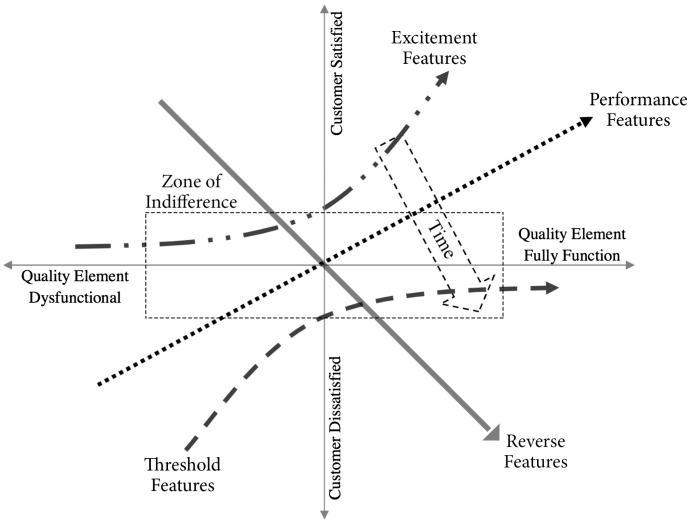


Fig. 1. Kano Model. Indicating the relation between the five categories [8].

Most customers can name the performance figures of a product and subsequently are most aware of such "performance" features (see middle function in Fig. 1). Information such as the maximum speed or the battery pack capacity of an electric vehicle is well understood. If they are executed well customer satisfaction increases and reversely, decreases satisfaction if executed poorly.

Features, which can be categorised under the "threshold" group, are essential to the product and subsequently can cause grave dissatisfaction with the product if missing

(see bottom function in Fig. 1 termed). Appropriate seats, an easy entry to the vehicle and sufficient handholds fall into this category. A product with all threshold features however will only meet the basic costumer expectation and will not add towards an increase of satisfaction.

The opposite can be observed with features that delight and excite a consumer (see top function in Fig. 1 termed "excitement"). For example, in-vehicle displays communicating contextual journey information may increase satisfaction considerably, but may not have a negative impact on satisfaction if they are not there. USB charging sockets in busses are likely to fall into this category as well as it provides an added service that is unrelated to the actual use, which is however of use to the user.

Two further categories can be included in the analysis, the "zone of indifference" and the "reverse features". The first category describes features which are present but only appreciated by very few consumers and thus can be excluded from the product. A lap-timing feature or a G-Force indicator in a road going sports car can be seen as an example for this category. Only few vehicle owners will ever make use of the feature yet it did create costs in development and implementation. "Reverse features" are not essential and cause dissatisfaction to the user due to their poor execution. An example for this could be an overly loud warning sound for closing doors on public busses.

In the context of the current paper, the Kano Model illustrates why this paper focuses on the interior design, human factors aspects and the vehicle HMI. Most users who are unfamiliar with the advances in technology required for level four autonomy, the technical specification of the vehicle (sensor and software specification) will be perceived as a "given" and therefore fall into the "threshold category" (see bottom function in Fig. 1). The expectation for the passenger is to be transported from A to B without any issues or disruption. Lengthy delays due to sensor malfunctions or arriving at the incorrect destination will likely lead to dissatisfaction.

Subsequently the attributes that may influence the user positively are the level of comfort provided for the journey, including the ease of use, which are mainly described by the performance and excitement categories [10]. However, one can argue that the basic aspects of comfort, such as providing adequate seating, extend to the threshold category, reinforcing the notion that passenger comfort is essential to a positive traveling experience.

Karlsson et al. [11] suggested that the challenge is that for those involved with the project or technology (e.g. designers, engineers) the difficulties with regard to achieving even the expected level of performance are so apparent that they fail to fully appreciate that the end user will be unaware of them and will subsequently expect more from the product.

The first generation of a new vehicle type, particularly one so strongly dependent on novel technology, is naturally centred on achieving the technological breakthrough and proving its capabilities. Whereas now, with the first trials in the public commencing, the focus will shift as the vehicle previously was assessed by experts, it is now the unaware end-user creating an initial impression of the technology. Kano Model Mano [12] and Brauer [13] already stated that the examination of utilitarian performance is not sufficient to understand customer satisfaction. Thus this state-of-the-art review analyses the attributes to ultimately facilitate user satisfaction.

For a user to be satisfied with a product several attributes have to be satisfied to some extent. With regards to last mile mobility vehicles these attributes can be grouped into three categories (see Merat et al. [13]): (1) *symbolic* attributes which cover the social influence of a product; (2) *hedonic* attributes describing the pleasure and fun derived from using the product; and (3) the *functional* attributes, the by far largest group covering the performance and effort expectancy as well as the related facilities and the price value [13]. This demonstrates that the passenger satisfaction is dependant on more than just the performance of the vehicle, further underlining the findings from the Kano Model and Mano [12] and Brauer [13] and allows for the vehicle attributes to be categorised according to the satisfaction attributes.

The authors go on to suggest that there are also significant related psychological criteria such as trust, which is strongly connected to the users expectations of the vehicle behaviour being met. The vehicle therefor is required to communicate its intentions to the passengers and outsiders, who can then for example predict the vehicle path. If the vehicle then behaves as the person predicted it would increase the trust into the vehicles abilities. A user, who is satisfied according to the previously mentioned criteria, is likely to express a growth in acceptance regarding last mile mobility vehicles.

The discussion demonstrates that a detailed understanding of comfort is required to design a vehicle, which will be accepted by the public as a viable product. Subsequently there is the need for an appropriate comfort model to be developed. We will therefor in this paper argue for a suitable proven model to be adapted for the specific application in last mile mobility vehicles.

3 Comfort Model

Although there are several schools of thought as to what we mean by comfort and discomfort, we here consider comfort a bipolar phenomenon whereby comfort is positioned at the extreme positive end, and discomfort at the extreme negative end of a continuum with a neutral point in between [14]. Thus we will regard comfort as a multifaceted experience influenced by a combination of physical, psychological, semantic, and social variables.

Therefore in this paper we use the following definitions:

Comfort: A pleasant state of well-being, ease, and physical, physiological and psychological harmony between a person and the environment.

Discomfort: A state where one experiences hardship of some sort, which could be physical, physiological or psychological.

Whilst there are varying methodologies to grade comfort, research by Ahmadpour et al. [9] demonstrated that a single graded evaluation scale is best suited to assess users perception of comfort and discomfort. In the paper the group explores the experience of flight comfort and discomfort as a combined experience and attempts to validate this thesis by two studies.

Passenger comfort forms a key user requirement in the aviation industry with 35% of passengers on intercontinental flights basing their choice of airline on comfort, placing it after flight schedules [15]. As a consequence, passenger comfort has been an important research area in aviation for some time. Recently, Ahmadpour et al. [9] tried to explore what factors in particular are important determining comfort and suggested eight key factors (see Fig. 2 right).

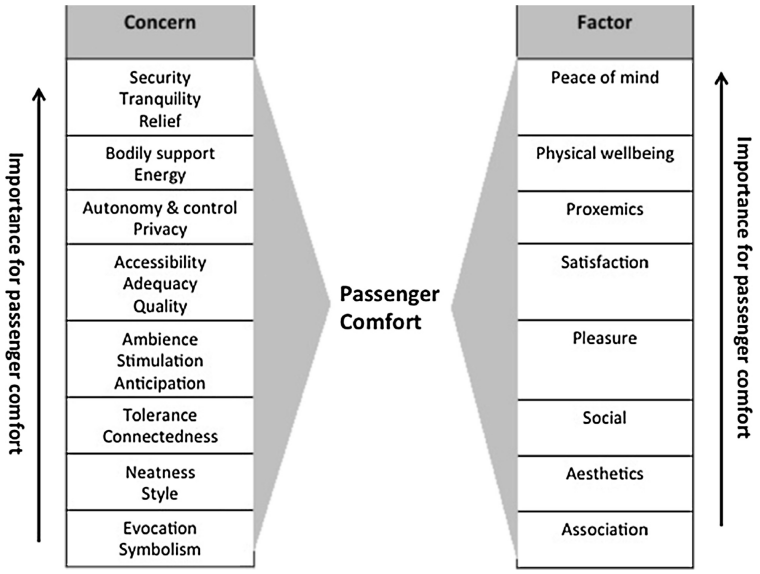


Fig. 2. Overview of the eight factors of passenger comfort in relation to their concerns [9].

We here explore if and to what extent these factors can be extrapolated across domains and be used to guide the design and evaluation of last mile mobility vehicles. The reason for this is that vehicle automation transforms the active driver into a passive monitor, and ultimately into a passenger. This shift from the driver to passenger experience, where the occupant is no longer in control of the vehicle, forms an essential aspect of vehicle automation [5]. Consequently, it will no longer be appropriate to consider traditional automotive comfort factors only and comfort factors in aviation will increasingly become of interest. At the same time, however, it is also important to realise that there are a number of fundamental differences between both modes and we cannot simply extrapolate the model to last mile mobility solutions.

There are several commonalities between the passenger experience across the two transport modes.

In both situations the passenger is not directly in control of the vehicle. Instead a journey to a prior chosen destination is undertaken during which information about the duration and location may be provided. However, unlike airplanes, last mile mobility vehicles may provide the option for users to alter the destination on route based on information about possible delays or difficulties. This demands for an interface that allows the passenger to interact with the vehicle and receive essential travel details such as

time to arrival. Further more, during the transit period, both modes may offer entertainment or connectivity to the passenger in a space, which is shared with other, possibly unknown to them passengers. Yet, a further difference is related to the fact that in contrast to an airplane in mid-air, last mile mobility vehicles will operate in environments containing a large number of other road users which are visible to the occupants.

On the basis of these similarities, the passenger comfort experience model appears to be well suited to explore the passenger expectations and needs whilst traveling in a driverless vehicle.

A similarity between these factors and the factors established by Karlsson [11] in their work regarding the assessment of vehicle interiors based on the semantic environment description (SMB) can be argued and further supports the theory of adopting the comfort model originating in aviation for an automotive context (Fig. 3).

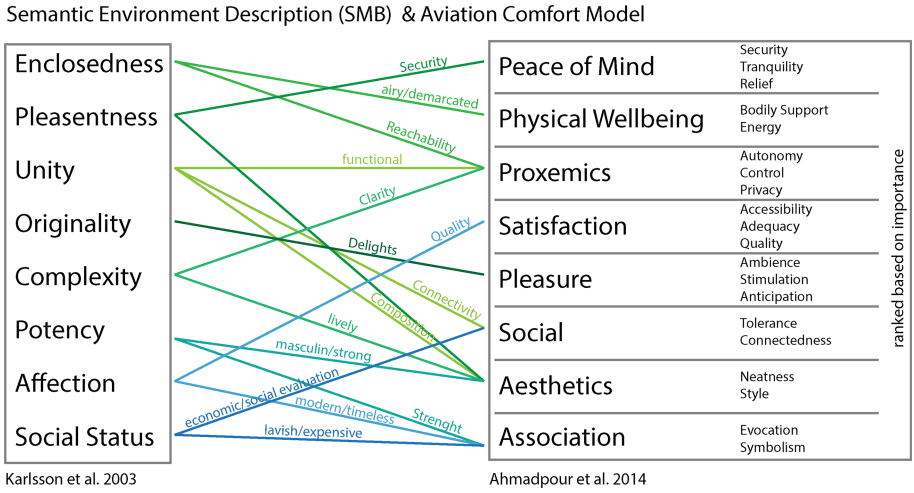


Fig. 3. Indication of similarities between Karlsson et al. [11] and Ahmadpour et al. [9]

However a number of differentiations are required prior to defining the factors for an automotive use, the service (catering) can be influential on the perception of comfort during air travel; however is currently unlikely to be experienced in an automotive environment. Equally the comfort criteria that become relevant with longer travel durations can be neglected for last mile mobility applications. This includes for example the requirement for passengers to be able to adjust their seats for greater comfort whilst resting or set up workstations, as the time spend in the vehicle is likely to be less than 15 min at a time.

Further consideration requires the time spend travelling as a survey of commuters in the Sacramento – Francisco Bay Area transportation corridor by Malokin et al. [16] indicates the wish to be able to work whilst commuting as it was stated as a reason to change the travel mode if it would permit to use the time productively. However a correlation between travel time as well as the reason for traveling and in travel time use requires further investigating, as there are conflicting responses to the willingness to work during transit. According to a survey in Germany, the demand to facilitate

activities such as working is low, as a minority only undertakes work; 77% of respondents never work on busses, 69% when traveling on trains. Instead, enjoying the landscape dominates the answers [17]. Schoettle and Sivak [18] add that the most frequent answer seen in their survey was; "Watch the road even though I would not be driving". Cyganski et al. [17] goes on to report that the morning commute is seen as particularity unpleasant, even less popular than work itself. This could be caused by the time spend traveling traditionally been seen as dead time, reducing the possibility to spend time more enjoyably. As previously mentioned, enjoying the landscape is the most frequent answer, however often mentioned activities are also listening to music, reading for leisure and relaxing, something that driverless vehicle have the potential to enable fully. This research contrasts with the common perception in regards to the main benefit of driverless vehicles as described in the current public opinion.

Exact definitions of all comfort factors identified by Ahmadpour et al. [9] are not achievable due to the fact that some attributes featured in more than one comfort category. However, in the below, the different categories as identified by Ahmadpour et al. [9] to contribute to the passenger comfort experience are discussed in the context of last mile mobility vehicles.

Peace of mind: Factor one according to the importance ranking by Ahmadpour et al. [9] is "Peace of Mind" to which attributes such as an efficient interior layout and the adjustability of the seats contribute. On a practical level, offering a secure storage for luggage and belongings is essential as well as a contribution to a passenger's peace of mind. Additionally the opportunity to rest physically and mentally contributes to the factor. Mental rest in the case of driverless vehicles in particular can be achieved in part by being able to observe the vehicle behaviour i.e. a good view to the exterior and a HMI informing the passenger. A novel factor for this use case is the ability to regain control over the journey, allowing passengers to prematurely terminate the route or choose a different path to the proposed one. As the initial vehicles will challenge the passenger's trust into the novel technology, this first factor will further gain in importance compared to traditional self-driven vehicles, over the initial trust barrier. In this regard, providing passengers with an indication that the vehicle is remotely monitored and an option to contact the human operator could also help lower the trust barrier. The discussion on trusting a self-driven vehicle extends beyond the interior and the passengers to those who will encounter the vehicle as road user. The vehicle has to communicate its intention, its chosen path or halting at a zebra crossing for example to those who are dependant on the information in order to react to it. A pedestrian will have to judge whether the vehicle has recognised him and understood his intentions. This understanding is critical on both sides particularly as these vehicles currently all operate in a shared space with pedestrians.

Physical Wellbeing: The second factor is "Physical Wellbeing" which is predominately focused on attributes providing good body support whilst giving a degree of privacy and distance from other passengers. Noted is also the amount of energy required by a passenger to utilise the vehicle, such as entering over a tall step or restraining oneself against vehicle movements. With the passengers being able to conduct non-driving tasks such as working or reading it is essential that sufficient body support is offered by the seating reducing the need to strain against uncontrolled movements caused by the

vehicle changing its path. It subsequently also aids in combating the induction of motion sickness caused in part by jolting vehicle movements.

Proxemics: "Proxemics" is the third factor, describing a feeling of autonomy and control over the space such as seats and HMI. Adjustability of the seating furniture is a reoccurring theme throughout various factors, arguable demonstrating its importance in relation to physical comfort. Beyond the physical aspects of the interior, Hall [19] categorises in the Theory of Proxemics the different zones of personal space, placing a scenario where occupants of a vehicle are seated next to each other as "intimate" which is normally reserved for affectionate contact between lovers. A reason why the central seat in a set of three is the least preferred option amongst travellers using public transport. To ease this subsequently often greatly uncomfortable experience Evans and Wener [20] propose the use of territorial furniture such as arm rests. This factor does in part contribute to the following factor of user satisfaction.

Satisfaction: The fourth factor "Satisfaction" consists, in this case of three aspects; symbolic, hedonic and functional. All contribute to a users satisfaction and are individually important and therefor the first two are discussed below in further detail as separate factors. The aspect of functionality, can be renamed in this instant as usability and is predominately based around the ease of interaction with a driverless vehicle through the HMI, easily accessible buttons and screens that can be operated without effort.

Pleasure: "Pleasure" was ranked fifth and as previously discussed, is in part related to the factor of proxemics, the individuals perception of personal space. Pleasure as a factor it greatly influences the overall user satisfaction and can be positively influenced by providing the passenger with the ability to make choices, adjustments and generally performing different activities during the transit. The haptic of the materials that the passenger comes in contact with influences the experience of pleasure as well. However, as this type of vehicle permits the occupants to focus on non-driving related tasks the quality of the infotainment system for example, included would be the ease of use and clarity, are a major contributing factor to a pleasurable journey.

Noteworthy, in regards to the previous three points, is the predicted change in the cause for the occupant's pleasure and satisfaction. Initially the passenger experience will be dominated by the novelty factor, however as the availability of the systems increases and it will become the norm to utilize these vehicles, passengers will focus on other attributes such as the way they are able to use their time and how easy the interaction with the vehicle is [2].

Social: On sixth position are "Social" factors, dominated by the circumstance that these vehicles are shared vehicles operating within a MaaS concept. This leads to a mixture of preferences such as being able to interact with others in the space or alternatively seeking personal space.

Increasingly the support of mobile network connectivity does also play a significant role as passengers want to go online, stay in touch via messenger applications or keep track on their journey [21].

Aesthetics: The "Aesthetics" are ranked seventh with attributes such as a clutter free and clean interior that is well maintained and a vehicle that conveys a sense of stability

and protection. From the outside it is important that the cabin appears inviting and that the onlooker can discern easily if there is space available to board.

Association: The eighth comfort factor is "Association", part of the symbolic satisfaction attributes, which similarly to aesthetics describes attributes such as a modern cabin, notably described as "car like" in the aviation model survey by Ahmadpour et al. [9] a cosy interior that does not agitate or tire a passenger is also a contributing factor to the association.

These factors were used as basis for developing a set of criteria to analyse the mobility concepts. The criteria can be grouped in several categories; the accessibility, analysing ingress and egress to the vehicle for the different passenger groups, the interior space, investigating the amount of space given to the passengers to move within the vehicle as well as the overall impression of the cabin, the seating provided for the users, also including standing options, the interaction with the vehicle through the HMI and the Emotional experience of traveling on-board of one of these vehicles.

Table 1. Rating scale

Not applicable	Unacceptable	Poor	Average	Good	Excellent
N/A	–	–	o	+	++
	−2	−1	0	+1	+2

Table 2. Benchmark of current last mile mobility vehicles

	ULTra	GRT	Pod Zero	Navya	EZ10	Olli
Accessibility						
• Ingress	−	+	−	o	o	+
• Egress	−	+	−	o	o	+
• Disabled access	N/A	N/A	N/A	From V4	+	N/A
Interior design						
• Visibility to exterior	−	+	o	++	+	+
• Luminosity	−	++	o	++	+	+
• Haptic	o	++	−	o	o	+
• Aesthetics	o	+	−	o	o	+
• Storage	−	−	–	−	o	−
Seating						
• Layout	o	o	o	o	o	o
• Shoulder space	−	+	−	−	o	−
• Wheelchair restraints	N/A	N/A	N/A	N/A	N/A	N/A
• Body support	+	++	−	o	+	−
HMI						
• Booking system	+	o	N/A	App	App	N/A
• Access system	+	o	o	+	+	+
• Ride information	o	−	−	o	o	−

(*continued*)

Table 2. (*continued*)

	ULTra	GRT	Pod Zero	Navya	EZ10	Olli
• Route control	N/A	o	−	−	−	−
Emotional experience						
• Sense of control	–	−	–	−	−	−
• Ability to observe	–	–	–	−	−	−
Social						
• Ability to interact with other passengers	o	o	o	o	o	o
• Ability to work	−	−	−	−	−	−
• Mobile network	N/A	N/A	N/A	N/A	N/A	N/A
Exterior design						
• Visibility to interior	−	+	o	+	o	o
• Aesthetics	+	+	o	+	o	+
• HMI	o	−	o	o	+	o

- Accessibility
  - Ingress,
  - Egress,
  - Disabled access.
- Interior space
  - Visibility to the exterior,
  - Luminosity,
  - Haptic,
  - Aesthetics,
  - Storage space.
- Seating
  - Layout,
  - Shoulder space,
  - Wheelchair restraining systems,
  - Body support.
- Interaction
  - Booking system,
  - Access system.
- Emotional experience
  - Sense of control,
  - Ability to observe (vehicle movements).

The research done by Ahmadpour et al. [9] in regards to the comfort factors "implies that comfort is greatly reduced in the presence of negative physical conditions, while an experience of joy enhances passengers' experience of comfort". Subsequently factors with the greatest potential for negative impact are the highest ranked factors according to importance to the passengers, thus require the greatest attention.

Based on this assumption some issues with the current prototypes become evident: ranking first in importance, it can be said for the factors regarding peace of mind that they are not yet addressed adequately. A good view to the exterior has been adopted by the more recent concepts (Olli, Arma and EZ10) however in all cases there is still a number of passengers seated backwards to the direction of travel and subsequently not benefiting from the improvement in visibility. Another aspect of the same factor prioritized by the travellers is a secure luggage storage, which has not been implemented in any other vehicle but the EZ10.

Physical Wellbeing as the second most important factor does need further consideration in all concepts. The problematic standing out is the lack of handrails and armrests, which could offer some stability and consequently evoke a more secure feeling in the passengers. This is a particularly important comfort aspect in driverless vehicles as passengers often cannot anticipate vehicle movements and strain themselves against it. An issue amplified by the lack of smoothness in automated driving. In hand with these considerations goes the amount of energy required to travel on this vehicle, high entry steps and seats without side support can be problematic to those who have greater mobility requirements such as the elderly and disabled.

In regard to the interior design there is a predominant public transport theme overarching all vehicles. This may be the result of designers wanting to emulate what the users are currently accustomed with from other forms of public transport to ease the transition. Recognising certain features will increase the trust in them, as they are seen as tried and tested and thus reliable. This does however bring a potential new problem with it; passengers are prompted to compare the vehicle and comfort to prior experiences with more traditional transit modes, which have been refined for many years. Driverless technology however promises to liberate the interior of vehicles from the vehicle controls and most importantly the driver, creating new space for novel arrangement that have not been fully explored in any of the concepts.

The exterior design holds a similar potential over the satisfaction of the customer, a capsule like design like the ULTra might create a sense of security due to its robust look although this example in turn has a negative impact on the interior. This signifies the importance of striking a balance between an open airy cabin for the passengers to observe the surroundings and a shape that symbolises a robust and protective vehicle.

The interaction between the vehicle and the passenger is still very limited, which may be due to most vehicles still being in the concept and trial stage, often limiting the vehicles to a set route. However as those vehicles become more advanced, the HMI will be the key point to the success of autonomy. The Olli showcases the first concept of a more liberal and intelligent destination choice, though only as an idea with the current vehicle still operating on pre-set routes. The designers will have to develop solutions to communicate between the passengers, the vehicle and other road users. Only if everyone involved feels like they are provided with sufficient information about what the other party is about to do, they will be able to trust and function alongside each other.

The most important issue with all concepts is the inadequacy for those who would benefit from the technology the most; the elderly and mobility impaired. Identified by Rode et al. [22] as an increasingly important demographic due to a decrease in car owner ship amongst them, a driverless last mile mobility service could be the key to

accessing public transport and regaining some autonomy. Therefore the accessibility to these vehicles needs to be improved with an even entry or ramp, adequate safety systems and a HMI that is within reach and not reliant on good eyesight alone.

4 Conclusion

In this paper we employed a proven comfort model in combination with known product development tools, as a basis for a new model, which details the comfort requirements for driverless last mile mobility vehicles. Using the model as guideline we analysed the current driverless last mile vehicles in regards to their passenger comfort and usability.

As result we can state there is the need for development regarding the interior as well as exterior, particularly with the passenger comfort in mind, to guarantee the success of the technology. With the importance of the meta product demonstrated previously in the review, it is evident why the developers of driverless last mile mobility solutions need to evolve their technical demonstrators into more refined products. With the number of concept trials growing in the recent months and subsequently the exposure to unaware users increasing, this is the right moment to create an acceptance for this new technology by demonstrating the benefits without creating discomfort.

However, it can be said that the vehicle concepts are not yet daring enough to break with previous ideas and attempt to fully benefit from the new liberty and space on offer. The proposed comfort model therefor is intended to provide the foundation to develop a design reference for driverless last mile mobility vehicles.

References

1. SAE International: Taxonomy and Definitions for Terms Related to On-road Motor Vehicle Automated Driving Systems. SAE International J3016_201401, Warrendale, PA (2014)
2. Nordhoff, S., Van Arem, B., Happee, R.: A conceptual model to explain, predict, and improve user acceptance of driverless 4P vehicles (2014). http://docs.trb.org/prp/16-5526.pdf. Accessed 1 Feb 2017
3. van der Laan, J.D., Heino, A., De Waard, D.: A simple procedure for the assessment of acceptance of advanced transport telematics. Transp. Res. C **5**(1), 1–10 (1997)
4. Bell, L.: Autonomous vehicles hit the street in the UK's first, 'driverless 31 pod' trial (2015). http://www.theinquirer.net/inquirer/news/2394681/autonomous-32vehicles-hit-the-streets-in-the-uks-first-driverless-pod-trial. Accessed 10 March 2017
5. Diels, C., Bos, J.E.: Self-driving car sickness. Appl. Ergon. **53**(Part B), 374–382 (2016). doi:10.1016/j.apergo.2015.09.009
6. Kano, N., Seraku, N., Takahashi, F., Tsjui, S.: Attractive quality and must-be quality. Hinshitsu **14**(2), 147–156 (1984)
7. Verdyn, D.: Discovering the Kano Model. http://www.kanomodel.com (2013). Accessed 20 Sep 2016
8. Lee, Y.C.H., Lin, S.B., Wang, Y.L.: A new Kano's evaluation sheet. TQM J. **23**(2), 179–195 (2011)

9. Ahmadpour, N., Robert, J.M., Lindgaard, G.: Aircraft passenger comfort experience: underlying factors and differentiation from discomfort. Appl. Ergon. (2016). doi:10.1016/j.apergo.2015.07.029
10. Turisová, R.: A generalization of traditional Kano model for customer requirements analysis. QIP J. (2015). doi:10.12776/qip.v19i1.407
11. Karlsson B. SA., Aronsson, N., Svensson, K.A.: Using semantic environment description as a tool to evaluate car interiors. Ergonomics **46**(13–14), 1408–1422 (2003)
12. Mano, H., Oliver, R.L.: Assessing the dimensionality and structure of the consumption experience: evaluation, feeling and satisfaction. J. Consum. Res. **20**, 451–466 (1993)
13. Brauer, K.: What is it worth? In: Aircraft Interior EXPO'06, 4 April 2006, Hamburg, Germany (2006)
14. Merat, N., Madigan, R., Nordhoff, S.: Human factor, user requirements, and user acceptance of ride-sharing in automated vehicles. In: International Transport Forum Roundtable on Cooperative Mobility Systems and Automated Driving, 6–7 December 2016
15. Vink, P.: Comfort and Design Principles and Good Practice. CRC Press, Boca Raton (2005)
16. Malokin, A., Circella, G., Mokhtarian, P.L.: How do activities conducted while commuting influence mode choice? Testing public transportation advantage and autonomous vehicle scenarios. In: Conference Paper TRB 94th Annual Meeting 2015. Paper No. 15-1179 (2015)
17. Cyganski, R., Fraedrich, E., Lenz, B.: Travel-time valuation for automated driving: a use case driven study. In: Conference Paper TRB 94th Annual Meeting 2015. Paper No. 15-4259 (2015)
18. Schoettle, B., Sivak, M.: A Survey of Public Opinion About Autonomous and Self-Driving Vehicles in the U.S., the U.K., and Australia. Report No. UMTRI-2014-21. http://deepblue.lib.umich.edu/bitstream/handle/2027.42/106590/102996.pdf?sequence=1> (2014). Accessed 20 Oct 2016
19. Hall, E.T.: The Hidden Dimension. Doubleday, Garden City (1966)
20. Evans, G.W., Wener, R.E.: Crowding and personal space invasion on the train: please don't make me sit in the middle. J. Environ. Psychol. **27**, 90–94 (2007)
21. CATAPULT Transport Systems: IMAGINE – Exploring Intelligent Mobility. Transport Systems Catapult (2016)
22. Rode, P., Hoffmann, C., Kandt, J., Smith, D., Graff, A.: Towards new urban mobility: the case of London and Berlin. In: Griffiths, P. (ed.) LSE Cities/InnoZ. London School of Economics and Political Science, London (2015)

Testing Scenarios for Human Factors Research in Level 3 Automated Vehicles

Christian Gold[1(✉)], Frederik Naujoks[2], Jonas Radlmayr[3], Hanna Bellem[4], and Oliver Jarosch[4]

[1] BMW AG, Knorrstraße 147, 80788 Munich, Germany
christian.gold@bmw.de
[2] University of Wuerzburg, Roentgenring 11, 97070 Würzburg, Germany
naujoks@psychologie.uni-wuerzburg.de
[3] Chair of Ergonomics, Technical University of Munich, Boltzmannstraße 15, 85748 Garching, Germany
jonas.radlmayr@tum.de
[4] BMW Group Research, Innovations, New Technologies, Parkring 19, 85748 Garching b. München, Germany
{hanna.bellem, oliver.oj.jarosch}@bmw.de

Abstract. The paper proposes a taxonomy for testing scenarios used in human factors research of Level 3 automated vehicles. Therefore, the literature was reviewed and testing scenarios were extracted. To categorize these scenarios, the four factors urgency, predictability, criticality and complexity of the driver response are introduced and defined. Furthermore, testing scenarios are suggested in dependence of the most important human factors research questions in Level 3 automated driving. The taxonomy thereby serves as a guidance and framework for the scenario selection and design of experiments in the context of Level 3 automated vehicle guidance.

Keywords: Highly automated driving · Automation · Take-over · Performance · Testing scenarios · Driving simulation · Conditional automation · Experimental design · Level 3 automation

1 Introduction

Automated driving is currently receiving tremendous amounts of attention in the media. While critical incidents have caused a more critical look, the public still perceives automated driving as a chance to increase safety, lower fuel consumption, increase efficiency of traffic flow and most importantly, offer new levels of comfort for the driver.

Current definitions from the SAE [1], the NHTSA [2] or the German BASt [3] suggest to split the range from manual driving to fully autonomous driving into 5 or 6 levels. Since technical limitations push the introduction of fully autonomous driving into the far future, current research focuses on Level 3 conditional [1] or high [3] automation. Both constitute that drivers will transfer full control of the vehicle dynamics to the system/the automation. There is no need to monitor the system, which

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_54

separates Level 3 conditional/high automation from another level, partly automated driving. In addition, drivers are allowed to engage in non-driving related tasks.

The driver is still in charge of taking over control in case system limits are met. This transition phase is understood to be the take-over scenario. The aforementioned definitions specify this transition phase to be completed "within a sufficient time budget" [3].

While various studies of automation effects on human performance in aviation research allow for first insights, the automotive context comprises smaller time budgets and safety margins. Thus, a thorough understanding of introducing higher levels of automation into vehicles is missing. Moreover, past research in the automotive domain has focused on distraction and the effects of secondary/tertiary tasks on the primary driving task. Lorenz et al. [4] showed that the introduction of Level 3 automated driving constitutes a change in the understanding of underlying effects: drivers and their interaction with the vehicle change from a parallel-task paradigm to a sequential-task paradigm. Drivers are not distracted anymore, but are allowed to switch their full attention to non-driving related tasks while still being responsible as fallback level. This introduces new research questions such as effects of vigilance decrements, drowsiness, fatigue or mode awareness issues in or prior to a safety-critical take-over scenario.

Obviously, these effects should always be evaluated in specific use cases to allow a thorough understanding. Any use-case will comprise the traffic scenario or situation in which a request to intervene (RtI) is issued. The definition of the specific testing scenario is thus of high importance in the evaluation of the human performance in take-over scenarios. This includes assessments of time and quality aspects of the required driver action to safely take-over manual control as well as the drivers' subjective assessments (e.g. comfort, understandability, acceptance, etc.) of the take-over itself.

Schneider [5] states that there is no clear definition of which parameters are essential for describing a specific traffic scenario. The temporal and spatial dilatation highly depend on the research questions, for example, which other vehicles are taken into account when defining the situation or which future timeframe is relevant for "this situation". For example, the maximum possible human performance in a take-over situation is unlikely to show in an uncritical take-over scenario, when the demand of the situation does not afford the driver to react as accurate as possible. On the other hand, comfort or acceptance ratings by drivers will show more reliable results in less critical scenarios. In addition, Schneider [5] differentiates objective performance data and how these are subjectively perceived by drivers.

The big variety of the situational context of testing scenarios underlines the need to distinguish different influencing factors and define suitable testing scenarios to examine relevant effects on and in a take-over.

This work defines and classifies testing scenarios that allow researchers in the project Ko-HAF (cooperative, highly automated driving), funded by the German Ministry of Economic Affairs and Energy, to address the human factors issues associated with Level 3 automated driving.

2 The Driver in a Level 3 Automated Vehicle

Human factors studies have demonstrated that automation usually affects operator state and performance, for example, by a loss of skill [6], loss of situation awareness [7] or overreliance on the automation [8], as the human's responsibility is shifted from an active performing to a passive monitoring task [8, 9]. During Level 3 automated driving, similar automation effects have been reported [10], calling for a careful design of automated driving functions.

Functional system limits (e.g., missing lane markings, high curvature, etc.) or system failures (e.g., sensor malfunctions) may render the automated vehicle guidance temporarily infeasible and necessitate manual vehicle control. However, vehicle automation has been shown to go along with changes in the drivers' state such as distraction from driving-related tasks by increased engagement in non-driving related tasks (NDRT) [11] as well as drowsiness and vigilance decrement. As the driver is relieved from the need to monitor the traffic situation continuously, it has to be expected that transitions to manual driving will consist of a task-switch from the NDRT to manual vehicle control. Task switches usually go along with performance decrements such as increased reaction times and error rates [12]. Therefore, safe and efficient transitions to manual driving are of paramount importance. This fact has motivated numerous studies on take-over performance and the design of efficient requests to intervene (RtI).

As undisputable as the safety of control transitions may be, the acceptance of Level 3 automated driving may also depend on other factors which have recently attracted considerable research interest, such as driving style (e.g. characteristics of acceleration or deceleration of the automated vehicle), driving strategy (e.g. cooperation with other road users) and design characteristics of the Human-Machine Interface (HMI). Disregarding the changing role of the human in these design aspects of Level 3 automated driving may well result in non-acceptance and disuse of the technology. For example, if the HMI does not support the driver to understand intentions and actions of the automation (e.g. overtaking, strategic driving manoeuvres), distrust and under-reliance on automated driving functions may prevent potential benefits of automated driving from being achieved.

3 Testing Scenarios in Human Factors Research

A growing body of empirical studies has been conducted on the research topics mentioned above, using a range of different testing scenarios. One example of system limits during Level 3 automated driving are missing or ending lane markings on motorways, which are normally needed for estimating the vehicle's future path reliably. System limits caused by secondary lines (e.g. because of work zones) have been included in studies by Melcher et al. [13] and Kerschbaum et al. [14]. Larsson et al. [15] investigated lateral control failures in which the lateral vehicle guidance was discontinued during a straight or curved road segment. Mok et al. [16] and Walch et al. [17] report findings on driver performance during take-over situations that are caused by high road curvature. Naujoks et al. [18] investigated three different lateral guidance

boundaries in which the lateral vehicle guidance could not be kept up because of missing lane markings, secondary lines due to a work zone or high curvature. Sudden obstacles on the road ahead, such as broken down vehicles or lost cargo, have also been investigated in driving simulator studies [19–21]. An overview of different testing scenarios, divided into system limits in longitudinal and lateral control as well as system failures, is provided in Table 1. We expect, although the available empirical evidence concerning human performance in the area of Level 3 automated driving has increased, there is still no common ground regarding which scenarios should be used when conducting empirical studies. In the following section, a taxonomy of testing scenarios is proposed. The taxonomy is meant to serve as guidance for human factors practitioners. At the same time, using comparable testing scenarios will increase comparability between studies and facilitate scientific knowledge building.

Table 1. Overview about test-scenarios used in research in automated driving.

Type of event	Description	References
System limit, longitudinal traffic	Pedestrian entering road in work zone	[22]
	Obstacle on the road (e.g. broken-down vehicle in lane of host vehicle)	[21, 23]
	Cut-in vehicle is not detected	[15]
	Road works require change to another lane	[24]
	Bad visibility because of fog	[25]
System limit, lateral control	Blurred lane markings	[26]
	Missing lane markings	[18]
	Secondary lane markings (e.g. because of a work zone)	[13]
	High road curvature	[18]
System failure	Insufficient brake intervention by automated driving function	[20]
	Insufficient quality of lateral vehicle control	[16]
	Sensor failure in straight or curved road segment	[18, 27]

4 Taxonomy of Testing Scenarios

This paper aims to develop a taxonomy of testing scenarios. Therefore, scenarios have to be structured and categorized regarding the main factors that determine the driver's response and behavior. Based on a review of the literature reported in the previous chapters, the factors urgency, predictability, criticality, and driver response have been selected to structure and categorize the testing scenarios.

Urgency. The urgency of a testing scenario indicates how fast a take-over reaction of the driver is required. This is associated with the time budget the driver has to take over control and react to the system limit. The time budget can be rather limited, for example if an obstacle is blocking the lane in front of the vehicle, or very long, for example in case of approaching a highway exit miles away. Taking over the vehicle is known to be

challenging, as the task switching from possible NDRTs to the manual driving task requires attentional resources and time-consuming re-configuration of the driver's physical and cognitive state to meet the demands of the respective scenario [26]. Restricting the available time for these processes may increase the speed of the take-over process, as the driver will invest cognitive and attentional resources, but may likewise decrease the quality of the subsequent driver input [19]. The urgency of a take-over situation is therefore considered to be a very important category in a taxonomy of testing scenarios.

Predictability. The predictability of a testing scenario refers to the timing of the detection of a take-over and thereby influences the timing of the RtI. Scenarios with a high predictability may be known in advance based on information in the backend. Assuming that a specific automated system is only available on highways, the end of a highway, and thereby the end of the automation scenario, would have a high predictability. Even dynamic system limits like the end of a traffic congestion, or insufficient weather conditions can be highly predictable, with help of car-to-car communication, or via a common database like the safety server in the Ko-HAF project [28]. Short term system limits, detected by onboard sensors, are considered to have a low predictability, like an accident occurring right in front of the automated vehicle.

The predictability also correlates with the time budget that can be provided by the system for the driver to take over vehicle control [29]. In many cases, a RtI can be displayed much earlier in scenarios with a high predictability. This determines the characteristic of the warning and thereby a driver's reaction, wherefore the predictability is considered for the taxonomy.

Criticality. The criticality of a testing scenario scales the cost when failing to take-over vehicle control in time. The impact of failing in a take-over scenario can be rather low, like missing to exit the highway, or very severe, for example when colliding with other road users or obstacles. Therefore, criticality is determined by the situation and the characteristic of the automated systems. A minimal risk maneuver for example can reduce the risk of a collision and thereby decrease the criticality of a testing scenario. The criticality determines the strategy and behavior of the driver in a take-over situation, as the risk involved determines the strategy and behavior [30], via the value system and personal aims of the driver.

Drivers' Response. The drivers' response is the required action in reaction to a RtI. In dependence of the testing scenario, different driver reactions may be necessary or appropriate. They can either be complex (e.g. evasive maneuver), or rather simple (e.g. stabilizing the vehicle in its lane) [18]. The complexity of the drivers' response influences the timing of the take-over reaction as well as the quality of the driver's input. The driver response is therefore considered another factor when categorizing testing scenarios for Level 3 automated vehicles.

Table 2 lists the four factors and proposes a three-level classification (low, medium, high) for each of the factors. Although the selected factors correlate in many cases, testing scenarios exist for almost any combination of urgency, predictability, criticality, and driver response.

Table 2. Classification of the factors urgency, predictability, criticality and driver response.

Factors	Low (1)	Medium (2)	High (3)
Urgency	High time budget	Medium time budget	Small time budget
Predictability	Near-term detection of the system limit	Predictable, but occurrence dependent on situational conditions	Known from backend, map, V2V-communication
Criticality	Low safety risk	Increased safety risk	High safety risk
Driver Response	Low complexity (e.g. stabilizing)	Medium complexity (e.g. steering)	High complexity (e.g. lane change)

Table 3 summarizes the scenarios listed in Chapter 3. The list is complemented with other scenarios from literature and new scenarios that have not been used in studies on automated driving yet. Table 3 also proposes a classification of the scenarios regarding the four factors described above. The classification was conducted by a board of experts from science and industry within the consortium of the Ko-HAF project.

Table 3. Classification of different testing scenarios

No.	Name	Urgency	Predictability	Criticality	Driver response
1	Sensor failure (Subsystem)	1	1	1	1
2	Sensor failure (Total)	3	1	2–3	1–2
3	End of highway	1	3	1–2	1–2
4	Lane change to deceleration lane not possible (e.g. because of traffic on target lane)	2	2	1	3
5	Lane change from entrance ramp not possible	3	2	2	3
6	Road narrows (known from backend)	1	3	2	2
7	Road narrows (detected by on-board sensors)	3	1	2	2
8	Danger zone/obstacle ahead (known from backend)	1	3	1–3	1–3
9	Danger zone/obstacle ahead (detected by on-board sensors)	3	1	1–3	1–3
10	Loss of reference signals (e.g. lane markings missing)	3	1	2–3	1
11	Predictable loss of reference signals (known from backend)	1	3	2–3	1

5 Recommended Testing Scenarios

Based on the classification introduced above and depending on the aim and hypotheses of a take-over study, different requirements for the testing scenario can be derived. When interested in the maximum driver performance within a take-over situation, demanding scenarios are needed, equivalent to a certain urgency, criticality and complexity of the driver response. Although these scenarios will be very rare in everyday usage of Level 3 automated systems, they are a valid tool for designing and evaluating Level 3 automated vehicle functions and for understanding the driver's behavior in a take-over situation. If the driver's comfort is in focus of the study, testing scenarios that are more uncritical are appropriate, which are characterized by a good predictability, low criticality and low urgency.

With the aid of the information given in the tables above, testing scenarios can be selected in dependence of the selected research focus. Furthermore, similar scenarios from a human factors point of view can be identified and used to introduce variance in the take-over situations but not in the drivers' response. This paper thereby provides guidance when designing take-over experiments and a framework for projects and research groups concerned with Level 3 automated vehicle guidance. Table 4 recommends testing scenarios in dependence of the most common research questions of Level 3 automated vehicles.

Table 4. Recommendation of testing scenarios in dependence of the research focus

Research focus	Urgency	Predictability	Criticality	Driver response	Corresponds to situation no. (Table 3)
Maximum driver performance	3	1	3	2–3	2, 9, 10
Mode awareness / HMI	1–3	1–3	1–3	1–3	May be relevant in any scenario
Driver comfort	1–2	3	1	1–3	3, 8

Acknowledgement. This work results from the joint project Ko-HAF - Cooperative Highly Automated Driving and has been funded by the Federal Ministry for Economic Affairs and Energy based on a resolution of the German Bundestag.

References

1. Committee, S.O.-R.A.V.S.: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Technical Report J3016_201401 (2014)
2. Blanco, M., Atwood, J., Vasquez, H.M., Trimble, T.E., Fitchett, V.L., Radlbeck, J., Fitch, G. M., Russell, S.M., Green, C.A., Cullinane, B.: Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts (2015)
3. Gasser, T.M., Westhoff, D.: BASt-study: definitions of automation and legal issues in Germany. In: Proceedings of the 2012 Road Vehicle Automation Workshop (2012)
4. Lorenz, L., Kerschbaum, P., Hergeth, S., Gold, C., Radlmayr, J.: Der Fahrer im Hochautomatisierten Fahrzeug. Vom Dual-Task zum Sequential-Task Paradigma. In: TÜV Tagung Fahrerassistenz, München (2015)
5. Schneider, J.H.: Modellierung und Erkennung von Fahrsituationen und Fahrmanövern für sicherheitsrelevante Fahrerassistenzsysteme. Fakultät für Elektrotechnik und Informationstechnik. Technische Universität Chemnitz, Chemnitz (2009)
6. Stanton, N.A., Marsden, P.: From fly-by-wire to drive-by-wire: safety implications of automation in vehicles. Saf. Sci. **24**(1), 35–49 (1996)
7. Endsley, M.R.: Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics **42**(3), 462–492 (1999)
8. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse, abuse. Hum. Factors J. Hum. Factors Ergon. Soc. **39**(2), 230–253 (1997)
9. Parasuraman, R., Mouloua, M., Molloy, R., Hilburn, B: Adaptive function allocation reduces performance cost of static automation. In: Jensen, R.S., Neumeister, D. (eds.), 7th International Symposium on Aviation Psychology, pp. 37–42 (1993)
10. Saffarian, M., De Winter, J., Happee, R.: Automated driving: human-factors issues and design solutions. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting (2012)
11. Naujoks, F., Purucker, C., Neukum, A.: Secondary task engagement and vehicle automation–comparing the effects of different automation levels in an on-road experiment. Transp. Res. Part F **38**, 67–82 (2016)
12. Allport, D.A., Styles, E.A., Hsieh, S.: Shifting Intentional Set: Exploring the Dynamic Control of Tasks. MIT Press, Cambridge (1994)
13. Melcher, V., Rauh, S., Diederichs, F., Widlroither, H., Bauer, W.: Take-over requests for automated driving. Procedia Manuf. **3**, 2867–2873 (2015)
14. Kerschbaum, P., Lorenz, L., Bengler, K.: Highly automated driving with a decoupled steering wheel. Proc. Hum. Factors Ergon. Soc. Annu. Meet. **58**, 1686–1690 (2014)
15. Larsson, A., Kircher, K., Hultgren, J.A.: Learning from experience: familiarity with ACC and responding to a cut-in situation in automated driving. Transp. Res. Part F **27**, 229–237 (2014)
16. Mok, B., Johns, M., Gowda, N., Sibi, S., Ju, W.: Take the wheel: effects of available modalities on driver intervention. In: 2016 IEEE Intelligent Vehicles Symposium (IV) Gothenburg, 19–22 June 2016 (2016)
17. Walch, M., Lange, K., Baumann, M., Weber, M.: Autonomous driving: investigating the feasibility of car-driver handover assistance. In: 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 11–18. ACM, Nottingham (2015)

18. Naujoks, F., Mai, C., Neukum, A.: The effect of urgency of take-over requests during highly automated driving under distraction conditions. In: Ahram T., Karowski W., Marek T. (eds.), Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014, AHFE Conference, pp. 2099–2106. Krakau (2014)
19. Gold, C., Damböck, D., Lorenz, L., Bengler, K.: "Take over!" How long does it take to get the driver back into the loop? Proc. Hum. Factors Ergon. Soc. Annu. Meet. **57**, 1938–1942 (2013)
20. Strand, N., Nilsson, J., Karlsson, I.M., Nilsson, L.: Semi-automated versus highly automated driving in critical situations caused by automation failures. Transp. Res. Part F **27**, 218–228 (2014)
21. Naujoks, F., Purucker, C., Neukum, A., Wolter, S., Steiger, R.: Controllability of partially automated driving functions—does it matter whether drivers are allowed to take their hands off the steering wheel? Transp. Res. Part F **35**, 185–198 (2015)
22. Gold, C., Lorenz, L., Damböck, D., Bengler, K.: Partially automated driving as a fallback level of high automation. 6. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren (2013)
23. Radlmayr, J., Gold, C., Lorenz, L., Farid, M., Bengler, K.: How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, (2014)
24. Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., Othersen, I.: The interaction between highly automated driving and the development of drowsiness. Procedia Manuf. **3**, 6652–6659 (2015)
25. Vollrath, M., Schleicher, S., Gelau, C.: The influence of cruise control and adaptive cruise control on driving behaviour—a driving simulator study. Accid. Anal. Prev. **43**, 1134–1139 (2011)
26. Zeeb, K., Buchner, A., Schrauf, M.: Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. Accid. Anal. Prev. **92**, 230–239 (2016)
27. Larsson, P., Johansson, E., Söderman, M., Thompson, D.: Interaction design for communicating system state and capabilities during automated highway driving. Procedia Manuf. **3**, 2784–2791 (2015)
28. Lüke, S.: Projekt Ko-HAF – Nächste Schritte zum hochautomatisierten Fahren. Kooperationsforum Fahrerassistenzsysteme, Stuttgart (2016)
29. Naujoks, F., Forster, Y., Wiedemann, K., Neukum, A.: A human-machine interface for cooperative highly automated driving. In: Advances in Human Aspects of Transportation, pp. 585–595. Springer, Cham (2017)
30. Michon, J.A.: A critical view of driver behavior models: what do we know, what should we do?. In: Human Behavior and Traffic Safety, pp. 485–524. Springer, New York (1985)

Autonomous Vehicles: Reliability of Their Perception of the World Around Them and the Role of Human Driver

Ata Khan(✉)

Carleton University, Ottawa, ON K1S5B6, Canada
ata_khan@carleton.ca

Abstract. According to the vision of technology promotors, in a fully autonomous vehicle design a driver need not be in the loop even in a complex driving environment. Recognizing the technological, legal, social and economic constraints, this vision may not become reality for many years into the future. However, a "cognitive vehicle" provides a realistic step towards full automation. A cognitive vehicle is an intelligent vehicle designed to have a seamless non-distracting interface with the driver and for reasons of liability and cost, the driver remains "in the loop". As a part of this vehicle's design, an adaptive driver assistance system, among many other features, can have the longitudinal and transverse control capability. The driver has the option to turn on the active safety feature and can override the automation feature to make difficult decisions. On the other hand, active feature takes over driving if the driver is highly distracted or disabled.

Keywords: Autonomous vehicle · Human-factors · Cognitive vehicle · Perception reliability · Automation in driving

1 Introduction

A cognitive vehicle design differs from the design of a fully autonomous vehicle in which, according to technology promoters, the driver need not be in the loop. This paper is intended to present the issues of accuracy and reliability in autonomous driving and suggests cognitive vehicle design features as solutions to these issues.

2 Levels of Technological Advances

Khan, Bacchus and Erwin [1] provided a projection of technological advances as a part of a paper on policy challenges of increasing automation in driving (Fig. 1). On the technology development front, innovations have progressed as projected along the continuum between conventional fully human-driven vehicles and autonomous vehicles which partially or fully drive themselves and which may ultimately require no driver at all.

Along this continuum are notable automated vehicle technologies that enable a vehicle to assist and make decisions for the human driver. The automated driving features can be a part of semi-autonomous or fully autonomous driving. Examples of

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_55

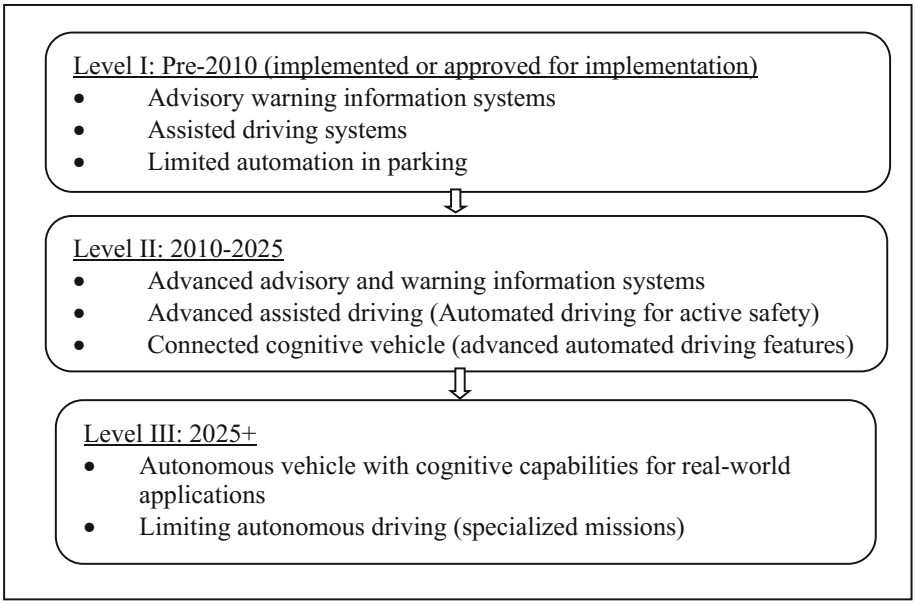


Fig. 1. Levels of technological advances and timeframe [1].

applications include operational assistant or autopilot in heavy traffic, keep-your lane system, automated parking system and advanced adaptive control systems. A subset of developing technologies are packaged as driving assistance system that include crash warning systems, adaptive cruise control, lane-keeping systems, and self-parking technology. The Society of Automotive Engineers (SAE) and the US National Highway Traffic Safety Administration (NHTSA) have created a taxonomy to help clarify this continuum [2, 3]. These are summarized in Table 1.

There are different views on how autonomous and connected-vehicle technologies relate to each other. Some technology developers believe that these are closely related, and others consider that autonomous vehicles can function on sensor data alone and need not obtain data from other vehicles [2]. Connected vehicles are designed to interact with other vehicles or the transportation infrastructure [4, 5].

If vehicles share driving environment information with each other within a certain distance via information technology, this could aid autonomous driving decisions. Taking a step further, if vehicles share sensor information with nearby vehicles, this could give an autonomous vehicle additional information that could be used as refined data for decision-making. In order to take advantage of these technological capabilities, it is a common belief among researchers that connected vehicle technology will be central to achieving automation in driving.

Table 1. Level of automation in driving Adapted from Anderson et al. [2] and SAE International (J3016 [3].

Level of automation	Description
Level 0	No-automation: The human driver is in complete control of all functions of the vehicle
Level 1	Function-specific Automation: One or more specific control functions are automated and these operate independently of each other. But, the driver has overall control
Level 2	Combined Function Automation: This level involves automation of at least two primary control functions designed to work in unison to relieve the driver of control of those functions (e.g. steering and acceleration); shared authority when the driver cedes active primary control in certain limited driving situations; the driver must remain attentive all the time
Level 3	Limited Self-Driving Automation: The vehicle enables the driver to cede full control of all safety-critical functions under certain traffic or environmental conditions; transition back to driver control under defined conditions. The driver is expected to be available for occasional control
Level 4	High Automation, Self-Driving Automation: The vehicle is designed to perform all safety-critical driving functions and monitor roadway conditions for an entire trip; the driver provides destination or navigation input; Driver can still take control
Level 5	Full Automation: Completely automated and may not even have a steering wheel or a brake and gas pedal. The human occupant enters a destination

3 Issues of Accuracy and Reliability

The issues of accuracy and reliability have to be overcome for driver acceptance of highly automated and fully autonomous vehicles. According to literature, lack of accuracy refers to crashes of autonomous vehicles as well as false alarms. The reliability attribute seeks to ensure that system performance does not deteriorate over time or under difficult driving environments characterized by geometric design features of facilities, traffic control, vehicular traffic density, cyclist and pedestrian traffic, and environmental conditions [6].

Driver trust and reliability are interrelated and from a human factors perspective, advanced technology vehicles can be designed that enhance driver trust (Fig. 2). As noted earlier, an example of improving accuracy is by minimizing false alarms. Stable performance is an indicator of reliability. Understandability implies provision of clear and useful information. Message framing refers to the instruction to be conveyed by the message (i.e. permissive, advise/inform, alert/warn, prohibit). Message coordination calls for coordination of various systems (i.e. vehicle-infrastructure, and in-vehicle systems) in order to improve driver trust of safety systems. Familiarity with the system and message can enhance driver trust and response [6].

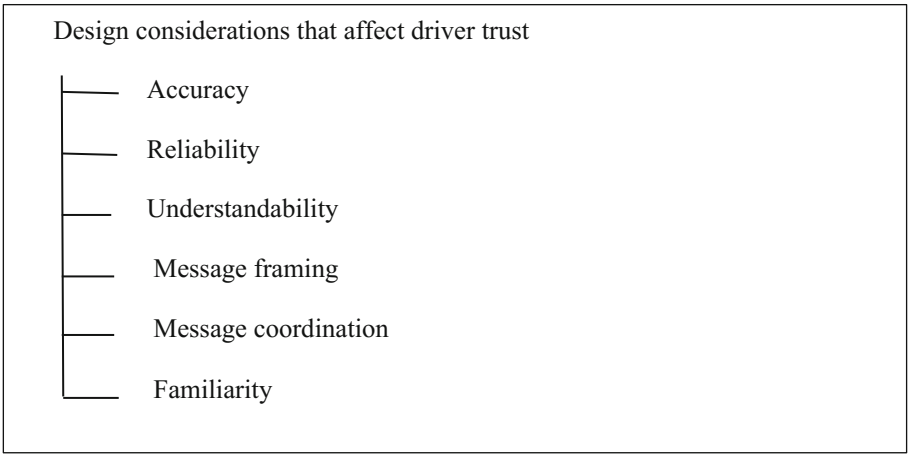


Fig. 2. Design factors for driver trust (Adapted from Reference [6]).

4 The Cognitive Vehicles

The role of the human driver and the driver-vehicle interface continue to be of research importance. At present, there is lack of consensus on the full autonomy for the vehicle. According to researchers and automotive industry experts, the next step is the development of a cognitive vehicle that will integrate intelligent technology and human factors for providing non-distractive interface for safety, efficiency and environmental sustainability in driving. A cognitive connected vehicle can function in highly automated and fully autonomous modes [7].

Table 2 describes the capabilities of a cognitive connected vehicle. Technological forecasts suggest that a number of cognitive vehicle features can be achieved with R&D efforts [8, 9].

The above brief review indicates that over the years, the vision of the *intelligent vehicle* became increasingly ambitious. An intelligent vehicle in its advanced form should have cognitive features that can perform non-distracted and non-aggressive driving tasks. A *cognitive vehicle* is intended to assist the driver, and if necessary in dangerous conditions, it has the capability to take corrective active safety action if the driver is incapacitated or highly distracted or if the driver wishes the vehicle to take over driving. However, driving an intelligent cognitive vehicle does not take the driver out of the loop. The design attributes of a cognitive vehicle are influenced by human factors in driving. According to a recent news article, development of 'human-like' self-driving technologies is attracting investor capital [10].

The above information suggests that design advances of the cognitive vehicle can potentially overcome the issues of accuracy and reliability while keeping the human driver in the loop. Descriptions of driver-vehicle interface, multi-functional driver assistance system, and advanced driving capability such as collaborative driving are provided in another paper by the author presented at this conference.

Table 2. Cognitive vehicle features for human and machine control Source: Khan et al. [7].

Cognitive vehicle features	Features for human control	Features for adaptive longitudinal and lateral control system
Awareness of position and surroundings	X	X
Ability to gather data and transmit data	X	X
Ability to process data	X	X
Ability to cooperate/collaborate	X	X
Communication for active safety		X
Informs the driver about situations (warnings, advice)	X	
Diagnostic capability	X	X
In case of crash, capability to send and receive information	X	X
Ability to provide non-distractive user interface	X	X
Infotainment capability	+	+

+ This feature does not relate to traffic service and safety objectives.

Awareness of position and surroundings can be enhanced by the integrated use of technology and methodology. The following are example technology components [11].

- Dedicated short range communication system (DSRC) and long term evolution (LTE) communication system (when required)
- Forward & side radars
- Road geometric feature processor
- Traffic data in subject and surrounding lanes obtained from multiple purpose sensors and communication system
- GPS-map/curve speed processor
- Lane tracking sensor processor
- Road weather scanning feature
- Static object location scanner

The situational awareness model integrates the data inputs and dynamically updates the driving environment information that is required by the driver assistance system part of the cognitive vehicle.

Technological developments continue to improve the driving assistance system. The initial version of this system did not interact with the human driver. Next came the current advanced driver assistant noted as Level II technology in Fig. 1. Further research is expected to deliver the new generation driver assistance system with additional capabilities, especially improved human driver-automation interface.

With an eye on the motorists' acceptance, the role of the human driver and the driver-vehicle interface continue to be recognized as areas of research importance. Although

there is a lack of consensus on the full autonomy for the vehicle in every day driving, researchers and automotive industry experts appear to be calling for a cognitive vehicle which will integrate intelligent technology and human factors for providing non-distractive interface for safety, efficiency and environmental sustainability in driving. Technological forecasts (Level III in Fig. 1) suggest that cognitive vehicle features can be achieved with R&D efforts.

Various research groups have expressed views on the attributes of the cognitive vehicle (8, 9, 11). In arriving at the views that are advanced here, current and recent developments in communication technology and the use of artificial intelligence in vehicle design have been taken into account. A cognitive vehicle is expected to have attributes that are noted in Table 2. It is assumed that even in the in the human control mode of operating the cognitive vehicle, the control of the vehicle can be taken over by the vehicle in emergency.

To further elaborate on the cognitive vehicle features that are required for the design of the crash avoidance part of an advanced driving assistance system the following requirements should be met [7, 11].

- Information availability on the complete situational context (consisting of vehicle, environment and driver) is necessary and is within reach.
- In designing systems for the cognitive vehicle, the technology and methodology should be able to deliver the capabilities of context awareness, adaptive interfaces, driver intention recognition, and driver monitoring.
- Ability to measure distance between vehicles on a real time basis and to define critical distance under the dynamic driving, road and environmental conditions.
- Self-calibration capability (i.e., updating the probability of the driver to perceive the correct distance between vehicles vis-à-vis the safe stopping distance in case of an emergency), monitoring driver attention or distraction on the basis of corrective actions taken (or not taken).
- Deciding when to issue alert and the nature of alert, and when to activate safety feature.

5 Bayesian Artificial Intelligence

Bayesian Artificial Intelligence (AI) can overcome the accuracy and reliability problems by serving as an overseer of automation functions and if carefully designed, it will provide well-defined alerts for the human driver to make highly complex decisions. In this section, an illustration based on data extracted from driving simulator tests is used to explain the concept and method.

Bayesian Artificial Intelligence (AI) has been used by the author to model driver actions and role of technology in taking corrective decision. In association with driving simulator data it is helping to advance surrogate safety measure models as aids to study technology and human factor-based solutions for the distracted driving problem [7].

Korb and Nicholson [12] view artificial intelligent as the “intelligence developed by humans, implemented as an artefact”. The Bayesian artificial intelligence can integrate two facets of problem solving in the design of cognitive features of connected and

autonomous vehicles. The first one models a human action (e.g., non-distracted non-aggressive driving). The second one is modelling rationality in the form of "optimal decision".

Based on the above noted features of the Bayesian AI, an algorithm was developed by Khan [11] for the major functions of the longitudinal and lateral control system of the cognitive vehicle (Fig. 3). It is intended for on-line use as a part of cognitive vehicle design.

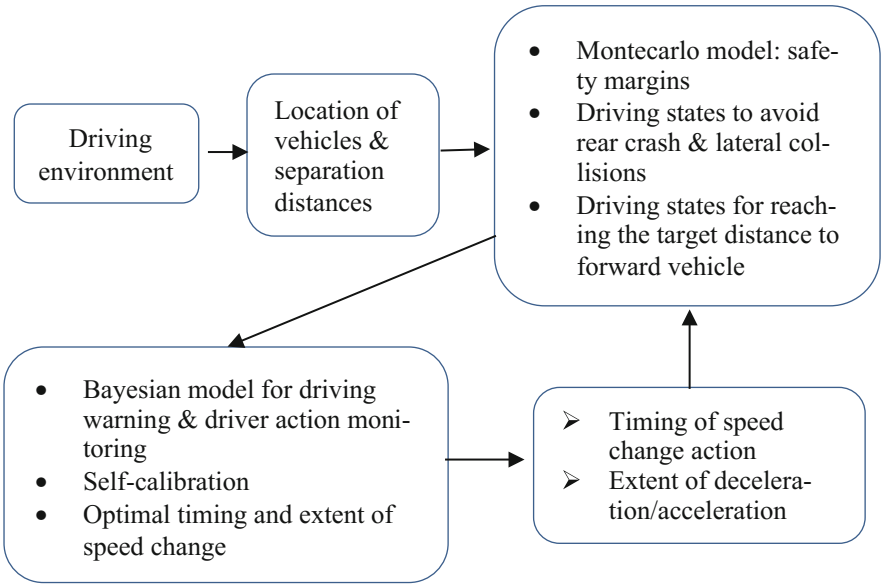


Fig. 3. Major functions of the longitudinal and lateral control system

With reference to Fig. 3, in the driving environment, frequently a driver within the traffic stream finds it necessary to decelerate at a very high rate in order to avoid a collision with the lead vehicle which applied brakes to reduce speed for some unknown reason. Another situation that necessitates hard braking is when a vehicle abruptly changes lanes and joins the traffic stream in front of the driver. In these and many other similar situations, drivers that are not assisted with automation features send shock waves in the traffic stream that sometimes result in traffic accidents. In such traffic environment, even the most favourable outcome is delays experienced by a large number of motorists.

Table 3 shows that when safety margin is violated, the probability of crash becomes unacceptably high.

Table 3. Monte Carlo simulation of safety margin

Scenario & speed of vehicles	Distance required: range & probability distribution	Available distance range & prob. distribution	No. of crashes out of 20 simulations	Probability of crash (based on all runs)
A 103.492 km/h	57–72 m Triangular mode = 57 m	42–70 m Triangular mode = 56 m	15	0.82
B 103.492 km/h	57–72 m Triangular mode = 64.5 m	42–70 m Triangular mode = 56 m	18	0.92
C 103.492 km/h	57–72 m Triangular mode = 72 m	42–70 m Triangular mode = 56	19	0.96

In order to illustrate driving without automation features, data from a driving simulator was used. A snapshot of driving trajectories was extracted from a simulation run (Figs. 4, 5, 6 and 7). This driving profile shows high levels of deceleration as well as acceleration. Also, severe drop in speed and headway distance can be seen. Next, high level of acceleration was applied to reach the preferred (target) speed. These driver actions exhibit much driver stress which could result in a collision. If the vehicle featured an advanced driving assistance system, the driving experience could be stress-free and the risk of a collision could minimized.

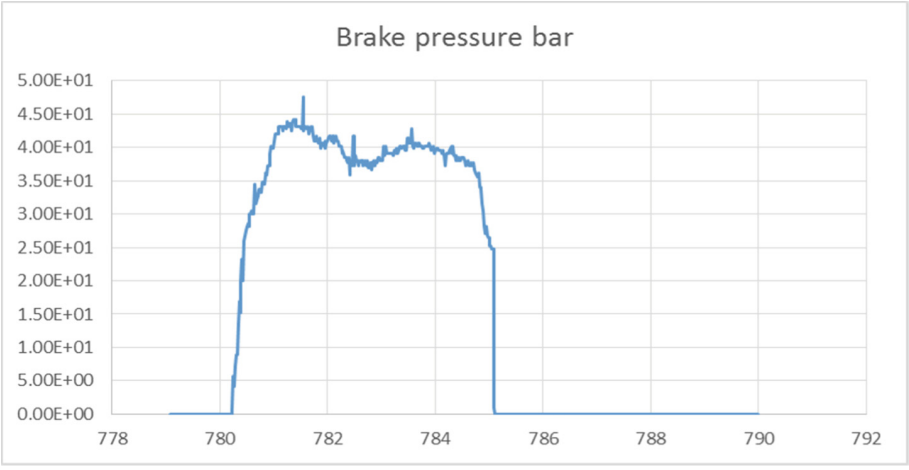


Fig. 4. Brake action under human control. Notes: (1) Horizontal axis is time stamp (sec). (2) Vertical axis is brake pressure. (3) Lead car brakes at time stamp 779.2 (sec).

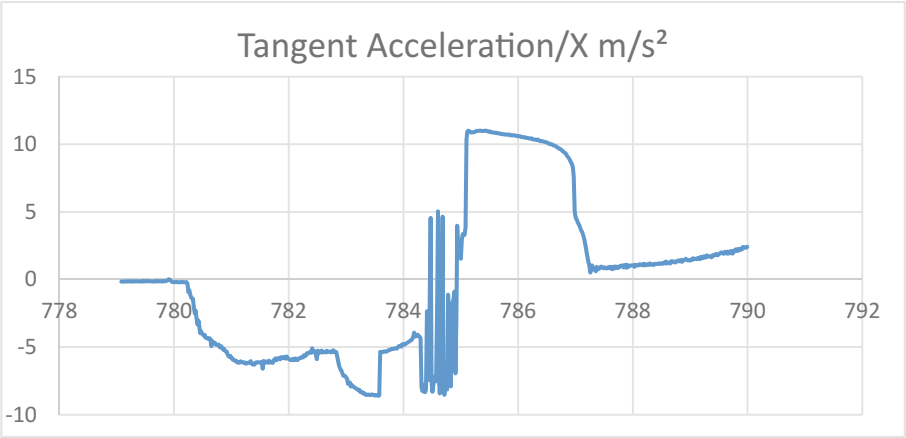


Fig. 5. Tangential acceleration under human control. Notes: (1) Horizontal axis is time stamp (sec). (2) Vertical axis is acceleration/deceleration (m/sec/sec). (3) Lead car brakes at time stamp 779.2 (sec).

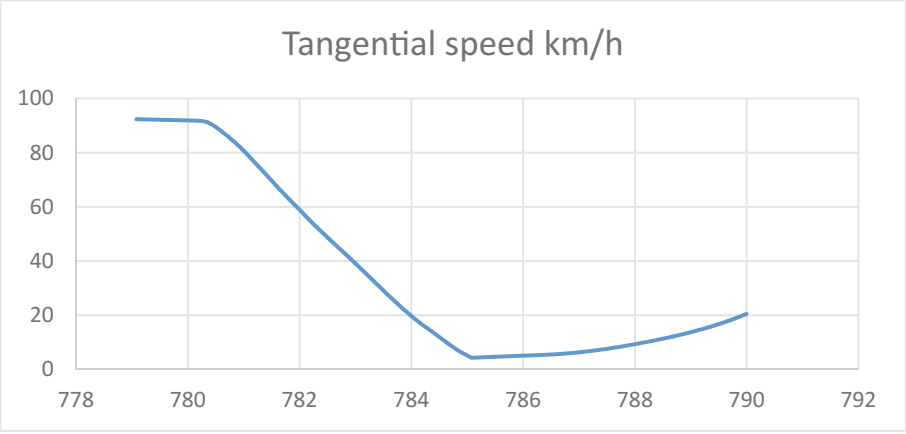


Fig. 6. Tangential speed (km/h) under human control. Notes: (1) Horizontal axis is time stamp (sec). (2) Vertical axis is tangential speed (km/h). (3) Lead car brakes at time stamp 779.2 (sec).

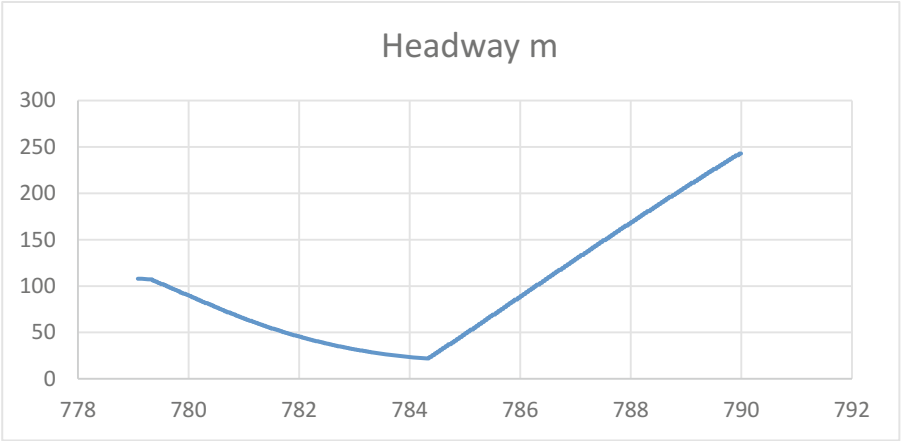


Fig. 7. Distance headway (m) under human control. Notes: (1) Horizontal axis is time stamp (sec). (2) Vertical axis is distance headway (m). (3) Lead car brakes at time stamp 779.2 (sec).

If vehicles are equipped with driving assistance or if vehicles are operating in the automated mode, the following driver actions can be avoided: (1) Distracted and/or aggressive driving. (2) Abrupt and risky actions such as a lane change manoeuver without sufficient gap in traffic. (3) Become distracted first and then to avoid a collision decelerate at a very high rate, (4) In extreme conditions, rear or lateral collision.

An important inference can be drawn from the above analysis that automation in driving supported by accuracy and reliability attributes can play an important role in maintaining a safe, efficient and productive driving environment for individual automated vehicles as well as platoons.

6 Conclusions

(1) Vigorous promotion of autonomous vehicles without checks on their accuracy and reliability may not achieve the safety and efficiency objectives.
(2) The cognitive vehicle design, if guided by human factors and supported by Bayesian AI, is likely to enhance driver acceptance and consequently safety benefits will be achieved.
(3) The cognitive vehicle design has potential to meet accuracy and reliability criteria that are instrumental to gain driver trust.

Acknowledgements. This paper is based on research sponsored by the Natural Sciences and Engineering Research Counsel of Canada (NSERC) and the Ministry of Transportation of Ontario (MTO). The views expressed are those of the author.

References

1. Khan, A.M., Bacchus, A., Erwin, S.: Policy challenges of automation in driving. IATSS Res. **35**(2), 79–89 (2012)
2. Anderson, J.M., Karla, N., Stanly, K.D., Sorensen, P., Samaras, C., Oluwatola, O.A.: Autonomous Vehicle Technology: A Guide for Policymakers. Rand Corporation, Santa Monica (2014)
3. Society of Automotive Engineers (SAE) International: Automated Driving, Levels of Driving Automation are defined in New SAE International Standard J3016, Warrendale, Pennsylvania and Tray, Michigan, U.S.A. (2014)
4. European Commission: EU accelerates connected vehicles. ITS International, 25 January 2016. And C-ITS Platform, Final Report, January 2016. Also, see European Commission: Connected cars. Business Innovations, Internet of Things, Case study 43, February 2015
5. Khan, A.M, Whelen, M., Elsafdi, O., Snobar, N., Jones, B., Arnold, P.: Automation in driving for enhancing resiliency in transportation system. In: CSCE Annual Conference, London, Ontario (2016)
6. Richard, C.M., Morgan, J.F., Bacon, L.P., Graving, J.S., Divekar, G., and Lichty, M.G.: Multiple Sources of Safety Information from V2V and V2I: Redundancy, Decision Making, and Trust – Safety Message Design Report, FHWA – HRT-15-007 (2015)
7. Khan, A., Bacchus, A., Erwin, S.: Surrogate safety measures as aid to driver assistance system design of the cognitive vehicle. IET Intell. Transp. J. **8**(4), 415–424 (2014)
8. Heide, A., Henning, K.: The "cognitive car": a roadmap for research issues in the automotive sector. Ann. Rev. Control **30**, 197–203 (2006)
9. Hoch,S., Schweigert, M., Althoff, F. The BMW SURF project: a contribution to the research on cognitive vehicles. In: Proceedings of the 2007 IEEE Intelligent Vehicle Symposium, Istanbul, Turkey, 13–15 June 2007, ThB1.26, pp. 692–697 (2007)
10. Traffic Technology Today: News: Autonomous vehicle software developer secures funding for 'human-like' driving, 5 February 2016
11. Khan, A.M.: Bayesian-Monte Carlo model for collision avoidance system design of cognitive vehicle. Int. J. Intell. Transp. Syst. Res. **11**(1), 23–33 (2013)
12. Korb, K.B., Nicholson, A.E.: Bayesian Artificial Intelligence. Chapman & Hall/CRC, Boca Raton (2004)

Connecting Rural Road Design to Automated Vehicles: The Concept of Safe Speed to Overcome Human Errors

Pasquale Colonna(✉), Paolo Intini(✉), Nicola Berloco, and Vittorio Ranieri

Department of Civil, Environmental Building Engineering and Chemistry, Technical University of Bari, via Orabona 4, 70125 Bari, Italy
{pasquale.colonna,paolo.intini}@poliba.it

Abstract. The human-road-vehicle automated system is a challenge to overcome human errors. Several rural road crashes happen due to loss of friction, unlikely predictable by drivers. The friction diagram method (FDM) by the authors, described in previous papers, is able to evaluate the skidding risk taking into account vehicle, environmental, road factors. An important variable is the vehicle speed. According to the FDM, the speed corresponding to the maximum friction used can be computed. If all vehicles will travel at speeds lower than that, all other safety checks being verified, then the skidding risk can be reduced. Automated vehicles could travel at the safe speed predicted for each section, by acquiring all the necessary information directly from the road. The algorithm can be customized according to the particular vehicle, tires and road conditions. Additional remarks about the shift from traditional road design practice to the driving automation are also given.

Keywords: Automated vehicles · Road friction · Human errors · Safe speed · Adaptive Cruise Control

1 Introduction

The introduction of the automated vehicles (AVs) in the market could have some potential impacts on the road design process. It is probably time to think to which road design rules will still be considered and/or modified, according to the different levels of automation and to the different time goals (short, medium and long-term).

For example:

- Should the approach to the calculation of stopping and overtaking sight distances be maintained, modified or reconfigured?
- Should the deceleration ratio considered for the design speed profile, in approach at specific points, be maintained, modified or reconfigured?
- Could the road friction be automatically related to the road design process?
- Could the lane and shoulder widths, which can be influential on the speed choice of traditional vehicles but almost likely null for AVs, be reduced?

The different levels of automation (see Table 1 in Sect. 2 for more details) evidently indicate the probable time sequence to be expected for the introduction of AVs in the

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_56

road networks. The sequence will likely be the following: (1) road hosting mixed traffic of AVs and traditional vehicles, (2) road network with lanes reserved to AVs, (3) road networks totally shaped for AVs.

The questions highlighted before should be progressively solved in those different stages of configuration of road networks, for what concerns the road design process. It is anyway evident, that these are not only problems regarding road design but also traffic management and operation. Currently, one should base the judgements on what can be observed and forecast at the moment for the future. Anyway, some of these matters will probably radically change in next few years. For example, the introduction of little vehicles lifted from the ground (e.g. aerotaxi, motodrones), governed by public/private network systems, show some potential to change traffic operation rules.

This article is evidently not oriented towards the possible assessment and solution of all these possible matters. Anyway, it aims at contributing to reflect, number and address some of them, especially from the point of view of road design and safety. In fact, these matters are very often considered only by IT, mechanics and electronics perspectives. However, the topic is of great interest for the future of road design too. Therefore, there is a need for starting to think to these matters from the road perspective. This article is a contribution in this sense. In particular, it tries to propose some applications of road design traditional concepts such as safe design speeds and road friction, to the AVs at different levels.

The remainder of the article is structured as follows. In Sect. 2: AVs and Road Design/Safety, a general background about the AVs and their potential impact on road design and safety is presented. The Sect. 3: The Safe Speed, is devoted to the explanation of how a safe speed can be currently determined on a road layout both from a traditional perspective and from a possible suggested perspective related to AVs. Finally, some conclusions are drawn in Sect. 4.

2 AVs and Road Design/Safety

In this section, a literature overview about autonomous vehicles (AVs) with particular regard to road safety is presented. The possible relationships between tire-road friction modelling and self-driving cars are explored, considering previous applications.

2.1 Driving Automation in the Future

The road towards the complete vehicle automation is characterized by subsequent steps. The National Highway Traffic Safety Administration (NHTSA) [1] classifies the degree of vehicles automation into 5 levels, based on the classification made by SAE [2].

However, the success in the short/long-term period of the AVs and their relative share in the traffic flow depend on several features, related to their real future costs and benefits [3]. The cited author, based on the evolution in time of the implementation of previous vehicle technologies, forecast that the AVs could be available for driving on public roads around 2020 and that they could represent most of the circulating vehicles only in 2050 s, depending not only on technical issues.

Table 1. Levels of automation of vehicles [1] and examples of application for each level (mainly based on [4]).

Level	Description	Examples of application
0	No automation	Braking systems, electronic stability control, anti-lock brake systems, dynamic brake support
1	Automation provides limited assistance in time and tasks	Cruise control (CC), Adaptive CC (ACC), Cooperative ACC (C-ACC), Traffic jam assist
2	Automation conducts some tasks but human keeps monitoring the environment	Super Cruise (General Motors), Highway Driving Assistant (Toyota), systems classified as still requiring driving supervision
3	Automation conducts some tasks and monitors the environment in some instances	–
4	Fully automation in some environments and conditions	Parking system (Audi), requiring no supervision by drivers, who can leave cars at garage entrance
5	Fully automation	Robotic Taxi or closed circuit shuttles/ delivery (see e.g. CityMobil2 [5] or Google Projects [6])

In fact, on one hand, there are a lot of technical matters to address, such as the development of safe self-driving vehicles, the implementation of sensors/wireless for vehicle-infrastructure connections (see e.g. [7]). On the other hand, apart from real benefit/cost analyses, the acceptance of AVs by the public is crucial for their diffusion. For example, based on surveys conducted in La Rochelle (France), Piao et al. [8] reported that only a quarter of the respondents think that AVs are safer than traditional vehicles and that only half of them stated that they will consider to use AVs in the future if available. Other possible developments in the field of transport automation can also suffer from scarce acceptance or misperceptions by the public. Vehicle platooning was proposed as a possible strategy for reducing transport, environmental and safety costs (see e.g. the EC SARTRE [9] or the California PATH projects [10]). Within the objectives of the ITS project, Tsugawa [11] estimated that an autonomous platoon of 3–4 trucks traveling on a dedicated lane, with a driver only in the first truck, could pursue energy saving goals through the use of both adaptive cruise control and vehicle-to-vehicle communication systems (Cooperative Adaptive Cruise Control, CACC [12]). However, he stated that freight industry could be also not prone to accept truck platooning because of the time loss due to waiting that all the trucks in the platoon are loaded before departing.

Therefore, the future diffusion of AVs depends on several features. In any case, it will probably modify the relationships between transport and territory. Zakharenko [13] e.g. estimates that an increased availability of AVs can have impacts on travel distances, parking, economic activities, rent prices and urban sprawl, modifying cities.

2.2 The Influence on Road Safety

It is widely accepted that a full driving automation will cause a reduction in road accidents (see e.g. [14]). In fact, human factors are related to about the 90% of road crashes (see e.g. [15]). The logical expectation is that, by reducing the possibility of human errors through automation, then a huge amount of traffic crashes could be reduced. This could be especially valid for a situation in which self-driving vehicles are largely developed rather than in a condition of mixed vehicles. In fact, Litman [3] estimates that the safety benefits related to AVs in terms of reduced traffic risk could reveal their impact between 2040s and 2060s.

AVs will benefit from scanning systems, able to recognize any obstacle, including other vehicles. They will be connected with other vehicles (V2V: Vehicle-to-Vehicle communication) and with the infrastructure (V2I: Vehicle-to-Infrastructure, I2V: Infrastructure-to-Vehicle communications). They will be provided by safety systems like adaptive cruise control, lane guidance and braking systems. This equipment will prevent human errors in the complete self-driving stage. Some of these systems can be currently found in commercial vehicles, consisting in support to the driver [advanced driver assistance system (ADAS)]. However, it was shown that in the first levels of automation, in which drivers still play an active role, these automated processes can induce habituation, reducing response times, as in the case of adaptive cruise control (ACC) [16].

In any case, the expected safety benefits from AVs could be considered as enormous with respect to the current situation. However, while thinking about traffic simulation of self-driving vehicles, Gora and Rüb [17], highlighted some novel arising problems. The connection between vehicles could suffer from distorted or missing data and from handling a huge quantity of data. Vehicles could be exposed to hacking attacks or software breakdowns. Therefore, some security issues will come out, replacing traditional safety matters, but representing challenges to overcome.

2.3 Different Scenarios for Road Design and Operation

The generic condition of traditional vehicles fully governed by drivers is therefore related to more problems from a road safety perspective. In fact, the driver is not aided by guidance systems different from his/her senses. Hence, a road well designed for the traditional vehicle could be adequate also for the AV. In order to obtain this condition, the AV should recognize the eventual driving risks in a minor or equal time period than the traditional user. Thus, the question to be considered is the following. Can the road geometry, the road signs, the GPS signal, the inputs from the on-board systems to recognize the road geometry, the other inputs from on-board/on-road sensors, the information about other vehicles, fixed and mobile obstacles, be acquired by the AV with at least the same effectiveness of a traditional driver? In case of positive answer, the traditional road design could be adequate for the AVs too.

However, an opposite matter arises in parallel. The AV should necessarily have characteristics allowing at least the same efficiency of the traditional vehicle. Hence, the problem is shifted towards the acceptance of some defined quality standards for the AVs. However, the AV could present a lack of efficiency (DE) in each moment of its life.

Thus, the problem of road design compatible with AVs could become the assessment of that quantity DE. How much small should be this quantity in order to let the traditional road design process be adequate for the AVs too? In other words, how much prudent should be AVs with respect to the traditional vehicles, in order to maintain the current road design standards as they are?

Another different scenario is represented by a road (or lane) specifically born for the AV, a sort of "AV-Native Road". In this case, different rules and control systems could be used. In case of AV-Native Road, a centralized control system could be thought, in order to obtain the optimization of capacity and travel times. In this case, vehicle platooning could be used, and the need for guaranteeing stopping and sight distances for vehicles could be eliminated. Moreover, the centralized control system should guarantee the absence of any obstacle on the road layout. The greater risk to address in this case should be the vehicle stability related to the road friction.

The introduction of an AVs-reserved lane, close to a traditional lane on an existing road, could eventually get the matter perplexed. In fact, in this case, the two above presented scenarios could be considered together. Hence, the difficult problem of managing two different control systems, with different rules and design criteria arises. The harmonization of the two concepts should be solved. A possible solution could be that the design of the reserved lane should follow the concepts of traditional design, by differentiating only the lane occupation type. However, in this way, some advantages of automation concerning the overcoming of human errors, leading to the crash reduction, could be lost. For this reason, future road designers should address the following issues:

- The definition of the characteristics of AVs-reserved lanes;
- The definition of the control systems for reserved lanes, interactive with the other parts of the road.

Finally, a reflection should be made about the change of the role of users in the driving task. In a fully autonomous system, centralized and less flexible, the vehicle could essentially become again only a mean of transport. The driving task could potentially lose the sense of being defined as a "task" requiring attention, in particular for familiar drivers (see [18–20]). In fact, the driver will become only a passenger, as in trains, buses or planes. Moreover, the vehicle representing a status symbol, having characterized all the past century, will probably decrease in importance with the transport automation. This could open social and psychological changes, hardly foreseeable.

2.4 AVs and Tire-Road Friction Estimation

As anticipated, the aim of the authors as road safety researchers, is to integrate different knowledge from both road/traffic and vehicle/automation engineering, considering the impact of road geometry on the in-vehicle systems to be implemented in AVs.

Several road crashes on rural roads are currently caused by loss of friction. Normally, drivers are not aware that a skidding will occur until it starts. Therefore, vehicle automation shows a great potential for reducing this accident type: intelligent AVs will be able to forecast the skidding before it happens, based on friction estimations.

There are two perspectives while looking at road friction: a micro-perspective, considering the physics of the tire-road interaction and a macro-perspective, considering the influence of road friction as a parameter on the road design practice. In fact, it plays a crucial role in the design of horizontal curves, with particular regard to the cross slopes and the radii of curvature. The different road geometric conditions (in terms of horizontal and vertical alignments, together with environmental and vehicle factors) determine the friction used by the vehicle with respect to the available tire-road friction.

Currently, some technologies are able to estimate the tire-road friction coefficient in real-time. They can use torque measurements but also simple GPS-based and accelerometers measurements, considering different models, surface and vehicle factors [21–26]. Real-time tire-road friction estimation can be used for controlling the vehicle dynamics such as the ABS, the Traction Control Systems (TCS) or the Active Yaw Control (AYW) [26, 27]. Chen and Wang [28] or Wang et al. [29] highlighted that these applications can be easier on electric ground vehicles, allowing many advantages due to more flexibility and control for four-wheel-driving and four-wheel-steering vehicles (4WD4WS).

All these technologies consider the application of micro estimates of the tire-road friction coefficient to the vehicle control systems. However, a macro perspective can be considered too, thanks to the I2V technology. If roads and vehicles can communicate, then it could be easy for vehicles to acquire geometric information from the road. Tire-road friction estimations based on existing models, together with macro remarks based on road design could be integrated in advanced ACC systems, including the road friction. This concept is explained in the remainder of the paper.

3 The Safe Speed

One of the main aims of the road designer is to individuate, for each section, the safe speed to be followed by the generic user. Moreover, the road should be designed in order to guide the user in the speed choice, coherently with the safe speed selected by the designer (concept of self-explaining roads, see [30]).

The proposal for a definition of the safe speed for the traditional road design is given in Sub-sect. 3.2. It is also based on a method previously developed by the authors regarding the prediction of skidding risk, which is presented below in Sub-sect. 3.1. A proposal of application of these concepts to the AVs is suggested in Sub-sect. 3.3.

3.1 The Friction Diagram Method

Within the possible safety checks to consider in the road design/enhancement stages, the authors developed a novel method [31], namely the Friction Diagram Method, in order to take into account the road friction in road safety. According to different environmental, vehicle and road conditions, this method provides a diagram, namely the Friction Diagram. In each section of the road, it compares the Friction Demand (F_D - the friction needed by the vehicle for not skidding) with the available Friction Potential (F_P) that the road can provide to the vehicle in some defined conditions. The comparison

is assessed through the introduction of the concept of Friction Used (F_{USED}), the ratio between the Friction Demand and the Friction Potential:

$$F_{USED} = \frac{F_D}{F_P} \times 100\ [\%] \tag{1}$$

When the F_D exceeds the F_P, the F_{USED} is greater than 100%. In this case, the conditions of safe driving are not ensured with respect to skidding.

An example of Friction Diagram, which is the representation of the F_{USED} along the road, is given in Fig. 1, for a two-lane rural road in the hinterland of Bari (Italy).

Through the determination of the F_{USED}, the diagram provides the exact measure of the safety conditions related to friction to the safety inspector. Thus, for each section, and for each combination of the horizontal and vertical alignments, the friction diagram provides the road risk assessment with respect to friction. The analysis can be conducted through different vehicle types. However, for each section, it is useful to define the vehicle showing the worst performance in terms of skidding risk (named by the authors: Design Critical Vehicle, DCV, updated in [32]).

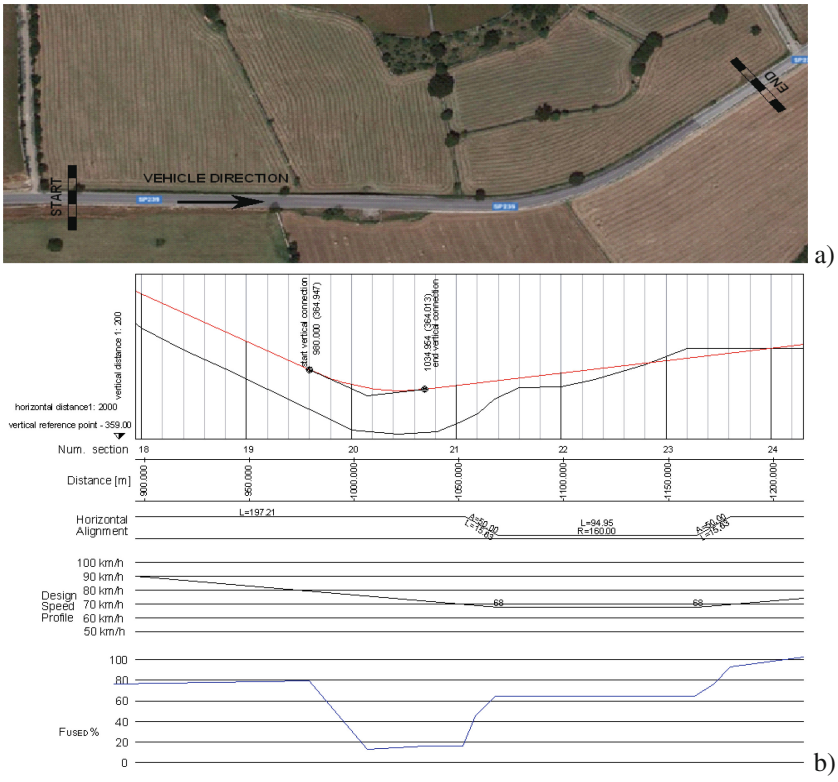


Fig. 1. (a) a section of a two-way two-lane rural road (S.P. 239, province of Bari, Italy); and (b) its friction diagram [31].

Taking the 3rd Lamm's safety criterion [33] as a reference, the construction of the Friction Diagram can be considered as its update, but based on new hypotheses. They allow the analysis of friction taking into account several different features related to the vehicle, the road design, the environmental conditions, and to the vehicle dynamics.

The Friction Diagram Method can be implemented in both the design and safety-based maintenance (enhancement) stages, as a further tool to improve safety of new and existing roads [34, 35]. It can lead to the determination of the safe speed to be guaranteed, for each section of a road layout, through the traditional design process (Sect. 3.2). In that case, the proposed method integrates the already existing safety checks. However, the concept of safe speed defined below for the traditional design, could be also applied to the AVs, as discussed in Sub-sect. 3.3.

3.2 The Safe Speed in the Traditional Design

The Friction Diagram shown in Fig. 1, can be used as a design/diagnosis tool for the road design/enhancement stages. The horizontal and vertical road alignments belonging to the traditional design can be modeled in order to ensure that the F_{USED} is always less than 100% for each road section. The diagram should be used together with the speed and the sight distance diagrams, in order to define a safe speed, within the design speed range, for each road section. It is named "safe", because it satisfies the three above mentioned checks, in both directions of travel and for different boundary conditions. The safe speed is intended as the lowest speed among the three speed values which satisfy the three checks about visibility, speed and friction [32]. Conversely, it can be also considered as the maximum speed among all possible speed values, which simultaneously satisfy the three safety checks.

An iterative procedure has to be implemented in order to determine the safe speed. Once the speed range of the road has been identified, the speed diagram can be designed (in both directions of travel) according to given rules defined by standards. The definition of the speed diagram also allows to plot the visibility diagram and the friction diagram (in both directions of travel), because they depend on speed. In the road sections for which at least one of the three checks based on the diagrams is not satisfied, the checks have to be repeated. In this case, the speed should be progressively decreased in order to obtain a value able to satisfy all the checks, namely the safe speed.

In a previous paper, an application was performed [32]. However, that practical example of definition of the safe speed is briefly recalled as follows. In a first stage, none of the three checks were satisfied, considering a maximum design speed equal to 100 km/h. Hence, the same checks were carried out with a lower speed of 70 km/h. In that case, the visibility and the speed checks were satisfied, while the friction check was not met in a small part of the road section and for only one direction of travel. Further calculations have been conducted iteratively in order to obtain the speed satisfying the friction check for all the road section. Its value is equal to 66.67 km/h (Fig. 2). This speed is the safe speed for the road inquired. It should eliminate the skidding risk on the inquired road section, for any vehicle in wet conditions.

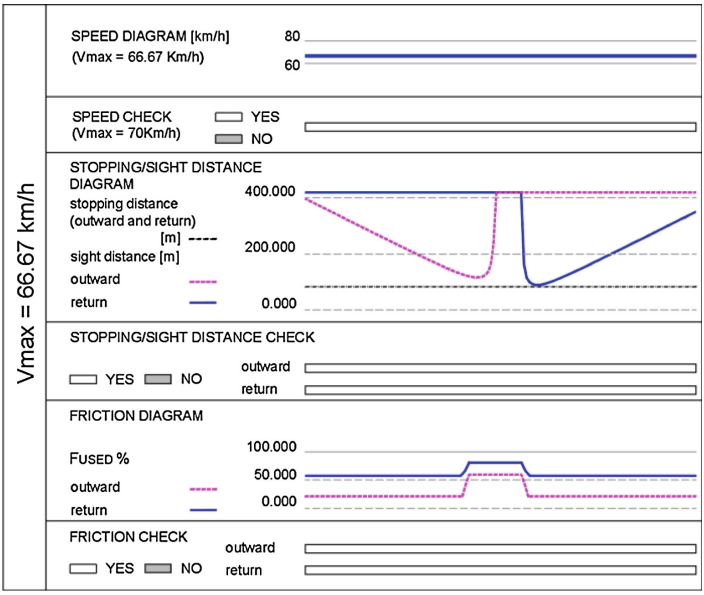


Fig. 2. Example of definition of the safe speed for a rural road segment in traditional design. Adapted from [32].

The study of the entire road infrastructure potentially leads to various safe speeds for each road section. A traditional countermeasure to be implemented on the road could be the selection of speed limits defined through this iterative procedure.

3.3 The Safe Speed for the Autonomous Vehicle

At this point, the safe speed for an AV, in the different defined levels of automation should be assessed. Should it be greater of the speed of traditional vehicles, in order allow a greater efficiency of AVs? Or, considering some unpredictable interactions between AVs and the environment, should it be lower?

By using the friction diagram, the safe speed is determined for each combination of the horizontal and vertical alignment, each vehicle and environmental condition, according to the friction available and required.

In case of roads traveled by traditional vehicles, the safe speed is obtained by comparing the checks regarding friction, speed and sight distance and choosing the minimum speed within the three speed values allowing the verification of the three checks (see Sect. 3.2). For each section of the road layout, it guarantees the safe travel.

The safe speed so obtained could be implemented in the AVs. In this way, they could travel in safe conditions for each road layout, being the vehicle parameters automatically available through on-board systems. However, each vehicle should be able to automatically recognize the road layout (from a geometry perspective) in order to allow the calculation of friction limit conditions in each possible combination. In fact, in the future, vehicles could have a full road geometry database implemented on-board. Alternatively,

or jointly, roads should be supplied with sensors in order to be read by the vehicles. This is already considered, since vehicles should be provided with guidance systems in order to choose the trajectory to follow.

However, the safe speed determined for the traditional design and vehicles is assessed through speed, sight distance and friction checks. Its application to the AVs should consider that the above defined checks will be most likely modified in advanced stages of automation. In fact, the concepts of design speed and the verifications about acceleration/deceleration of subsequent road elements are thought for human drivers. Therefore, the function of a speed check will be probably changed for AV-Native Road design, however still considering some comfort requirements due to accelerations/decelerations. Similarly, sight distances checks will be most likely modified due to the fact that the AV will detect obstacles through sensors and not through a human eye. Hence, it is evident that, even if the concept of safe speed can be considered as applicable to the AVs, there are some challenges to overcome due to the different developments of the possible road scenarios (see in Sect. 2.3).

In the intermediate stages of automation, in which users can still play an active role in the driving task, the concept of safe speed can still be considered. In fact, an aided guidance system could be implement on-board, suggesting to the driver, section by section, the safe speed to follow on the road layout, according to its traditional determination (see in Sect. 3.2). It can be considered as an Adaptive Cruise Control (ACC) system, to be implemented in-vehicle.

The application of these concepts could be potentially valid for all the road environments. However, it is particularly devoted to the rural environment, for which the determination of an automated safe speed can be more demanding, especially in absence of platooning technologies (probably devoted in a first stage only to some specific main arterial roads). In fact, in urban environments, which are most often related to the introduction of AVs, the setting of a safe speed is less applicable. In this case indeed, the speed should be normally low and it is not referred to a vehicle in free flow conditions, since the normal condition is the continuous interactions with other vehicles, pedestrians and/or obstacles. Hence, in this specific application related to road friction, the rural road environment was particularly considered.

4 Conclusions

Some concepts regarding the introduction of the automated vehicles (AVs) and their possible impacts on road design, safety and operation were presented. Some possible scenarios of development about the road networks were discussed and related to the levels of vehicle automation currently defined. Several features and problems related to the evolution of road design with respect to the AVs need to be further addressed and separately studied. In particular, from the perspective of road design, the role of traditional safety checks will be likely modified. However, their evolution is dependent from the parallel introduction of the AVs and their impact on road operation. Land use and transport planning could be affected as well (see e.g. [36]), since the introduction of AVs will have an impact on transport planning.

The concept of safe speed was explained, starting from the traditional design context which includes speed, sight distance and friction safety checks. The possible application of this concept was suggested for the AV, particularly in rural environments. It is evident anyway that also the details about the possible implementation of the safe speed prediction in the AVs will be dependent on the evolution of AVs themselves and on the available technologies. However, it could be considered also in the intermediate stages of automation, as an Adaptive Cruise Control (ACC) system, able to guide the driver in the selection of the safe speed for each vehicle in each road section.

References

1. National Highway Traffic Safety Administration: Federal Automated Vehicles Policy. Accelerating the next revolution in roadway safety. DOT HS 812 329. Washington, DC (2016)
2. SAE International. On-Road Automated Vehicle Standards Committee: SAE Information Report: (J3016) Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving System, Warrendale, PA (2014)
3. Litman, T.: Autonomous Vehicle Implementation Predictions. Implications for Transport Planning. Victoria Transport Policy Institute, Victoria, Canada (2017). http://www.vtpi.org/avip.pdf
4. Christensen, A., Cunningham, A., Engelman, J., Green, C., Kawashima, C., Kiger, S., Prokhorov, D., Tellis, L., Wendling, B., Barickman, F.: Key considerations in the development of driving automation systems. In 24th Enhanced Safety Vehicles Conference. Gothenburg, Sweden (2015)
5. CityMobil2 Project. Cities demonstrating automated road passenger transport. http://www.citymobil2.eu/
6. Google self-driving car project (2009), currently by Waymo, Alphabet Inc. (2016) https://waymo.com/
7. Lu, N., Cheng, N., Zhang, N., Shen, X., Mark, J.W.: Connected vehicles: Solutions and challenges. IEEE Internet Things J. **1**(4), 289–299 (2014)
8. Piao, J., McDonald, M., Hounsell, N., Graindorge, M., Graindorge, T., Malhene, N.: Public views towards implementation of automated vehicles in urban areas. Transp. Res. Procedia **14**, 2168–2177 (2016)
9. SARTRE Project. Safe road trains for the environment. http://www.sartre-project.eu/
10. PATH Project. Partners for advanced transportation technology. http://www.path.berkeley.edu/
11. Tsugawa, S.: Results and issues of an automated truck platoon within the energy ITS project. In: Intelligent Vehicles Symposium Proceedings, 2014 IEEE, pp. 642–647 (2014)
12. Nowakowski, C., Shladover, S., Cody, D., Fanping, B., O'Connell, J., Spring, J., Dickey, S., Nelson, D.: Cooperative adaptive cruise control: testing drivers' choices of following distances (California PATH Research Report No. UCB-ITS-PRR-2010-39), Berkeley, CA (2010)
13. Zakharenko, R.: Self-driving cars will change cities. Reg. Sci. Urban. Econ. **61**, 26–37 (2016)
14. Fagnant, D.J., Kockelman, K.: Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transp. Res. Part A **77**, 167–181 (2015)
15. Treat, J.R., Tumbas, N.S., McDonald, S.T., Shinar, D., Hume, R.D. Tri-level study of the causes of traffic accidents. Executive Summary (No. DOTHS034353579TAC Final Rpt). (1979)
16. Rudin-Brown, C.M., Parker, H.A.: Behavioural adaptation to adaptive cruise control (ACC): implications for preventive strategies. Transp. Res. Part F **7**(2), 59–76 (2004)
17. Gora, P., Rüb, I.: Traffic models for self-driving connected cars. Transp. Res. Procedia **14**, 2207–2216 (2016)

18. Colonna, P., Intini, P., Berloco, N., Ranieri, V.: The influence of memory on driving behaviour: how route familiarity is related to speed choice. an on-road study. Saf. science, **82**, 456–468 (2016)
19. Intini, P., Colonna, P., Berloco, N., Ranieri, V.: Measuring trade-offs between risk and travel time Based on experimental speed data. In: Advances in Human Aspects of Transportation, pp. 1103–1116. Springer International Publishing, New York, (2016)
20. Intini, P., Colonna, P., Ryeng, E., Berloco, N., Ranieri, V.: The relationships between familiarity and road accidents: some case studies. In: Transport Infrastructure and Systems: Proceedings of the AIIT International Congress on Transport Infrastructure and Systems (Rome, Italy, 10–12 April 2017). CRC Press (2017, in press)
21. Hahn, J.O., Rajamani, R., Alexander, L.: GPS-based real-time identification of tire-road friction coefficient. IEEE Trans. Control Syst. Technol. **10**(3), 331–343 (2002)
22. Li, L., Wang, F.Y., Zhou, Q.: Integrated longitudinal and lateral tire/road friction modeling and monitoring for vehicle motion control. IEEE Trans. Intell. Transp. Syst. **7**(1), 1–19 (2006)
23. Grip, H.F., Imsland, L., Johansen, T.A., Fossen, T.I., Kalkkuhl, J.C., Suissa, A.: Nonlinear vehicle side-slip estimation with friction adaptation. Automatica **44**(3), 611–622 (2008)
24. Rajamani, R., Phanomchoeng, G., Piyabongkarn, D., Lew, J.Y.: Algorithms for real-time estimation of individual wheel tire-road friction coefficients. IEEE/ASME Trans. Mechatron. **17**(6), 1183–1195 (2012)
25. Choi, M., Oh, J.J., Choi, S.B.: Linearized recursive least squares methods for real-time identification of tire–road friction coefficient. IEEE Trans. Veh. Technol. **62**(7), 2906–2918 (2013)
26. Li, L., Yang, K., Jia, G., Ran, X., Song, J., Han, Z.Q.: Comprehensive tire–road friction coefficient estimation based on signal fusion method under complex maneuvering operations. Mech. Syst. Signal Process. **56**, 259–276 (2015)
27. Singh, K.B., Taheri, S.: Estimation of tire–road friction coefficient and its application in chassis control systems. Syst. Sci. Control Eng. **3**(1), 39–61 (2015)
28. Chen, Y., Wang, J.: Adaptive vehicle speed control with input injections for longitudinal motion independent road frictional condition estimation. IEEE Trans. Veh. Technol. **60**(3), 839–848 (2011)
29. Wang, X., Zhao, J.: Switched adaptive tracking control of robot manipulators with friction and changing loads. Int. J. Syst. Sci. **46**(6), 955–965 (2015)
30. Theeuwes, J., Godthelp, H.: Self-explaining roads. Safety science **19**(2), 217–225 (1995)
31. Colonna, P., Berloco, N., Intini, P., Perruccio, A., Ranieri, V.: Evaluating the skidding risk of a road layout for all types of vehicles. J. Transp. Res. Rec. **2591**. 94–102 (2016)
32. Colonna, P., Berloco, N., Intini, P., Ranieri, V.: The method of the friction diagram: new developments and possible applications. In: Transport Infrastructure and Systems: Proceedings of the AIIT International Congress on Transport Infrastructure and Systems (Rome, Italy, 10–12 April 2017). CRC Press (2017, In press)
33. Lamm, R., Psarianos, B., Mailaender, T.: Highway Design and Traffic Safety Engineering Handbook. McGraw-Hill Inc. (1999)
34. Colonna, P., Berloco, N., Intini, P., Ranieri, V.: A scientific approach searching for a protocol to design and improve safety of existing roads. Paper presented at the 1st European Road Infrastructure Congress, 18–20/10/2016, Leeds (UK) (2016)
35. Colonna, P., Berloco, N., Intini, P., Ranieri, V.: Sicurezza stradale. Un approccio scientifico ad un problema tecnico e comportamentale. WIP Edizioni (2016)
36. Colonna, P., Berloco, N., Circella, G.: The interaction between land use and transport planning: a methodological issue. Procedia: soc. Behav. Sci. **53**, 84–95 (2012)

A Longitudinal Simulator Study to Explore Drivers' Behaviour During Highly-Automated Driving

David R. Large[1(✉)], Gary Burnett[1], Andrew Morris[2], Arun Muthumani[2], and Rebecca Matthias[3]

[1] Human Factors Research Group, University of Nottingham, Nottingham, UK
{david.r.large,gary.burnett}@nottingham.ac.uk
[2] Loughborough Design School, Loughborough University, Loughborough, UK
A.P.Morris@lboro.ac.uk, marunpsg@gmail.com
[3] Jaguar Land Rover Research, International Digital Laboratory, University of Warwick, Coventry, UK
rmatthia@jaguarlandrover.com

Abstract. Six experienced drivers each undertook five 30-min journeys (portrayed as 'daily commutes' i.e. one on each of five consecutive weekdays) in a medium-fidelity driving-simulator engineered to mimic a highly-automated vehicle. Participants were encouraged to act as they might in such a vehicle by bringing with them their own objects/devices to use. During periods of automation, participants were quickly engrossed by their chosen activities, many of which had strong visual, manual and cognitive elements, and required postural adaptation (e.g. moving/reclining the driver's seat); the steering wheel was typically used to support objects/devices. Consistently high subjective ratings of trust suggest that drivers were unperturbed by the novelty of highly-automated driving and generally willing to allow the vehicle to assume control; ratings of situational awareness varied considerably indicating mixed opinions. Qualitative results are discussed in the context of the re-design of vehicles to enable safe and comfortable engagement with secondary activities during high-automation.

Keywords: Highly-automated driving · Driving simulator · Trust · Situation awareness · Secondary tasks · Transfer of control

1 Introduction

Fully autonomous cars are expected to deliver plentiful socioeconomic benefits, including improvements to traffic flow and road safety, and 'mobility for all' [1]. Relinquishing the responsibility for vehicle control also provides the opportunity for drivers to engage in personal or social activities, unrelated to the driving task. This is likely to make time 'spent' driving more enjoyable and productive, and is expected to transform the comfort and convenience of everyday car travel [2]. Given that UK drivers spend, on average, 235 h a year behind the wheel [3], understanding the type and variety of activities that they may undertake within their vehicles during periods of automation – and the likely impact that engagement in such activities will have on drivers' behaviour and continued

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_57

engagement with the primary driving task – is of particular interest. A key consideration (during highly-automated driving) is that drivers should be able to comfortably and effectively undertake secondary activities of their choosing, while the vehicle is in control, but must be able to resume control in situations not covered by the automation. This will naturally impact and inform the design of next-generation vehicles – potentially inspiring radically new vehicle interior designs – and appears key to the acceptance of highly-autonomous cars [4].

1.1 Automation

The driving task is made up of myriad individual subtasks [5]. Automation can be applied to many of these subtasks individually, thereby augmenting (e.g. anti-lock braking systems) or replacing (e.g. adaptive cruise control) elements of traditional manual control. If automation is applied simultaneously across all subtasks, a vehicle may be considered autonomous (or 'driverless'). However, current automated systems may be unable to cope with all possible driving situations and eventualities [6], and oftentimes, drivers are still therefore required to monitor the system and resume manual control.

There are a number of automation taxonomies that define the allocation and demarcation of system functionality between the driver and automated subsystems, e.g. [7, 8]. Whilst these differ in the specific nomenclature, narrative and number of levels of autonomy they profess, there is general consistency between taxonomies in the adoption and utilisation of three key descriptors: *partial*, *high* and *full*.

During 'partial' automation, the human driver retains full responsibility for safe vehicle control and driving practice, and must therefore permanently monitor the automated system and maintain a readiness to resume control at any time. 'High' automation removes the need to permanently monitor the system, but the driver must still be prepared to resume manual control within a predefined time-period in certain circumstances (e.g. in situations not covered by the automation, or following minor technical malfunctions, such as sensor failure). During 'full' automation, safe operation rests solely on the automated vehicle system and the driver is not expected to be available for control at any time during the journey [7].

Exploring and understanding human factors issues associated with vehicle automation has been a keen area of interest since the 1970s [9]. On first inspection, it may appear that removing drivers from the control-feedback loop and relieving them of their duties and responsibility for safe vehicle operation means that human factors are of limited applicability in automated vehicles. However, far from making human factors redundant, the potential of driverless vehicles inspires new areas of investigation [10]. For example, as vehicles operate at increasing levels of autonomy (i.e. drivers relinquish an increasing degree of control to automation), complex driver-vehicle interaction patterns will exist, and therefore understanding issues such as trust and over-reliance, workload, skills degradation and situation awareness [11–13] – particularly during manual control recovery (MCR) [14] – remain firmly rooted on the research agenda.

Additionally, in situations of high or full-automation, drivers – relieved of their duties and responsibilities for manual control – may be inclined to undertake secondary tasks while the vehicle is in control [15]. This is likely to have a direct impact on vehicle

design. For example, automotive manufacturers may recline or relax the driver's seating position, withdraw or collapse the steering wheel, or provide additional interior lighting to enhance comfort, space and visibility for drivers during autonomous control.

1.2 Overview of Methodological Approach

Understanding what drivers are inclined to do during periods of high or full automation – and also how this will impact vehicle interior design and their consequent ability to resume manual control if required – is important. Nevertheless, exploring this before such vehicles have reached wide-scale deployment and acceptance is inherently difficult. Previous investigations concerning drivers' behaviour and their engagement in secondary activities during periods of automated control, e.g. [14, 15], have focussed on exploring changes with increasing level of automation and/or have provided limited exposure/artefacts: some of these approaches rely largely on speculation, and may therefore lack face and ecological validity. Therefore, a new methodological approach is required.

We devised a novel simulator study, requiring that participants attended at the same time on each of five consecutive days (Monday to Friday). During each visit, participants completed the same 30-min motorway journey – portrayed to participants as a regular, daily journey that they might undertake, such as their commute to work – within the driving simulator. The simulator was modified to mimic a highly-automated vehicle, thus relieving drivers of lateral and longitudinal control actions (and all associated cognitive elements). Prior to attending, participants were asked to envisage what activities they might undertake in an autonomous vehicle and to bring with them any objects or devices that they would require to achieve these tasks. During each journey, participants were then able to use these freely while the vehicle was in control.

The qualitative study primarily aimed to highlight the type and range of activities undertaken by drivers (and the artefacts they used to enact these) during periods of high-automation, and identify any behavioural adaptations that these demanded or encouraged. However, it also provided the opportunity to explore the impact that these activities had on primary task engagement, explored through the constructs of situational awareness and trust.

2 Method

Six participants took part in the study, with ages ranging from 29 to 55 years. All participants held a valid UK driving licence and were experienced and active drivers, with more than 5 years driving experience. Participants were self-selecting volunteers who responded to advertisements placed around the University of Nottingham campus and were reimbursed with shopping vouchers as compensation for their time. Due to the risk of simulator-induced motion sickness, participants were screened for the history of motion sickness, migraines, epilepsy, dizziness or blurred vision before taking part in the study, and provided written informed consent.

The study took place in a medium-fidelity, fixed-based driving simulator at the University of Nottingham (Fig. 1). The simulator comprises an Audi TT car located

within a curved screen, providing 270° forward and side image of the driving scene via three overhead HD projectors. Side and rear view mirror images are captured digitally and relayed to two 7-inch LCD screens, located to replicate the side mirrors, and a 55-inch curved LED television positioned behind the vehicle and visible using the existing interior rear-view mirror. A Thrustmaster 500RS force feedback wheel and pedal set are faithfully integrated with the existing Audi steering wheel and pedals, such that drivers interact with the original controls. In addition, the indicators and electrical controls remain fully operational within the vehicle. The driving scenario was created using STISIM Drive (version 3) software to replicate a standard UK motorway, with speed information etc. exported dynamically to a bespoke Java application, and subsequently re-presented on a 7-inch LCD screen within the vehicle to mimic the instrument cluster.

Fig. 1. Driving simulator showing motorway scenario.

Participants were invited to attend at the same time on each of five consecutive weekdays, with each visit representing the same daily commute journey to work. Participants had previously been advised that the car and route supported highly-automated driving, and as such were told that they were not required to permanently monitor the system status or driving situation, but may be requested to resume manual control, given appropriate notice. During each visit, participants began driving manually. Automated control (at 70 mph) was activated shortly into the drive at the request of the participant using a voice-command: "start automated driving"; this initiated a 5.0-s hand-over transition. During the transition, a voice message: "starting automated driving" followed by a punctuated tone (5 beeps), was used to indicate progress and completion of the hand-over.

During periods of automation, participants were provided with a system 'health-bar' presented visually on a 12-inch touchscreen Microsoft Surface Pro tablet computer located in the centre console of the vehicle. Participants were advised that the health-bar indicated the 'overall status' of the automated system (encompassing vehicle sensors, road situation etc.). The 'health' status ranged from *green*, indicating 'system

ok' to *amber* ('minor problem, no driver intervention required') to *red*, indicating 'major problem, driver intervention may be required', with the number of markers also reducing with diminishing health (Fig. 2). The health-bar was designed to correspond with similar 'status' representations in existing technology (e.g. battery discharge indicator, Wi-Fi signal strength etc.) and cultural expectations (i.e. red, amber and green classifications), and intended to provide an 'intuitive' visual representation of vehicle status; as such, participants were not specifically made aware of each individual representation before taking part. In reality, the health-bar indicator remained *green* ('system ok') for most of the drives, with only occasional, short-lived departures into *amber* ('minor problem').

Fig. 2. Automated system 'health-bar' showing (left to right) green, amber and red status, with associated declines in the quantity of 'health' markers.

Drivers were provided with a verbal take-over request towards the end of each drive, comprising a warning message delivered at 60 s prior to hand-over ("approaching take-over") followed by an action alert: "resume manual control", delivered 5.0-s before the provision of manual control. Drivers therefore completed each drive manually. They followed the same route every day with each drive lasting approximately 30 min.

At the end of each drive, participants completed the 'Trust in Automation' [16] and 'Situation Awareness Rating Technique' (SART) questionnaires [17]. Participants were also video-recorded during each drive. These recordings were subsequently analysed and coded to identify common themes, with a coding scheme that emerged during the analysis. In addition, general comments, captured from participants throughout the study were used to elucidate observations.

3 Results and Discussion

Results are presented and discussed in the following sections: *activities and artefacts*, *posture* and *trust and situation awareness.* These sections present some of the salient features of the behaviour that were observed; design implications are discussed throughout. The intention in presenting this work is not to provide an exhaustive account of the entire range of activities that drivers undertook during highly-automated driving, but rather to expose novel, qualitative insights, based on the empirical observation of drivers' behaviour over an extended period of observation, and provide a broad and rich palette for designers of next generation vehicles with 'self-driving' capabilities.

3.1 Activities and Artefacts

A number of distinct activities and artefacts featured during the study. The most common activities were: reading an article or magazine (constituting 22% of analysis time), reading articles or magazine on an iPad (or similar) (16.2%), social networking activities using a mobile device (8.0%), web-browsing (7.4%) and watching programmes/films on a laptop or iPad (5.4% of the time). Drivers also engaged in mobile phone texting/conversations, eating/drinking, taking notes and personal grooming activities, such as applying hand cream/nail polish etc. The most common artefacts used by participants were: paper documents (articles, magazines, books etc.) and computing devices, such as laptops and mobile technology (mobile phones, iPads etc.).

Although there was some commonality between participants, the same activities were not necessarily undertaken to the same degree by all participants. For example, two participants spent more than 80% of their time reading, whereas another participant spent 43% of the time accessing websites and 30% of their time watching films/programmes using a video-on-demand (VOD) provider (notably, Netflix and YouTube). Another participant spent nearly 36% of their time engaged in social networking activities using their mobile phone (i.e. using Facebook, Twitter, etc.).

Many of the activities undertaken by participants involved watching or reading from a secondary device or paper copy. For these activities, the steering wheel provided a natural resting surface (particularly for books, papers, iPad etc.) (Fig. 3), whereas in other situations (e.g. using a laptop) the steering wheel hindered activities and drivers were required to recline or adjust their seating position (or place their laptop elsewhere, e.g. the passenger seat) to account for this.

Fig. 3. Different configurations for holding secondary artefacts employed by drivers.

Utilising the steering wheel as a natural anchor point was a common strategy – this also provided good visibility in that the item was placed in the driver's direct line of sight. For example, in Fig. 3, participants can be seen supporting an iPad against the steering wheel, leaning a magazine article against it and resting a mobile phone on top of it. The participant holding the paper document can actually be seen holding this off

centre, so that it aligned with the health-bar display, thereby ensuring that both items were in the same field of view.

Although using the steering wheel clearly eased the burden of physically holding items, drivers are still likely to physically tire even employing this strategy, particularly on longer journeys. Therefore, a key consideration emerging from the study is whether vehicle manufacturers should provide some form of support for secondary devices during periods of high-automation. This could be achieved by either by withdrawing the steering wheel and replacing it with a different device, or by utilising the steering wheel as a mount (as people appeared naturally inclined to do during the study). An associated concern is that some of these activities (e.g. reading a book or printed article) may require additional lighting. Therefore, ambient in-vehicle lighting that activates during periods of automation (e.g. to accompany the provision of a 'steering-wheel' mount), could also assist drivers.

It is assumed, of course, that if a vehicle interior adapts during periods of automation, it should be capable of rapidly reconfiguring for manual driving – if and when the need arises. In particular, physical changes to re-enable manual control should be made quickly but gracefully, for example to avoid an iPad (hitherto supported by a reclined steering wheel) being forcibly thrust back into a driver's face.

Successfully resuming manual control is also contingent on drivers being able to safely discharge and secure secondary devices. A notable observation during the study was that some drivers found it difficult to store their secondary devices before resuming control. A further problem was revealed by one of our older participants, for whom aged-related presbyopia dictated that they wore reading glasses while attending to their in-vehicle activities. The issue became particularly apparent when they were required to resume manual control of the vehicle – the final hand-over announcement provided sufficient time for this driver to safely locate their laptop, but was apparently insufficient time for them to remove and safely store their glasses. Thus, the participant continued for the remainder of the drive wearing their reading-glasses – in the driving simulator, this did not necessarily present a problem and they were still able to attend to the road situation, but in a real-world environment, this may cause a problem.

Many of the activities also required some degree of 'connectivity' (e.g. streaming content, web-browsing, accessing social networks etc.). Whereas the University wireless network was accessible to participants during the study (largely by default rather than design), and in a real-world environment some technology may have data roaming capabilities, the provision of wireless network access in automated vehicles appears to be paramount and in line with expectations.

Participants generally spent very limited time attending to the 'health-bar'. For most participants, this constituted less than 3% of their time. This suggests that drivers were generally prepared to delegate full control to the vehicle, despite their awareness that it only offered 'high' automation (i.e. they may be required to resume control in some situations). This decision may have been influenced by the simulated environment in which the study took place and a different response may occur in a real world driving situation.

The limited time spent attending to the health-bar in general (i.e. by all participants) also suggests that drivers prioritised their secondary activities above maintaining

awareness of the driving situation. This would obviously impact their ability to resume manual control of the vehicle during an emergency situation. Consequently, it has been suggested that 'hands-free' driving tends to imply 'eyes-free'. To address this concern, some authors (e.g. [14]) suggest that automated systems could disengage automated control if the driver removes their visual attention away from the road centre for longer than a defined period of time. This would encourage drivers to return their gaze to the road to regain situational awareness periodically throughout an automated drive. However, this strategy requires that drivers are aware of the need to do so and also restricts the range of secondary activities available to drivers, e.g. to activities where visual attention can be successfully 'chunked' in this manner. Interestingly, several participants recognised this dilemma during the study. For example, during the post-study interview, one of the participant commented: "*I often used to check the status of health bar when I was involved in activities such as checking email, Facebook* etc.*, but when I watched video episodes on YouTube or Netflix, I hardly noticed the health bar"*.

The nature of secondary activities is clearly of fundamental importance if we are to expect drivers to maintain the ability to resume control, given appropriate notice, in situations of high-automation. While restricting drivers to a limited set of activities might seem feasible 'on paper', the realities of driving mean that this is unlikely to be a viable option in the 'real-world' (consider the current difficulties restricting texting-while-driving in the UK – which is already against the law and carries a significant fiscal and 'points' penalty). A better solution therefore may be to develop other HMI (human-machine-interface) solutions that can keep drivers appraised of the current driving situation (road, traffic, weather etc.), through other means/modalities (e.g. auditory, haptic experiences etc.), without the pre-requisite of visual attention. Nevertheless, it is recognised that drivers may also be highly-engaged in secondary activities from a cognitive perspective, and therefore other solutions employing sound etc. may be equally ineffective. Needless to say, this is a 'hot' topic for further investigation.

3.2 Posture

In situations where the vehicle was in control, most participants (though notably not all) used this opportunity to avail themselves of the 'freedom' from the driving task by changing their posture, reclining their seat etc. In particular, drivers were observed 'relaxing' their posture immediately after relinquishing manual control, e.g. by reclining and/or adjusting the position of their seat. In contrast, drivers were seen overtly re-assuming strict driving posture by sitting upright, holding the wheel in both hands and 'bracing' themselves, as if preparing for impact, shortly prior to resuming manual control (i.e. on hearing the final verbal instruction).

Drivers also changed their posture regularly during periods of automated control. This primarily occurred due to the adoption of new tasks that required different postural configurations (e.g. to provide space for a laptop, to use the internal mirror to apply make-up etc.), but also appeared to happen 'on a whim' throughout the journey. This is therefore likely to be a strategy employed to alleviate discomfort, boredom, fatigue etc. – and is notably something that it is not possible during manual driving. Several examples of the postures adopted by drivers during periods of automated control are shown in Fig. 4.

Fig. 4. Examples of different postures adopted during periods of automated driving.

Postural changes were also observed in response to changes in the driving situation. For example, one of the participants was watching a film on their computer, which was placed on their lap. On noticing changes to the health bar, they repositioned their laptop on the passenger seat, thereby allowing them to both continue watching the film, while also enabling an effective resumption of the physical aspects of manual control should the need arise, i.e. eliminating the need to move the laptop during the transfer-of-control request.

Recognising the different postures that drivers adopt during periods of automated control is important to designers of future vehicles with autonomous capabilities. For example, the transition to automated control may be accompanied by the automatic relaxation of seating position, or the withdrawal of the steering wheel, to enable drivers to adopt a more comfortable posture and provide the space to undertake secondary tasks of their choosing.

3.3 Trust and Situation Awareness

Given that the study was primarily intended to provide a qualitative description of the behaviour and activities of a small number of drivers, trust and situation awareness ratings are provided primarily to illustrate the attitudes and opinions of the six drivers who took part in the study, and are not necessarily generalizable to a wider population.

Trust ratings show that drivers generally placed high levels of trust in the system. Moreover, ratings were generally high from the outset (i.e. day one) and drivers did not require an acclimatisation period, in which to 'test' or gain familiarity with the system. A notable exception to this was that all drivers noticeably delayed removing their hands from the steering wheel after transitioning to automated mode on day one; however, this did not occur on subsequent days.

Trust is an important determinant of system performance and predictor of automation use [18]. However, high levels of trust in automation can have a negative impact on manual control recovery (MCR), particularly in an emergency situation [11]. The high trust ratings observed in the current study (from the outset) is therefore an interesting

observation, which, although potentially confounded by the low-risk perception in the simulated driving environment, could suggest that drivers are likely to adapt very quickly to automated control. On the one hand, this is good news, in that it provides general support for the move to automated vehicles – previous investigations and theory tends to suggest that some drivers may initially be sceptical of relinquishing control to driverless cars. On the other hand, high levels of trust are likely to encourage drivers to quickly delegate full responsibility for vehicle control to the 'system'. In situations of full automation, such behaviour may be acceptable. However, during intermediate solutions (e.g. partial-automation and high-automation), such trust may be misplaced and could result in drivers relying on the systems for aspects of the driving task for which it is not primarily intended.

Using the SART subjective rating technique [17], we also asked participants to provide an indication of their situation awareness (SA) following each drive. During manual control, drivers require good situational awareness for safe control of their vehicle, and integration with the road situation and other road users: this requires an understanding of both the current situation as well as the projection of future actions [6]. Automation can have a negative effect on SA due to over-reliance, reduced vigilance and lack of understanding of the system capabilities [6]. In particular, lapses in situational awareness (typically combined with high levels of trust in the automation) are most likely to present themselves as problematic when resuming manual control.

SART ratings varied considerably throughout the study, suggesting that drivers held mixed opinions about the potential impact of their secondary tasks on situation awareness, and raising potential concerns about their ability to successfully resume manual control if required. Nevertheless, participants' comments indicate that they were well-aware of the need to re-engage with the driving task prior to resuming manual control: "*In automated driving, I feel more relaxed and engaged myself in secondary activities. During the … [hand-over]…, I felt like I need to quickly transfer from my relaxed state (low demand) to an extremely alert state (high demand)".* Another participant, who was seen actively monitoring the system health-bar during periods of automation – and consequently reported high levels of situation awareness – was evidentially well prepared for potential take-over requests: "*I noticed the amber coloured bars started to reach the minimum level … which prompted me to stay more active and expect something might going to happen. Then I placed my hands on the steering wheel to ensure that I could effectively respond to the situation that is going to occur."*

3.4 Further Considerations

Inviting drivers to undertake the same route over an extended five-day period in a driving simulator – to mimic their daily commute driving – is a highly novel approach to study their long-term behaviour during automated driving. Both the methodological approach – as well as the findings themselves – are therefore likely to be of value to the automated driving community. However, there are inevitable limitations associated with simulation that may have influenced participants' behaviour, such as risk perception in a simulated environment. Moreover, some in-car features (that drivers may have reasonably expected in their own vehicle) were notably absent during the study, for example, the radio/CD

player was not enabled. In a real-world situation, listening to music may have been a primary activity. However, we were keen to discover the artefacts that drivers would select (given complete freedom to choose) and did not preclude drivers from using their own music players during the drives.

Furthermore, there were no motion cues in the simulator. In a real-world situation, road vibrations and vehicle motion may limit drivers' ability to engage in tasks demanding high levels of visual attention, such as reading – due to the inevitable mismatch between the visual perception and the vestibular sensation of motion – and will therefore likely result in increased incidence of motion sickness.

Finally, the HMI used to provide system health status was designed for experimental convenience and was necessarily sparse in order to avoid distracting drivers. Again, in a real-world situation, far more detailed information concerning vehicle status, road conditions, status of automation etc. would likely be presented.

4 Conclusion

Long-term, qualitative observations suggest that drivers are likely to engage in a variety of secondary activities using a range of artefacts during periods of high-automation. Many of the activities observed had a strong visual element and required postural adaptation (e.g. moving or reclining the driver's seat to watch a film on a laptop): for some drivers, activities also necessitated the wearing of reading glasses, which had previously not been required during manual driving. The study differs to previous investigations in that drivers were exposed to automated driving over an extended period, involving multiple visits to the simulator to complete the same 'commute' journey. Moreover, participants were instructed to bring with them any items or devices that they wanted to use. Although the study was conducted using a vehicle designed to operate at high-automation, many of the results are equally applicable to situations of full-automation. The findings have clear implications for the design of next generation vehicles with self-driving capabilities, although further investigations are required. In particular, studies should consider the effect of road vibrations and vehicle motion on the selection and engagement with secondary activities during periods of automation. Further investigations are also required to explore the carry-over effects of secondary activities, particularly in emergency take-over situations.

Acknowledgements. The study was independently funded and conducted as part of the postgraduate degree in Ergonomics and Human Factors at Loughborough University, in collaboration with the Human Factors Research Group at the University of Nottingham. The authors would like to thank Rebecca Matthias, Human Factors Specialist at Jaguar Land Rover, who provided an industrial perspective and support throughout the study.

References

1. Mui, C., Carroll, P.B.: Driverless cars. Trillions are up for grabs. Cornerloft Press, Seattle (2013)
2. Fagnant, D.J., Kockelman, K.: Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transp. Res. A Pol. Pract. **77**, 167–181 (2015)
3. Department for Transport: The pathway to driverless cars: summary report and action plan (2015)
4. Poulin, C., Stanton, N.A., Cebon, D., Epple, W.: Response to: autonomous vehicles. Ingenia **62**(9) (2015)
5. McKnight, A.J. Adams, B.B.: Driver Education Task Analysis. Volume II: Task Analysis Methods (1970)
6. Endsley, M.R., Kaber, D.B.: Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics **42**(3), 462–492 (1999)
7. Gasser, T.M. Westoff, D.: BASt-study: definitions of automation and legal issues in Germany. In: Road Vehicle Automation Workshop, Irvine, CA (2012)
8. SAE International: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (2016)
9. Sheridan, T.B.: Big brother as driver: new demands and problems for the man at the wheel. Hum. Factors J. Hum. Factors Ergon. Soc. **12**, 95–101 (1970)
10. Banks, V.A., Stanton, N.A.: Going from a 'Driver Driving' to a 'Driver Monitoring' and back again: Modelling driving automation using EAST. In: Human Factors in Automated Systems (submitted) (2017)
11. Payre, W., Cestac, J., Delhomme, P.: Fully Automated driving: impact of trust and practice on manual control recovery. Hum. Factors **58**(2), 229–241 (2016)
12. Ma, R., Kaber, D.: Situation awareness and workload in driving while using adaptive cruise control and a cell phone. Int. J. Ind. Ergon. **35**(10), 939–953 (2005)
13. Parasuraman, R., Sheridan, T., Wickens, C.: A model for types and levels of human interaction with automation. IEEE Trans. Syst. Man Cybern. A Syst. Hum. **30**(3), 286–297 (2000)
14. Merat, N., Jamson, A.H., Lai, F.C.H., Daly, M., Carsten, O.M.J.: Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. Transp. Res. F Traffic Psychol. Behav. **27**, 274–282 (2014)
15. Jamson, A., Merat, N., Carsten, O., Lai, F.: Behavioural changes in drivers experiencing highly- automated vehicle control in varying traffic conditions. Transp. Res. C Emerg. Technol. **30**, 116–125 (2013)
16. Jian, J., Bisantz, A.M., Drury, C.G.: Foundations for an empirically determined scale of trust in automated systems. Int. J. Cogn. Ergon. **4**(1), 53–71 (2000)
17. Taylor, R.M.: Situational awareness rating technique (SART): the development of a tool for aircrew systems design. In: Situational Awareness in Aerospace Operations, pp. 1–17 (1990)
18. Lee, J., Moray, N.: Trust, self-confidence, and operators' adaptation to automation. Int. J. Hum. Comput. Stud. **40**(1), 153–184 (1994)

Understanding and Applying the Concept of "Driver Availability" in Automated Driving

Claus Marberger[1(✉)], Holger Mielenz[1], Frederik Naujoks[2], Jonas Radlmayr[3], Klaus Bengler[3], and Bernhard Wandtner[4]

[1] Robert Bosch GmbH, Corporate Research,
Robert-Bosch-Campus 1, 71272 Renningen, Germany
{claus.marberger,holger.mielenz}@de.bosch.com
[2] Center for Traffic Sciences (IZVW), University of Wuerzburg,
Roentgenring 10, 97070 Wuerzburg, Germany
naujoks@psychologie.uni-wuerzburg.de
[3] Chair of Ergonomics, Technical University of Munich, Boltzmannstraße 15,
85747 Garching, Germany
{jonas.radlmayr,bengler}@tum.de
[4] Adam Opel AG, Advanced Technology EE, Bahnhofsplatz 1, 65423 Ruesselsheim, Germany
bernhard.wandtner@opel.com

Abstract. Several levels of automated driving functions require the human as a fallback driver in case system performance limits are exceeded. Human factors research in this area is especially concerned with human performance in these take-over situations and the influence of the driver state. Based on work of the publicly funded project Ko-HAF the paper introduces a comprehensive model of the transition process from automated driving to manual driving and specifies relevant time stamps and time windows. The concept of *Driver Availability* is regarded as a quantitative measure that relates the estimated time required to safely take-over manual control to the available time budget. A conceptual framework outlines potential influencing factors on driver availability as well as ways to apply the measure in a real-time application.

Keywords: Automated driving · Control transition · Driver availability · Driver model · Human factors

1 Introduction

"Will I be allowed to sleep? Can I watch a movie? What about gaming on my phone or computer?" Researchers working on human factors of automated driving often hear these or related questions from potential users. It shows the big interest the topic has seen in the media and society. In consequence, researchers are working towards offering answers for an understanding of the psychological effects of automated driving.

The current definitions of automated driving distinguish the range between manual control of the vehicle and fully autonomous systems in 5 [1] or 6 levels [2]. In order to hand back vehicle control to the driver within a sufficient time budget, so-called "Requests to Intervene (RtI)" are issued by the technical system to the driver. This is a

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_58

key characteristic of SAE Level 3 (L3) "conditional automation" (according to [2]) or BASt Level 3 (L3) "high automation" (according to [1]) but the issue may also be relevant for SAE Level 2 and Level 4 automation. This work applies to all automation levels in which RtIs are part of the interaction concept. When L3 automation is in control of the vehicle dynamics, the driver is allowed to engage in non-driving related tasks (NDRT). While past automotive research on task allocation has been focused on secondary/tertiary tasks parallel to the primary driving task, conditional/highly automated driving represents a paradigm change since the parallel task turns into a sequential task paradigm [3].

One of the main motivations of past research efforts was to investigate the maximum time needed to resume manual driving and to better quantify the "sufficient time budget" [4, 5]. Results of various studies (for an overview see [6]) point towards a large spectrum of maximum reaction times from about 2 to 15 s. Now it is one of the most obvious research tasks to explain the variance and to identify the contributing factors. Gold [7] provided first insights into factors like age, traffic density and NDRT modality for a fixed take-over scenario. Nevertheless, he also concludes that further aspects, especially related to further driver states, should be taken into account in order to better explain the variance in take-over times.

Human factors researchers often use the term "driver availability" as the main topic when it comes to automated driving systems with human drivers as fallback [8–10]. However, what does this concept actually mean? Is there a way to operationalize it and how can it be used in automated driving systems?

Within the German publicly funded project Ko-HAF ("cooperative highly automated driving") one of the objectives is to develop a deeper understanding of the concept "driver availability". The concept serves as a framework for a series of human factors experiments with the goal to utilize a "driver availability measure" within a highly automated driving system. Ko-HAF especially addresses take-over performance after long automation periods, thus putting driver availability and it's contributing factors in the center of research. Two of the project goals are to quantify relevant aspects of driver availability with respect to their effect on take-over performance and to design human-machine-interfaces that enable safe and comfortable transitions from highly automated driving to manual control. This paper aims to introduce a conceptual framework for specifying the concept of driver availability with an outlook how to apply it in system design.

2 Take-Over Process Model and Terminologies

Since driver availability focuses on performance aspects in transition situations, we start to take a closer look at the transition process. Based on the work of [4, 5, 7, 11, 12] this paper proposes a comprehensive model (see Fig. 1) to cover transitions initiated by a Request to Intervene (RtI). It is based on two parallel streams of state transitions – one focusing on basic automation states and one addressing aspects of human performance and state. The objective is to precisely define time windows and related start/end conditions within the transition process.

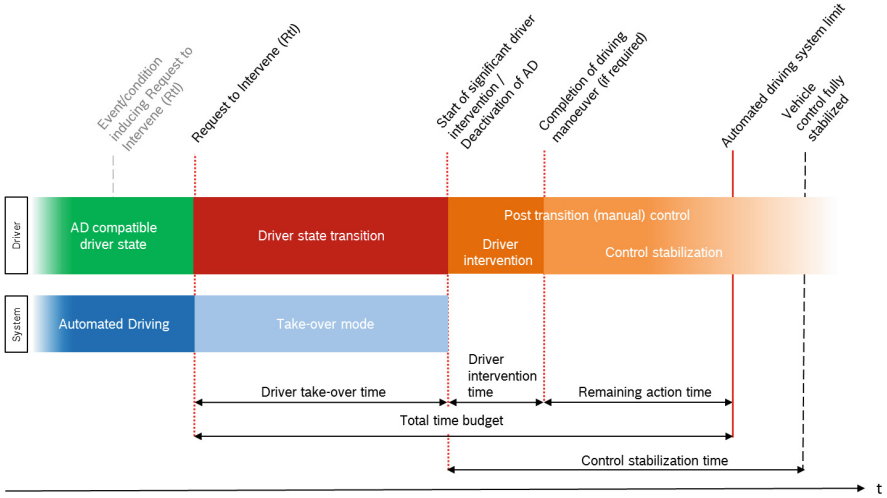


Fig. 1. System-initiated transition model from automated to manual driving

The following sections briefly describe the distinct phases of the transition process (see Fig. 1) and observable events that allow separating them:

AD (Automated Driving) Compatible Driver State: Driver state that is required due to performance limits of the automated driving system. Requirements on the compatible driver state are strongly dependent on the automation level (e.g., "driver in-the-loop" for SAE L1/2, driver take-over after sufficient lead time for SAE L3). Moreover, the required driver state may also be dependent on the current automation capability and the environmental conditions.

Request to Intervene (RtI): A system generated warning message that asks the driver to immediately take-back control after a system limit has been recognized. The RtI time stamp is essential to analyze the subsequent transition process as it represents the beginning of the driver state transition process.

System Limit: Situation that can be specified in time and space which the automated driving system cannot handle automatically and which causes a RtI. Examples are: sensor failure/no available sensor data, accident/critical object ahead. The time window starting with the RtI and ending with the system limit is referred to as "total time budget". It represents the maximum time window within the driver has to finish all required actions. Due to the dynamic nature of the take-over situation, especially the remaining (rather than the total) time budget plays a crucial role in assessing the driver state.

Take-Over Mode: The term refers to the remaining performance of a given system after a RtI has been issued. The remaining performance generally depends on the automation level and on the design of a particular system. SAE L3 systems must be able to continue operation for a sufficient period of time to bridge the gap between RtI and manual

driving. Depending on the type of system limit system operation might be of degraded nature. SAE L2 systems may be fully deactivated in this phase.

The *driver state transition* process can be analyzed on multiple levels such as on a sensory, motoric and cognitive level. Relevant time markers within this phase are: interrupting/finishing non-driving related task, start of visual reorientation, gaze on road center, hands on wheel/feet on pedal, etc. Details on this particular state transition are further described in chapter 4 of this paper.

Start of Significant Driver Intervention/Deactivation of AD: For SAE L3/4 the automated driving system is supposed to remain active in take-over mode. In these cases significant driver intervention corresponds to the moment in time, when the automated driving system is deactivated by the driver. The way how manual control is resumed depends on the system design, but usually comprises regular deactivation mechanisms (e.g., based on dedicated buttons or switches) and significant driver intervention on primary vehicle controls. The user-initiated deactivation of automation is considered an essential timestamp in the transition process as it is linked to the definition of "driver take-over time". SAE L2 systems may already release vehicle control at the time of issuing the RtI. In this case specific thresholds for significant driver intervention on primary vehicle controls such as steering and pedals can be used to measure this time stamp [7].

Post Transition (Manual) Control: An extended time window to analyze the quality of manually controlled driving after a RtI has been issued. The post transition driving phase can be decomposed in the sub-phases *driver intervention* and the *control stabilization*.

Completion of Driving Manoeuver: The kind of take-over action that is expected by the driver to successfully handle the system limit. The intervention depends on the demands of the take-over situation. Examples range from easy tasks such as bringing back hands on steering wheel to follow the lane to more complex manoeuvers such as manual lane changes or brake manoeuvers etc. Depending on the take-over situation it can be difficult to determine the exact end of the driver intervention as there is a gradual transition to the control stabilization phase.

Vehicle Control Fully Stabilized: After driver intervention is completed manual vehicle control performance may initially be below the average performance of an individual driver [13]. The time period until vehicle control is fully re-established on an individual level is referred to as "control stabilization".

3 The Concept of Driver Availability

In order to objectively assess the success of the transition process both timing aspects (as specified in Fig. 1) and quality aspects (e.g., maximum accelerations) of the driver's actions need to be considered [7]. The authors agree that the combined description of timing and quality aspects determine the performance of a driver-vehicle system in a transition situation. One important implication of assessing *performance* is, however,

that it can only be done after the transition has been fully accomplished. By contrast, the concept of *driver availability* is introduced in order to assess the driver state before or during the transition process. The assessment is based on a model-based prediction of an individual take-over performance for a given system limit. By taking into account the requirements of a particular take-over situation, driver availability is not just an absolute description of the current driver state but a relative assessment that is influenced by the situational demands. This understanding is in line with recent definitions of driver distraction and inattention which also focus on the relation between the occupation of attentional resources and the requirements for safe driving [14, 15].

In the context of the Ko-HAF project driver availability is operationalized as a quantitative, time-based metric that relates the available total time budget at time t with the predicted duration of a successful transition to manual control for any particular moment t during automated driving (see formula 1). Quality demands of the transition performance are considered to be implicitly ensured as the underlying transition process model refers to a safe, controlled, and hence, successful take-over. The proposed measure DA (Driver Availability) indicates insufficient driver availability whenever the predicted duration of a safe take-over process (driver take over time + driver intervention time) exceeds the available time budget (see Fig. 1). This directly corresponds to the extent remaining action time (see Fig. 1) is available to carry out the take-over.

$$DA \sim t_{tb} - (\tilde{t}_{dto} + \tilde{t}_{di}) = \begin{cases} < 0, \ driverunavailable \\ > 0, \ driveravailable \end{cases} \quad (1)$$

with

$DA = Measure\ of\ Driver\ Availability$
$t_{tb} = time\ budget$
$\tilde{t}_{dto} = predicted\ driver\ take - over\ time$
$\tilde{t}_{di} = predicted\ driver\ intervention\ time$

The calculation of the proposed driver availability measure is expected to be possible at any point in time during automated driving. As long as no specific system limit is detected, a hypothetical take-over situation (e.g., the system limit most likely to occur) is assumed to assess the current state of driver availability. In case a RtI has already been issued the requirements of that particular situation are considered. In the latter case the driver availability measure could be continuously updated according to the predicted remaining duration of the required driver reaction.

The presented formula serves as one example how to relate aspects of the current driver state and situational demands - other expressions (e.g., as ratio) may be equally suitable, especially for transforming a time-based measure into a normalized DA index.

4 Modelling Driver Availability

4.1 Modeling Framework

A modelling framework is proposed in this chapter with the aim of bringing together and relating all relevant influencing factors on the take-over process. We assume that the "time budget" will be primarily based on the situational analysis module of the automated driving function [16, 17], while estimations of the "predicted driver take-over time" as well as the "predicted driver intervention time" are modelled by cognitive task analysis (see Fig. 2).

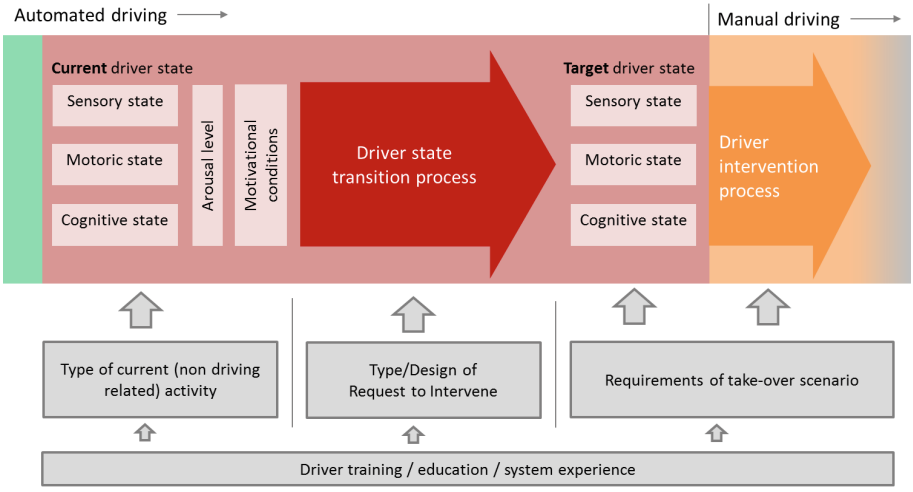


Fig. 2. Conceptual model of the processes required for assessing "driver availability"

During automated driving, the driver is likely to disengage from driving-related activities and engage in some sort of non-driving related task (NDRT) [18]. Taking back manual control over the vehicle, the so-called "driver state transition process" consists of a task switch from the NDRT to executing the driving tasks [19]. We propose that this task switch affords time consuming re-configuration of the drivers' *sensory*, *motoric* and *cognitive* state to meet the demands of manually controlling the vehicle. For example, re-configuration of the *sensory* state may consist of taking one's eyes off the NDRT, looking at the driving scene and attending to relevant HMI displays [9]. *Motoric* re-configuration could take place by freeing one's hands and taking them back on the steering wheel [20], while *cognitive* re-configuration takes place in the form of reconfiguration of mental task sets or response rules [21, 22].

We thus assume that the predicted duration of an individual take-over process can be modelled by considering all above-mentioned sub-processes required to reach the "target driver state" and by taking into account the type of required driver intervention. The proposed "driver state transition process" as presented in Fig. 2 considers multiple paths for *sensory*, *cognitive* and *motoric* transitions under specific arousal levels and

motivational conditions of the driver. Disengagement from driving related activities can lead to a state of passive fatigue [23–25], especially when there is no NDRT available to maintain a suitable arousal level [19], which may ultimately further prolong the state transition process. On the contrary, strongly motivating NDRTs may also have detrimental effects on the state transition process. The order of sub-processes involved in the re-configuration as shown in Fig. 2 may vary inter-individually and by the specific take-over situation. Non-interfering sub-processes may also be performed in parallel. Based on these assumptions, the framework also allows deriving hypotheses on how external factors influence the transition process:

- It is likely that the possibility to engage in NDRTs will be a major motivation for humans to make use of automated driving. As outlined above, the specifics of the NDRT processed during an automated ride will have a major impact on the duration of the transition process, as they determine the driver's *sensory*, *motoric* and *cognitive* state. A comprehensive overview on this topic is provided in [19].
- Driver characteristics such as experience with transition requirements [21, 26], risk tolerance [18, 27], or cognitive processing capabilities [27] are considered to influence the current driver state and how effectively an RtI is processed by the driver.
- The RtI itself can have significant effects on the driver state transition as it can be designed to communicate different levels of affordance and urgency [28]. Following assumptions of the multiple resource model, noticing and interpreting the RtI timely will to a large degree depend on whether the modalities of the NDRT and RtI do not draw on the same information processing channels [29].

4.2 Assessing the Current Driver State by Drive State Monitoring

To model the predicted duration of the transition process, in a first step the current driver state has to be assessed. Driver state monitoring systems allow for such an online/real-time assessment of relevant driver state variables (e.g., [30]). In accordance with the proposed modelling framework, the following building blocks of driver state monitoring are proposed:

For optimal performance an appropriate *arousal level* is necessary [31, 32]. Thus, the detection of drowsiness/fatigue (for definitions see [33]) is one key aspect for driver state assessment. Eye tracking systems can be used to obtain different reliable metrics, usually based on eye-lid movements over a given time interval (e.g., PERCLOS; see [34] for an overview). Additional information can be derived from physiological data (e.g., heart rate, EEG) and behaviors associated with drowsiness (e.g., yawning, change in body posture).

Another factor that influences the driver state is engagement in NDRTs. The driver's *sensory state* is affected by the modalities (e.g., visual, auditory) of a given NDRT. Visual distraction can be assessed with camera based head tracking (e.g., position and orientation [30]) and gaze detection (e.g., looking on/off the road, glance duration, glance patterns). Engagement in a visual task is usually associated with an increased number and duration of off-road glances, while an auditory task may lead to an increasing

gaze concentration to the road center [35]. For highly automated driving [11] showed that drivers' visual attention allocation is a predictor for take-over performance.

Driver's gaze behavior is also linked to his or her *cognitive state*, as the perception of the traffic environment is a basic prerequisite for maintaining situation awareness [36]. In addition, pupil diameter seems to be a sensitive measure for cognitive workload [37]. Apart from cognitive aspects, NDRT characteristics may also affect driver's *motivational conditions*, potentially leading to delayed task interruption in take-over-situations.

Lastly, the driver's *motoric state* plays a role for modeling the transition process from automated to manual driving. Longer periods of hands-off driving do not necessarily lead to performance deficits in take-over situations [20] but NDR-tasks with high motoric demands seem to be associated with increased take-over times (e.g., [5]). Driver's motoric state may be assessed in a direct way by using camera tracking with gesture recognition or indirect by using data from driver's interaction with in-vehicle information systems or attached devices [30].

4.3 Predicting the Target Driver State

The second step in predicting the duration of the transition process consists of the estimation of the required target driver state to solve the respective situation successfully. The target driver state is regarded to be influenced by the requirements of the take-over scenario such as the complexity of the take-over situation and the required driver tasks to successfully handle the critical situation:

- The available time budget of a given situation will determine how fast drivers need to react to RtIs and establish a suitable driver state to complete the situation manually [38]. On the other hand, situations of higher urgency may also speed up the driver state transition process, given that the situation's urgency is effectively communicated to the driver [28].
- The complexity of the required driver intervention (e.g., stabilising the vehicle vs. executing an avoidance manoeuvre) [28] as well as the complexity of the driving environment (e.g., traffic density [39]) will determine the requirements on the driver's *sensory*, *cognitive* and *motoric* state to handle the situation safely.

Although a growing number of studies has been conducted that can be used to estimate the required driver state (as was presented in the introductory chapter), there is still a need for significantly more research in this area.

5 Applications

One goal of the Ko-HAF project is to develop and validate a driver availability model that is able to predict human take-over performance under various driver states and situational conditions. The model is expected to summarize and integrate individual research findings not only to better understand the topic as such but also to actively use it as a software component, running in real-time as part of an automated driving system.

According to the calculated level of driver availability, it will be possible to adapt the vehicle behavior and the driver-vehicle interface/HMI during automated driving and to optimize the system for take-over situations.

Acknowledgements. This work results from the joint project Ko-HAF - Cooperative Highly Automated Driving and has been funded by the Federal Ministry for Economic Affairs and Energy based on a resolution of the German Bundestag.

References

1. Gasser, T.M., Westhoff, D.: BASt-study: Definitions of Automation and Legal Issues in Germany. In: Proceedings of the 2012 Road Vehicle Automation Workshop (2012)
2. Committee, S.O.-R.A.V.S.: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Technical report J3016_201401 (2014)
3. Lorenz, L., Kerschbaum, P., Hergeth, S., Gold, C., Radlmayr, J.: Der Fahrer im Hochautomatisierten Fahrzeug. Vom Dual-Task zum Sequential-Task Paradigma. In: TÜV Tagung Fahrerassistenz München (2015)
4. Damböck, D.: Automationseffekte im Fahrzeug – von der Reaktion zur Übernahme. München, Technische Universität München, Dissertation (2013)
5. Petermann-Stock, I., Hackenberg, L., Muhr, T., Mergl, C.: Wie lange braucht der Fahrer? Eine Analyse zu Übernahmezeiten aus verschiedenen Nebentätigkeiten während einer hochautomatisierten Staufahrt. 6. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren (2013)
6. Eriksson, A., Stanton, N.A.: Takeover time in highly automated vehicles - noncritical transitions to and from manual control. Hum. Factors (2017). doi: 10.1177/0018720816685832
7. Gold, C.G.: Modeling of take-over performance in highly automated vehicle guidance., München, Technische Universität München, Dissertation (2016)
8. Miller, D., Sun, A., Johns, M., Ive, H., Sirkin, D., Aich, S., Ju, W.: Distraction becomes engagement in automated driving. Proc. Hum. Factors Ergon. Soc. Annu. Meet. **59**, 1676–1680 (2015)
9. Gold, C., Damböck, D., Bengler, K., Lorenz, L.: Partially automated driving as a fallback level of high automation. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren **28**, 2013 (2013)
10. German Road Safety Council. http://www.dvr.de/dvr/beschluesse/ft_automatisiertesfahren_engl.htm
11. Zeeb, K., Buchner, A., Schrauf, M.: What determines the take-over time? An integrated model approach of driver take-over after automated driving. Accid. Anal. Prev. **78**, 212–221 (2015)
12. Blanco, M., Atwood, J., Vasquez, H.M., Trimble, T.E., Fitchett, V.L., Radlbeck, J., Fitch, G.M., Russell, S.M.: Automated vehicles: take-over request and system prompt evaluation. In: Meyer, G., Beiker, S. (eds.) Road Vehicle Automation, vol. 3, pp. 111–119. Springer, Cham (2016)
13. Merat, N., Jamson, A.H., Lai, F.C.H., Daly, M., Carsten, O.M.J.: Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. Transp. Res. F **27**(Part B), 274–282 (2014)
14. Victor, T.W., Engström, J., Harbluk, J.L.: 10 Distraction assessment methods based on visual behavior and event detection. Driv. Distraction **135** (2008)

15. Regan, M.A., Hallett, C., Gordon, C.P.: Driver distraction and driver inattention: definition, relationship and taxonomy. Accid. Anal. Prev. **43**, 1771–1781 (2011)
16. Weidl, G., Breuel, G.: Overall probabilistic framework for modeling and analysis of intersection situations. In: Advanced microsystems for automotive applications 2012, pp. 257–268. Springer, New York (2012)
17. Weidl, G., Madsen, A.L., Kasper, D., Breuel, G.: Optimizing Bayesian networks for recognition of driving maneuvers to meet the automotive requirements. In: Intelligent Control (ISIC), 2014 IEEE International Symposium on, pp. 1626–1631. IEEE (2014)
18. Naujoks, F., Purucker, C., Neukum, A.: Secondary task engagement and vehicle automation–comparing the effects of different automation levels in an on-road experiment. Transp. Res. F **38**, 67–82 (2016)
19. Naujoks, F., Befelein, D., Wiedemann, K., Neukum, A.: A review of non-driving-related tasks used in studies on automated driving. In: Paper presented at the 8th Conference on Applied Human Factors and Ergonomics, Los Angeles (2017)
20. Naujoks, F., Purucker, C., Neukum, A., Wolter, S., Steiger, R.: Controllability of partially automated driving functions–does it matter whether drivers are allowed to take their hands off the steering wheel? Transp. Res. F **35**, 185–198 (2015)
21. Allport, D.A., Styles, E.A., Hsieh, S.: Shifting intentional set: Exploring the dynamic control of tasks. In: Moscovitch, C.U.M. (ed.) Attention and performance 15: conscious and nonconscious information processing, pp. 421–452. The MIT Press, Cambridge (1994)
22. Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A.M., Koch, I.: Control and interference in task switching - a review. Psychol. Bull. **136**, 849–874 (2010)
23. May, J.F., Baldwin, C.L.: Driver fatigue: the importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Transp. Res. F **12**, 218–224 (2009)
24. Körber, M., Cingel, A., Zimmermann, M., Bengler, K.: Vigilance decrement and passive fatigue caused by monotony in automated driving. Proc. Manuf. **3**, 2403–2409 (2015)
25. Neubauer, C., Matthews, G., Langheim, L., Saxby, D.: Fatigue and voluntary utilization of automation in simulated driving. Hum. Factors **54**, 734–746 (2012)
26. Larsson, A.F.L.: Driver usage and understanding of adaptive cruise control. Appl. Ergon. **43**, 501–506 (2012)
27. Körber, M., Bengler, K.: Potential individual differences regarding automation effects in automated driving. In: Proceedings of the XV International Conference on Human Computer Interaction, pp. 1–7. ACM, Puerto de la Cruz, Tenerife (2014)
28. Naujoks, F., Mai, C., Neukum, A.: The effect of urgency of take-over requests during highly automated driving under distraction conditions. Adv. Hum. Asp. Transp. **431** (2014)
29. Wickens, C.D., Liu, Y.: Codes and modalities in multiple resources: a success and a qualification. Hum. Factors **30**, 599–616 (1988)
30. Rauch, N., Kaussner, A., Boverie, S., Giralt, A.: HAVE it deliverable D32.1 report on driver assessment methodology (2009)
31. Yerkes, R.M., Dodson, J.D.: The relation of strength of stimulus to rapidity of habit-formation. J. Comp. Neurol. Psychol. **18**, 459–482 (1908)
32. Stanton, N.A., Young, M.S.: Automotive automation: investigating the impact on drivers' mental workload. Int. J. Cogn. Ergon. **1**, 325–336 (1997)
33. Phillips, R.O.: A review of definitions of fatigue – and a step towards a whole definition. Transp. Res. F **29**, 48–56 (2015)
34. Friedrichs, F., Yang, B.: Camera-based drowsiness reference for driver state classification under real driving conditions. In: 2010 IEEE Intelligent Vehicles Symposium, pp. 101–106. (2010)

35. Victor, T.W., Harbluk, J.L., Engström, J.A.: Sensitivity of eye-movement measures to in-vehicle task difficulty. Transp. Res. F **8**, 167–190 (2005)
36. Endsley, M.R.: Final reflections: situation awareness models and measure. J. Cogn. Eng. Decis. Mak. **9**, 101–111 (2015)
37. Marquart, G., Cabrall, C., de Winter, J.: Review of eye-related measures of drivers' mental workload. Proc. Manuf. **3**, 2854–2861 (2015)
38. Beukel, A.P.V.D., Voort, M.C.V.D.: The influence of time-criticality on Situation Awareness when retrieving human control after automated driving. In: 16th International IEEE Conference on Intelligent Transportation Systems, pp. 2000–2005. (ITSC 2013) (2013)
39. Gold, C., Korber, M., Lechner, D., Bengler, K.: Taking over control from highly automated vehicles in complex traffic situations: the role of traffic density. Hum. Factors **58**, 642–652 (2016)

What to Expect of Automated Driving: Expectations and Anticipation of System Behavior

Johanna Josten[1(✉)], Teresa Schmidt[2], Ralf Philipsen[2], Lutz Eckstein[1], and Martina Ziefle[2]

[1] Institute for Automotive Engineering, RWTH Aachen University, Steinbachstr. 7, 52074 Aachen, Germany
{josten, eckstein}@ika.rwth-aachen.de
[2] Human Computer Interaction Center, RWTH Aachen University, Campus-Boulevard 57, 52074 Aachen, Germany
{schmidt, philipsen, ziefle}@comm.rwth-aachen.de

Abstract. A realistic view on what new technology can or cannot handle safely should enhance the interaction of driver and technology from first activation on. Wrong assumptions on how technology functions might result in critical driving situations or prevent the activation of an otherwise useful system. Addressing different forms of information given to a driver, a two-tiered research approach was conducted. A focus group study (N = 11) on users concerns regarding automated driving revealed the need of information provided to users to gain an understanding of the function. Further, the usefulness of specific initial information to enhance prediction of system behavior in selected driving situations was focused on in an online survey (N = 108). The results are discussed with regard to further use in behavioral studies on take-over anticipation and intention to use automated systems.

Keywords: Human factors · Automated driving · Initial information

1 User Centered Research for New In-Vehicle Technology

Currently, vehicle automation is one of the most prominently discussed research fields in automotive engineering. While first partially automated driving functions are introduced into the market, open issues remain from the human factors' view, concerning, for example, the factors that influence a successful interaction of drivers with different knowledge and expertise on assistance systems. A large body of studies has already been conducted with focus on the process of control shifts from vehicle to driver, i.e. on take-over situations (e.g. [1, 2]). Research foci were, amongst others, on different time intervals for take-over requests [3] or on driver state during automated driving (e.g. [4, 5]). While the resulting take-over quality is one of the most detailed researched measures of a successful and safe interaction between system and driver, research should also consider the user's initial concerns and mental concepts on an automated system, as they likely influence whether the system will be accepted, trusted and used.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_59

Using adaptive cruise control (ACC) as an exemplary driver assistance system, initial information on the system's capabilities and characteristics was shown to influence attitudes and trust of drivers towards the system [6]. The same authors showed that experiencing system failures had no negative influence on trust and acceptance if correct information on system limits was provided beforehand.

Regarding the first contact with automated driving, the lack of experience combined with wrong expectations and assumptions on the system's capabilities could result in the occurrence of non-expected take-over situations. Non-anticipated take-over requests might however lead to longer intervention times of the driver, probably increase the criticality of the situation and lower the trust in the system, similar to the decrease in trust observed after initially incomplete information on ACC [6].

Aspects of a system's functionality, such as the degree of freedom the driver is given when using the system, also have a direct influence on the interaction in take-over situations: Monitoring a partially automated system [7] without keeping the hands on the steering wheel results in persistently higher take-over times in non-anticipated situations [8]. Allowing a higher degree of freedom, i.e. a lesser direct engagement with the system, could result in a higher proneness for secondary tasks, even when the system still should be monitored by the driver. Only behavioral effects with strict control over driver distraction were investigated in [8] and no overall decrease in monitoring performance could be observed for hands free driving. However, it remains unclear whether introducing a system initially with a higher degree of freedom leads the driver to assuming a different importance of monitoring outside an experimentally controlled setting or raises the need for information from the system.

Users might also have concerns about shifting control to an unknown system and exhibit initially low trust. Familiarity, control and knowledge about risks have for example been investigated as factors for the perceived risk of hazards [9]. Furthermore, for new V2X-technology, user concerns regarding the data processed and stored by the system as well as fear for loss of control over data were pointed out [10]. To increase the quality of user interaction from first contact on, it is important to address users concerns and to understand factors that influence assumed system characteristics and could enhance the correct anticipation of system behavior.

To address both of these research needs, i.e. perceived user requirements for trust and control in automated driving and the influence of initial information, a two-step procedure was designed. Integrating insights from laypeople (non experts), the estimation of the system's capabilities was tested in three selected highway driving situations.

2 Understanding Expectations

A qualitative approach was chosen to collect input on which information needs are requested by laypeople before using functions of higher automation levels for the first time. Furthermore, expectations and concerns regarding automated functions were collected.

2.1 Method

Two focus groups were conducted by a trained interviewer in October 2016. The discussions were structured in two parts lasting 90 to 100 min. First, an introduction video about automated driving was presented to the participants. It focused on currently not available, higher automation levels, i.e. on functions that do not require the driver to monitor the function, and contained information about new possibilities for the driver (e.g. hands free driving), an overview of sensor networking and some design elements. This way, considerations about needed information, functionality, trust and control were based on a common basic knowledge of all participants. The discussion afterwards consisted of (a) informative questions (e.g. about own experience with driver assistance systems), (b) general questions (e.g. which information is important for trusting the system) and (c) specific questions (e.g. influence of transparent information usage on trust).

Eleven laypeople, who had never worked in a field related to automated driving or automotive engineering, were invited to discuss their attitudes about automated driving. Eight women (f) and three men (m), between 22 and 27 years old, participated in the two focus groups. All of them worked in and/or studied non-technical fields. The participants attended voluntarily and were not compensated for their participation. They were chosen from a group of interested persons based on their habits to represent the generation of early adopter millennials.

2.2 Results

Overall, the participants were in all cases hesitant to trust an automated system and showed diverse levels of requirements. When it comes to control, a strong longing for information as well as quick and easy ways to shift control (and responsibility) back to the driver was communicated. The majority had theoretical knowledge about several driver assistance systems. When it came to actual experience, the groups showed only little expertise.

Concerning control, the general impression gained from comments was that all the participants demand control or the possibility to regain control immediately. One participant expressed the feeling, that she wants "to keep the control anytime" (f, 27 years) and wished for a reassurance that "the system works properly". Further, another participant stated similar thoughts:

> *"I think it should be accessible in every case. Especially, if I grab the steering wheel, it should be clear that I take back the control from the system. Otherwise it would not be an option for me to buy such a car" (m, 24 years)*

Regarding a system in control, participants discussed whether age would be another reason to mistrust an automated function to one's safekeeping:

> *"(…) never, especially old people would never say 'I let the car control itself'. So that would never happen" (m, 24 years)*

A further broadly discussed topic concerning control and trust was the possibility of being (technically) attacked. Here, some of the participants were unsure about the

technical process of being hacked and proposed a practical solution, e.g. "there is a button" (f, 27 years) to regain control over the vehicle. The argumentation was that the vehicle has to have a (haptic) function, which leads to controlling the car again.

Several persons mentioned a general skepticism about the interaction with other (non-automated) vehicles:

> *"(…) I am rather skeptical, because I would only trust [the technology], if every car would drive automated and as long as people are driving by themselves or there are speeders, I would not put myself in there in full confidence" (f, 23 years).*

Another participant questioned if there would be information provided about function failure in her own car or in other vehicles. Her argumentation referred to sensor failures, while e.g. watching a movie on the road.

Addressing the trust in automated driving functions, a diverse picture evolved. The majority of the participants agreed that recommendations of friends and/or family members would strengthen their trust in the system. Different traffic scenarios, like driving on a highway or using automated functions in urban traffic, led to diverse opinions about the capabilities of a system:

> *"(…) I would say all the time, if the highway is crowded, that I would like to have control and see for myself which lane is the fastest" (f, 24 years).*

Further, the question which information needs to be provided to the user in order to trust an automated car was discussed. The participants had several ideas about that topic, e.g. "a voice, which is talking to me or at least the car gives feedback" (f, 23 years). Another participant addressed the possibility to set own preferences to gain knowledge about what the system is currently performing:

> *"(…) you can simply adjust what you want to hear, that one has the possibility to do so. Because, like was already said, ok, there is the end of a traffic jam or some-thing failed. Or you can say beforehand, I want to get this message, but not that one" (m, 25 years).*

Overall, the studies revealed a general uncertainty about what automation can do, e.g. finding a suitable lane on crowded highways, and in which situations it might be useful. Further, the focus groups voiced concerns regarding the use of automated system, i.e. perceived loss of control or fear for hacking, which might prohibit the use of such a system. Furthermore, the need for selected information displayed by the system was expressed.

3 Initial Information and Degree of Freedom

Different expectations were voiced in the focus groups regarding the suitability and benefit of higher levels of automated driving, albeit providing a common basis of understanding using an instruction video. Different expectations might however lead to problems when interacting with a specific automated function for the first time, as the system will most probably work differently than is expected by the user. Especially critical in this regard could be a short-termed take-over request issued by the system and not expected by the driver. Providing detailed initial information about a system before first use might enhance the understanding of what is to be expected from the

system and which responsibilities remain with the driver. This might in turn lead to higher trust in the system as following experience would strengthen the correct initial mental model, similar to what has been observed for ACC [6].

To investigate the influence of initial information on system limits and the allowed degree of freedom when using a system, an online survey was conducted. The presented results are part of a more extensive survey on intention to use, data sharing, expected behavior of systems as well as on overall information needs when using automated systems. The survey, in contrast to the focus group, included two systems with different responsibilities of the driver, i.e. requiring him to be in the loop at all times or to shift responsibility for monitoring to the system, as well as the necessity to keep contact with the steering wheel at all times.

Data of 108 participants who answered more than 50% of situational assessment questions are reported. Participants (49.1% female) are between 18 and 77 years old (M = 30, SD = 10). Mean driving experience per year is 9426 km (SD = 10235). Two of the participants do not hold a driving license (German license type B). In difference to the focus group sample, 35.2% of participants had already used ACC at least once and 75% had used Cruise Control (CC) at least once. Due to exclusion of missing values, participants are not equally distributed into the four information groups of the survey shown in Table 1. Group sizes varied between N = 19 and N = 39.

3.1 Method

The introduction of systems was manipulated between participants with a focus on details of system limits, in the following referred to as *initial information*. Furthermore, the *degree of freedom* the driver was given while using the system was manipulated, i.e. either having to keep physical contact with the steering wheel (*hands on*) or to be able to drive hands free (*hands off*). Each participant received one of four possible combinations of introduction and rated a partially (PAD) as well as a conditionally automated system (CAD) [7] based on the information provided (see Tables 1 and 2).

Table 1. Manipulation of initial information and system description in the survey

Rental car scenario – First use of a new system Different kinds of information provided to participants (divided into four groups):			
Unspecific information System gives back control if system limits are reached		*Specific information* System gives back control if system limits are reached, i.e. when lane markings are difficult to detect or when own lane ends	
Hands off allowed	Hands off not allowed	Hands off allowed	Hands off not allowed
Partially automated system (PAD) [7] Driver responsible for monitoring, take-over required immediately if system limits are reached			
Conditionally automated system (CAD) [7] System responsible for monitoring, take-over required after notification by system			

This paper will focus on system assessment and information needs in specific situations based on different initial information. Three situations were selected for display in the survey, differing in visual cues relating to the instructed system limits. To investigate whether take-over expectations are related to situational assessment when driving manually, one supposedly uncritical situation for manual driving was included with direct relation to take-over relevant information. The static driving situations were taken from video-material collected on German highways, shot with full view of all lanes ahead. Each situation included visible traffic ahead. Two of the situations included explicit take-over related cues, based on details of the given specific initial information. Take-over relevant cues were the end of one's own lane, referred to as *lane ends*, and missing lane markings for separation of lanes on the highway, referred to as *lane markings*. The third situation presented unclear cues to both the specific and the unspecific information group, with temporary, narrow lane markings due to construction works as well as a speed-limit change indicated by dynamic traffic signals above the road. This situation is referred to as *construction*.

The situation *lane markings* was expected to be perceived as clearly uncritical in manual driving whereas a system's reaction to this situation might be unclear to uninformed drivers. One supposed side effect of detailed initial information on system limits would be for drivers to exhibit a higher awareness for the necessity of being in the loop in situations directly linked to take-over reasons. This should be especially true when using a partially automated system where the responsibility for monitoring lies with the driver. Therefore, ratings for all three situations were analyzed regarding the subjective importance of the driver having a detailed overview over the current situation when using the automated system. To further control for a more direct effect of the information manipulation, certainty to predict the system's behavior in the current situation as well as the probability for a take-over was compared between situations. It was expected that the specific information given would be linked to the situations displayed, resulting in higher subjective certainty how the system would behave in a specific situation. To gain insights into which information participants believe is furthermore necessary to better predict system behavior, demanded information was sorted into different categories. This was done for each situation in order to take into account the additional cues in the construction zone that were not referred to initially in any of the groups.

3.2 Results

If necessary, corrected values due to unequal variances between groups are reported. Pair wise comparisons are reported with corrected p-values if performed as post-hoc tests. Probability, e.g. for take-over, and certainty of system behavior was measured in percent (0–100%). Other variables, e.g. the importance of the driver being in the loop, were measured on a 6-point likert scale (1 = "do not agree at all"; 6 = "totally agree"). Only participants with valid ratings in each category of comparisons are included.

As a first step of evaluation, ratings when driving manually in the displayed situation were compared regarding the probability for critical driving maneuvers, the complexity of the situation as well as the probability for driving errors. A first comparison of the mean ratings between the four information groups revealed no significant

differences, indicating that no general differences in criticality assessment ($F < 1$) and perceived complexity of the situation [$F(3,106) = 1.1$, $p = .34$] exist between groups.

However, rating differences were found between the three situations regarding the probability for critical driving situations [$F(1.8,168.8) = 32.7$, $p < .001$, $\eta_p^2 = .25$], the complexity of the situation [$X^2(2) = 40.3$, $p < .001$], and the probability for driving errors caused by the driver [$F(2,194) = 36.4$, $p < .001$, $\eta_p^2 = .27$]. In comparison to the situations *lane ends* and *construction*, the situation *lane markings* was found to have a significantly lower probability for critical maneuvers (pair wise comparisons both $p < .001$; mean difference to *lane ends*: 18.5%; mean difference to *construction*: 21%) as well as being significantly less complex than both other situations (pair wise comparisons both $p < .001$). The probability for driving errors was rated significantly higher for situation *lane ends* compared to both other situations ($p < .001$; mean difference to *lane markings*: 24.9%; mean difference to *construction*: 12.1%) as was *construction* in comparison to *lane markings* ($p < .001$; mean difference: 12.7%). Thus, missing lane markings are not rated as critical for manual driving, whereas the end of one's lane signals higher probabilities for critical maneuvers. If situation criticality is a relevant decision criterion for take-over anticipation, even the unspecific information group might rate the end of the own lane as a possible take-over situation. In comparison, missing lane markings might only be noticed as a reason for system initiated take-over by the group receiving specific information on system limits. Despite significant differences, all situations received only low to medium ratings, e.g. the probabilities for critical driving maneuvers were 39.8% for *construction*, 37.2% for *lane ends* and 18.8% for *lane markings*.

The mean importance of the driver being in the loop, i.e. the importance of the driver being aware of the traffic situation, was compared between groups over all three situations ($N = 104$). The level of automation did not have an influence on the subjective importance of being in the loop in the three situations ($F < 1$), although instruction clearly specified that the CAD system would monitor the situation itself. No significant effect of initial information or any interaction with initial information or automation (all $Fs < 1$) was found either. However, the importance of being in the loop increased significantly for hands off monitoring [$F(1,100) = 6.8$, $p < .05$, $\eta_p^2 = .06$]. As can be seen in Fig. 1, the overall subjective importance of the driver being in the loop was very high in all groups.

The most direct assessment of system behavior related to initial information on system limits, i.e. take-over probability, was assessed for each of the two situations with visual cues directly related to take-over requests. As initial information on system limits was kept the same between CAD and PAD, a mean rating was analyzed for each participant rating both system ($N = 98$).

As expected, the situation rated uncritical in manual driving was also rated as significantly less likely to result in a take-over request [$F(1,94) = 46.7$, $p < .001$, $\eta^2 = .33$]. However, a take-over request was never expected with very high certainty ($M = 64\%$ probability for take-over when lane ends; $M = 37\%$ probability for missing lane markings). The main effect for initial information was only marginally significant [$F(1,94) = 3.3$, $p = .07$], as was the interaction between initial information and type of situation ($F(1,94) = 3.4$, $p = .07$) with the difference between information groups being larger in the manually uncritical situation, i.e. *lane markings* (see Fig. 2). The

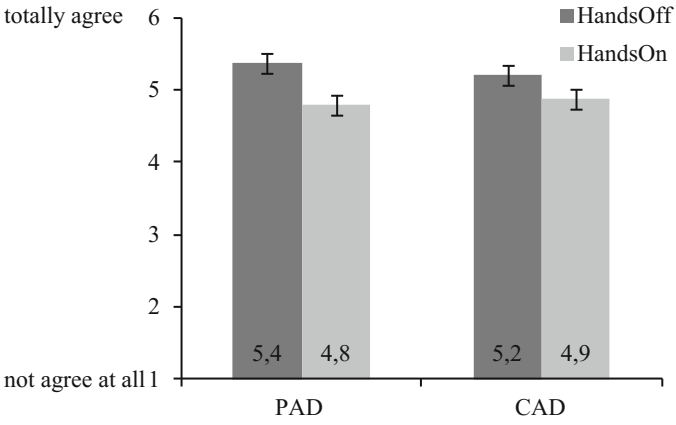


Fig. 1. Subjective importance of the driver having a detailed overview over the current situation in partially (PAD) and conditionally (CAD) automated driving for the two degrees of freedom (HandsOff and HandsOn).

interaction at least indicates a transfer of the information read to the ratings given, as the subjectively more uncritical situation in manual driving is assigned a higher take-over probability by the specific information group. No other significant effects were found.

As an additional, more general measure, the overall certainty of the system's behavior in the displayed situation was analyzed descriptively. As expected after analysis of take-over probability, overall certainty was rather low and did not show differences between specific and unspecific information groups (61% for specific information, 62.5% for unspecific information). The same pattern already observed for take-over probability could also be observed for overall certainty, with the only difference between information groups being visible in case of missing lane markings (68.4% for specific information, 73.1% for unspecific information). Here, the difference is however opposite to that shown for take-over probability, showing slightly lower certainty for the group receiving more information.

Table 2 lists different areas of information necessary for a better prediction of the system's behavior as requested by participants. Information reported as important to improve prediction was sorted thematically for the specific and unspecific information group for each situation by two evaluators. Only 20% of participants (8 out of N = 41) in the unspecific information group demanded information on whether the system could detect the upcoming lane change for the situation *lane ends*. For missing lane marking, again 20% of the group receiving unspecific system limits demanded information on how the system's orientation on the road is realized without lane markings and whether the system needed lane markings at all.

Although the specific information group was told that the end of one's lane and problems to detect lane markings would lead to a take-over situation, this information was apparently not sufficient to predict the systems behavior with a high certainty as can be seen in the analyses conducted so far. Analysis on information needs revealed

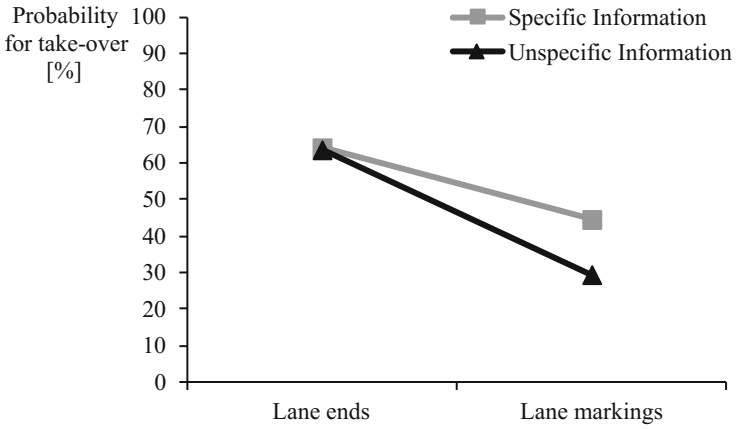


Fig. 2. Mean subjective probability for take-over in the two situations including information directly related to initial information. Separate means displayed for the groups receiving specific and unspecific information on system limits.

Table 2. Types of information as requested by participants for a better prediction of the system's behaviour in the three situations *lane ends*, *lane markings* and *construction*

Situation specific information	General requested information
Detection of lane change necessity	Detected objects and quality of detection
Ability to handle lane change	Surrounding traffic (e.g. behavior, V2V possibilities, automation level)
Orientation on road without lane markings	Planned maneuvers
Ability to handle construction markings	Own speed
Detection of speed limit	Priority of information processed by system

that some participants explicitly remarked that they were not sure how the system would react to partly missing lane markings or to temporary lane markings in different colors (14.9%; 10 out of N = 67). This might lead to lower certainty of system behavior even with higher detail in initial information, as observed for the situation *lane markings*. Thus, general information on system limits might not be enough to link the described system behavior to specific situations with slightly different features each.

Other information deemed relevant in the eyes of participants was the detection of surrounding traffic as well as the prediction of behavior of surrounding traffic, i.e. information relevant for interaction with other traffic users. For the situation *construction*, only 3.7% of all analyzed participants (4 out of N = 108) demanded information on system behavior related to the changing speed limit. It remains unclear whether the speed limit was not deemed important or not noticed by the other participants.

Overall, only half of the participants (51.2% of the unspecific group, 43.3% of the specific group) indicated for at least one situation which specific information is lacking to predict the system's behavior more precisely. Additionally, 11.1% of participants indicated for at least one situation that they could not name any specific missing information as basic experience with the system would be required first.

4 Discussion

In order to familiarize drivers with automated systems and raise their initial trust, concerns and information demands regarding the interaction of driver and vehicle need to be investigated. Expectations on automated driving were collected in focus group studies, whereas the influence of information given to the driver about the system was analyzed in an online survey.

The focus groups revealed that initial ideas and expectations about highly automated functions differ greatly between participants with low expertise regarding state-of-the-art driver assistance systems. Use in mixed-traffic, i.e. traffic including automated and non-automated vehicles, was seen as especially problematic. Interaction with other automated traffic raised the need for information on correct functioning in the focus group and in the survey.

The online survey was designed to deliver first input on how detailed initial in-formation and the driver's degree of freedom change the subjective importance of the driver being in the loop when the system is active as well as the anticipation of the system's behavior. The anticipation of take-over situations was analyzed in a first selection of three highway situations. At least in PAD, where the driver is still responsible for monitoring the system and environment, a high subjective relevance of being in the loop would be desirable from first activation on.

The importance of the driver being in the loop was however perceived as high regardless of provided initial information and rated automation level. Even the conditional automated system was thus not seen fit to dismiss the driver from active visual involvement with the driving task. The equally high importance of the driver being in the loop might be an expression of low trust in the system due to a lack of experience and could also be related to the perceived high loss of control reported by the focus group. It further remains unclear whether the difference between the two automation levels, PAD and CAD, was sufficiently clear to participants.

Allowing hands free driving does increase the subjective importance of being visually in the loop prior to actually using the system. This effect is positively surprising as take-over was shown to be significantly slower in non-anticipated driving situations when monitoring the system hands free [8]. A higher awareness of monitoring the system in combination with a correct mental model could supposedly enforce early anticipation of a take-over and dissolve the negative effects found previously for hands free monitoring.

Higher detail in information on system limits does only marginally change the anticipation of a take-over for situations linked directly to instructed system limits. This effect might also be due to characteristics of the situations selected for this survey. Although the situations differed significantly in their criticality and complexity,

the overall ratings for criticality in manual driving were very low for all three situations. Furthermore, one participant commented on the situation with missing lane markings as being supposedly safe even when driving with an automated system. One explanation could be that the situation without lane markings was too uncritical. However, the situation rated less critical in manual driving was also rated as significantly less likely to lead to a take-over request, indicating that expectations on successful situation management by a driver were transferred to a successful handling by the function.

As only half of the participants requested additional information in the situations to better anticipate system behavior, general conclusions are difficult to draw. Comments, however, showed that general information on system limits was not enough to gain overall higher certainty on how the system would behave in specific situations. Online information on what the system is processing and which maneuvers can be expected might be required instead, as well as displaying online information on system confidence, which was found to have positive effects on the interaction with ACC [11].

Furthermore, the conducted survey has multiple limitations such as sample size and low ecological validity. As only a small sample of three static driving situations was investigated, more situations with higher differentiation regarding criticality in manual driving and their reference to system initiated take-over requests should be taken into account in follow-up studies. In focus of further studies should therefore be the question whether initial information can raise take-over performance in first encounter with critical situations and whether information detail has an influence on the intention to use a specific system. Studies including dynamic material and driving simulation will be used to investigate when drivers decide for a higher take-over probability or when they detect possible take-over related cues, similar to studies on hazard perception (e.g. [12]).

Acknowledgments. This research was supported by the Profile Area ICT of RWTH Aachen University (RWTH-PH-ICT).

References

1. Zeeb, K., Buchner, A., Schrauf, M.: Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. Accid. Anal. Prev. **92**, 230–239 (2016)
2. Merat, N., Jamson, A.H., Lai, F.C.H., Daly, M., Carsten, O.M.J.: Transition to manual: driver behaviour when resuming control from a highly automated vehicle. Transp. Res. Part F **27**, 274–282 (2014)
3. Gold, C., Damböck, D., Lorenz, L., Bengler, K.: "Take over!" How long does it take to get the driver back into the loop? Proc. Hum. Factors Ergonom. Soc. Ann. Meet. **57**, 1938–1942 (2013)
4. Merat, N., Jamson, A.H., Lai, F.C.H., Carsten, O.M.J.: Highly automated driving, secondary task performance, and driver state. Hum. Factors **54**, 762–771 (2012)
5. Goncalves, J., Happee, R., Bengler, K.: Drowsiness in conditional automation: proneness, diagnosis and driving performance effects. In: IEEE 19th International Conference on ITSC, pp. 873–878 (2016)

6. Beggiato, M., Krems, J.F.: The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. Transp. Res. Part F **18**, 47–57 (2013)
7. SAE On-Road Automated Vehicle Standards Committee. Information Report J3016: Taxonomy and Definitions (2014). https://www.sae.org/misc/pdfs/automated_driving.pdf
8. Josten, J., Zlocki, A., Eckstein, L.: Untersuchung der Bewältigungsleistung des Fahrers von kurzfristig auftretenden Wiederübernahmesituationen nach teilautomatischem, freihändigen Fahren. FAT-Schriftenreihe 289 (2016)
9. Mullet, E., Duquesnoy, C., Raiff, P., Fahrasmane, R., Namur, E.: The evaluative factor of risk perception. J. Appl. Soc. Psychol. **23**, 1594–1605 (1993)
10. Schmidt, T., Philipsen, R., Themann, P., Ziefle, M.: Public perception of V2X-technology – evaluation of general advantages, disadvantages and reasons for data sharing with connected vehicles. In: IEEE Intelligent Vehicles Symposium, pp. 1344–1349 (2016)
11. Stockert, S., Richardson, N.T., Lienkamp, M.: Driving in an increasingly automated world – approaches to improve the driver-automation interaction. Proc. Manuf. **3**, 2889–2896 (2015)
12. Crundall, D., Chapman, P., Trawley, S., Collings, L., van Loon, E., Andrews, B., Underwood, G.: Some hazards are more attractive than others: drivers of varying experience respond differently to different types of hazard. Accid. Anal. Prev. **45**, 600–609 (2012)

The Design of a Vibrotactile Seat for Conveying Take-Over Requests in Automated Driving

Sebastiaan M. Petermeijer[1(✉)], Paul Hornberger[1], Ioannis Ganotis[1], Joost C.F. de Winter[2], and Klaus. J. Bengler[1]

[1] Lehrstuhl Für Ergonomie, Technische Universität München, Boltzmannstraße 15, 85747 Garching, Germany
{petermeijer,bengler}@tum.de
[2] Department of Biomechanical Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands
j.c.f.dewinter@tudelft.nl

Abstract. The driver of a conditionally automated car is not required to permanently monitor the outside environment, but needs to take over control whenever the automation issues a "request to intervene" (i.e., take-over request). If the driver misses the take-over request or does not respond in a timely and correct manner, a take-over could result in a safety-critical scenario. Traditionally, warnings in vehicles are conveyed by visual and auditory displays, though recently it has been argued that vibrotactile stimuli could also be a viable approach to present a take-over request to the driver. In this paper, we present a vibrotactile seat designed to convey dynamic vibration patterns to the driver. The seat incorporates 48 vibration motors (eccentric mass rotation) that can be individually controlled. One 6×4 matrix, with an average inter-motor distance of approximately 4 cm, is located in the seat back and one in the seat bottom. The DC-voltage to the motors is controlled by three Pulse Width Modulation (PWM) drivers, which in turn are controlled by an Arduino microcontroller. A study with 12 participants was conducted to investigate (1) at which vibration intensity participants find a vibratory stimulus annoying and whether this threshold changes over time, (2) how well participants are able to discriminate vibratory stimuli as a function of spatial separation, and (3) which of six dynamic vibration patterns are regarded as most satisfying. Results showed that participants' annoyance threshold reduced when they were repeatedly exposed to vibrotactile stimuli. Second, the percentage of correct responses in the two-point discrimination test increased significantly with increasing inter-stimuli distance (i.e., from 4 to 20 cm). Third, participants seemed to be more satisfied when more motors were activated simultaneously (i.e., more spatial overlap). Overall, the results suggest that participants are well able to perceive vibrotractile stimuli in the driver seat. However, the results suggest that repetitive exposure to vibrotactile stimuli may evoke annoyance, a finding that should be taken into account in future designs of vibrotactile displays. Future studies should investigate the possibility to convey complex messages via the vibration seat.

Keywords: Vibrotactile interface · Conditionally automated driving · Human factors

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_60

1 Introduction

1.1 Conditionally Automated Driving and the Take-Over Process

Conditionally automated driving may be introduced on public roads within the next decade [1, 2]. SAE International defines "conditional automation" as automation that controls the driving task, with the expectation that the driver takes back control when a "request to intervene" is presented [3]. This means that the driver will be allowed to take his hands and feet off the steering wheel and pedals, and may engage in non-driving tasks such as reading or making a phone call.

A take-over request is presented to the driver when the automation reaches its operational limits. Examples of such situations are when the automation enters a complex traffic situation it cannot solve, when a sensor has failed, or when rules and regulations require human involvement in the driving task. After receiving a take-over request, the driver is expected to take back control in a timely and safe manner. In this process the driver performs the following four temporally overlapping actions: (1) shifting the attention towards the road, (2) cognitively processing the current traffic situation and selecting an appropriate action, (3) repositioning him/herself to take back control (i.e., feet on pedals and hands on steering wheel), and (4) performing the selected action [4–7].

Gold et al. [5] presented drivers with an auditory-visual take-over request, and found that the eyes-on-road reaction time (i.e., time to shift of attention back to the road) was on average 0.8 s. Furthermore, drivers made a steering or braking action after 2.1 or 2.9 s for lead times (i.e., TTC at the moment of the take-over request) of 5 and 7 s, respectively. Thus, shorter lead times yielded faster response times. However, shorter lead times also yielded a diminished take-over quality defined as the level of lateral and longitudinal accelerations. A well-designed interface could assist the driver not only to react faster to the take-over request, but also to improve the take-over quality.

1.2 Vibrotactile Displays in Conditionally Automated Driving

With the introduction of conditionally automated cars, the role of the driver shifts from an active controller to a passive supervisor. That is, the driver does not directly control the car for most of the driving time, resulting in human factors issues, such as a loss of situational awareness [8]. How to get a disengaged operator effectively back into the loop after a prolonged period of passive monitoring is an extensively studied topic in the field of human factors [9].

The introduction of conditionally automated driving has major consequences for the design of displays in the car. Non-driving tasks, such as reading or talking on the phone, are generally visually or auditory demanding. Accordingly, it has been argued that traditional visual and auditory displays might be ineffective [10]. Vibrotactile feedback is considered an attention-capturing feedback modality that is relatively underutilized in driving [10–13]. Accordingly, vibrations could be used to complement auditory and visual warnings in order to provide a redundant warning.

Previous studies have investigated vibrotactile feedback manual driving on the steering wheel (e.g., forward collision warning system) [14], the gas pedal (to encourage

ecological driving) [15], or the driver's seat (e.g., lane departure warning system) [16]. In conditionally automated driving, the possibilities for presenting vibrotactile stimuli are limited, because the driver is not required to touch the steering wheel and pedals anymore. The seat and seatbelt are evident choices for presenting vibrotactile stimuli to a driver. Still, one cannot ignore the possibility that drivers of conditionally automated cars sometimes leave their seat, for example, to grab something from the backseat [17, 18], which could make them miss a signal.

This paper considers that vibrotactile stimuli are comprised of four basic coding dimensions, namely: (1) location, (2) amplitude, (3) timing (on/off pattern), and (4) frequency of the vibrations [19, 20]. Each of the four dimensions can be static (i.e., not changing over time) or dynamic (i.e., changing over time) [7]. Accordingly, a spatially static pattern is a vibration presented at one location, whereas a dynamic pattern changes location as a function of time. The four dimensions can be used to create distinguishable vibration patterns. For example, providing a stimulus at a certain location can provide a directional cue [10, 21]. A spatially dynamic pattern can provide the illusion of a moving stimulus (analogous to the phi-phenomenon [22]), and an increase of amplitude or a reduction of inter-stimulus interval can increase perceived urgency [23]. Location and timing are considered to be more suited for encoding information than frequency and amplitude [20, 24], and see [24] for an overview of tactile perception, vibrotactile technologies, and applications. In this paper, the term 'intensity' is used for a combination of frequency and amplitude of the vibrations.

1.3 Use Cases of a Vibrotactile Seat in Conditionally Automated Driving

In a meta-analysis, Prewett et al. [25] found that for basic reaction time tasks, vibrotactile stimuli elicit faster reaction times than visual stimuli and equally fast reaction times as auditory stimuli [16]. Schwalk et al. [26] showed that, among six spatially dynamic vibration patterns, a pattern that moved from the top of the seat back towards the front of the seat bottom was judged by participants as the most adequate one to convey a take-over request.

During the take-over process, vibrotactile displays could be used not only as a warning, but also to assist the driver in the first two phases (i.e., shift of attention and cognitive processing) of the take-over process. Meng and Spence, for example, argued that in-vehicle vibrotactile stimuli are promising as spatial warning signals [10]. Straughn, Gray, and Tan [27] showed that it is possible to direct the driver's attention towards the left or right by means of a directional vibrotactile warning. A directional take-over request (i.e., a vibration on the left or right side of the seat) could be used to indicate the direction of a potential danger zone, assisting the driver to direct his/her attention and interpret the traffic situation.

It might also be possible to communicate complex messages to the driver by means of vibrations. For example, an approaching car in the blind spot of the driver could be represented by a dynamic pattern. In this case, the location of the car is mapped to a certain vibration location in the seat (e.g., when the car approaches, the vibration moves accordingly). The mapping of a vibrotactile stimulus to a certain location has also been called "tactification" [28], which is analogous to "sonification" for auditory stimuli.

As mentioned above, vibrations can have a looming effect by increasing the frequency or decreasing the inter-pulse interval. Looming vibration patterns could support the driver in making a fast and accurate decision.

1.4 Aim of This Study

This paper presents the design of a vibrotactile display that presents spatially and temporally dynamic vibration patterns in the driver's seat. Moreover, this paper presents a small psychophysics study to evaluate the first prototype of the seat.

2 Design of the Vibrotactile Seat

2.1 Requirements of the Vibrotactile Seat

In order to build a vibrotactile seat that can provide spatially and temporally dynamic vibrotactile stimuli, the following functional requirements need to be met:

- The matrix of the motors should cover most of the contact area between seat and driver.
- The timing and intensity should be independently controllable for each vibration motor.

Additionally, for the purpose of evaluating the display:

- The seat should be compatible with different simulators and test vehicles, and therefore the display needs to be able to communicate with different software environments.
- The vibration motors should be easy to replace.
- The locations of the vibration motors need to be reconfigurable.

2.2 Signal Architecture

The vibrotactile seat is developed for use in simulators and test vehicles. The seat consists of three main parts, namely:

- vibration motors, located in the seat bottom and seat back.
- a control unit that controls the intensity and timing of the individual motors.
- software that determines which vibration pattern has to be present to driver and that communicates this information to the control unit. For example, the software may receive information about upcoming road works and determines that the driver should resume control. The software then communicates the take-over request to the control unit, which activates the motors with the appropriate timing and intensity.

2.3 Hardware

Motors. Eccentric rotating mass (ERM) motors produce vibrations by rotating a mass outside a rotation axis. The rotational speed of the motor is controlled by the voltage to the motor. The rotational speed determines the frequency and amplitude of the vibration, which are therefore not independently controllable. The motors that were used in this prototype were Precision Microdrives™ (type: Pico Vibe, model: 307-100) [29].

Motor Matrix. A series of Velcro strips were used to create a mat, which could be placed in the driver's seat (Fig. 1a). The vibration motors were placed between the Velcro strips so that the motors formed two 6×4 matrices (i.e., 48 motors in total). One matrix was located in the seat bottom and the other in the seat back (Fig. 1b). Accordingly, the configuration of motors can be changed if needed, and the seat mat is interchangeable between simulators or real vehicles.

Fig. 1. (a) Vibration mat in the driver's seat of a simulator. (b) Motor configuration in the seat mat, consisting of 6×4 matrices in the seat bottom and back

Control Unit. The motors are controlled by three Pulse Width Modulation (PWM)-drivers (Texas Instruments, TLC5940NT) [30], which are connected to an Arduino microcontroller (Arduino Mega 2560) [31]. The PWMs control the motors by a series of on/off pulses, which vary the duty cycle (i.e., percentage of time that the signal is on per cycle). The pulses (de)activate the transistor and consequently control the average DC voltage to the vibration motor. The Arduino in turn can be connected to the software environment through a USB or Ethernet connection (Fig. 2a).

Electrical Circuit. The DC-voltage to the motors is controlled by a PWM signal to a transistor (Fig. 2b). The resistor (R1, 100 Ohm) is connected to the base of a PNP transistor (Q1, P2N2907A). The emitter and collector of the transistor are connected to the ground and a motor (M), respectively. A diode (D1, 1N4001) is connected in parallel with the motor to prevent inductive motor spikes flowing back to the transistor. Power to the Arduino, PWM drivers, and motors is provided by an AC/DC converter (Traco-Power, TXL 100-3.3S), which converts 230 V AC from the power network to 3.3 V DC.

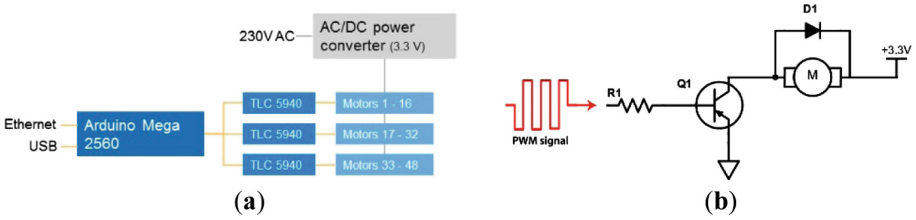


Fig. 2. (a) Functional diagram of the vibrotactile seat. Orange lines represent digital connections, whereas grey lines are power connections. (b) Electrical circuit of a single vibration motor that is PWM-controlled

2.4 Participants

Twelve participants, seven men and five women, between 19 and 31 years old ($M = 24.2$, $SD = 3.1$) took part in the experiment. All participants were students or employees of the Technical University of Munich and had a valid driver's license.

2.5 Apparatus

During the experiment, the participants were sitting in the driver's seat of a fixed-base driving simulator. The simulator featured a driver's seat, steering wheel, and three pedals (gas, brake, and clutch). Three 4 K High Definition screens were placed in front of the driver offering him a viewing angle of approximately 160°. The simulator was not running during the experiment, but used solely to provide a realistic driving position of the participant.

For this study, a prototype of the presented vibrotactile seat above was used (for a study with the presented seat see [32]). This prototype version featured 12 vibration motors. The motors were placed in two straight columns of 6 motors (one column of 6

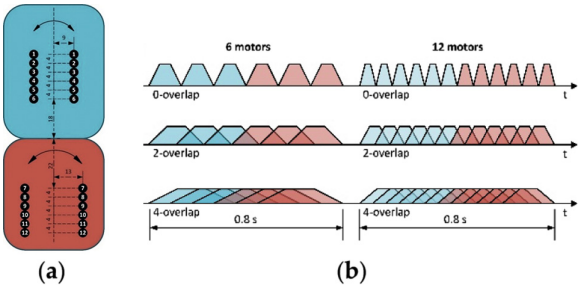


Fig. 3. (a) Dimensions of the 12 × 1 motor matrix on one side of the seat. The seat mat would be turned over to position the motors on the other side for the second trial. The blue area represents the seat back and the red area represents the seat bottom. (b) Activation sequence of the six vibration patterns that were investigated. Three vibration patterns activated six (*left*) or twelve (*right*) motors in sequence, with an overlap of 0, 2, or 4 motors. Blue: motors located in the seat back. Red: motors located in the seat bottom

motors in the seat back, one column of 6 motors in the seat bottom), with a 4 cm inter-motor distance (Fig. 3a). The columns were placed on one side of the seat mat. By turning over the mat, the columns were located on either the left or right side of the seat.

2.6 Procedure

After entering the lab, participants were presented with a consent form, which explained the goal and procedures of the experiment. The study consisted of three parts. The first part was aimed to determine the appropriate vibration intensity for the different motor positions in the seat. The second part investigated the spatial discrimination ability of the participant in the seat back and bottom. The third part concerned the apparent motion of spatially dynamic patterns. Participants were informed that they could leave the experiment at any moment.

Part 1. The PWM duty cycle, and with that the vibration intensity, of a single motor was increased linearly from 0% to 100% over 15 s. The participants were asked to press a button on the steering wheel when they perceived the vibration to be annoying or irritating. When the participant pressed the button or the maximum intensity was reached the vibration would stop.

After a short pause (random between 3 and 6 s), the process was repeated using a different motor. All twelve motors were activated once and the activation order of motors was randomized. Once the participant had received all twelve vibrations, the procedure was repeated but the intensity was inversed. That is, the duty cycle PWM would start at 100% and decrease to 0%. The participants were asked to press the button when the vibration no longer caused any discomfort or irritation.

Part 2. The spatial resolution was investigated using a two-point discrimination test. A pair of motors, both located in either the seat bottom or seat back, was activated simultaneously. The participant was asked to press the button if he/she felt the vibration at two distinct locations. Six motors result in 15 (i.e., $[6 \times 5]/2$) possible motor pairs, which were presented both in the seat back and in the seat bottom (thus 30 pairs in total). All motor pairs where activated once in a random order. The motor pairs were activated at 25% of the duty cycle for 2 s and with an interval of 3 s.

Part 3. Six dynamic vibration patterns were presented to the participant. The patterns differed in the amount of activated motors and in the amount of simultaneously activated motors. A pattern activated either 6 (motor: 1, 3, 5, 7, 9, and 11) or 12 motors (see Fig. 3b) and activated none, two, or four motors simultaneously. All six patterns moved down from the top of the seat back, and forward in the seat bottom. Each vibration pattern lasted 0.8 s.

The intensity of one motor was modulated to increase in a short amount of time, then was constant, and decreased swiftly. That is, the shape of the stimulus was trapezoidal. The three phases (i.e., increase, constant, and decrease) each represented one third of a single vibration time, see Fig. 3b. For an increasing overlap, the stimulated area on the participant increased as well. It has been shown that humans are more sensitive for larger

stimulus areas (i.e., spatial summation [33]). Therefore, the maximum intensity of the vibration was decreased for an increasing overlap. That is, for the 0-, 2-, and 4-overlap the maximum intensity of a single motor was 35, 25, and 20% respectively (illustrated by the height of the trapezia in Fig. 3b). After each vibration pattern the participant was asked to fill out the satisfaction scale of a questionnaire on the acceptance of automotive technology [34]. The scale contained four items on a five-point Likert-scale, ranging from −2 to +2 (1. Unpleasant – Pleasant, 2. Annoying – Nice, 3. Likeable – Irritating, 4. Desirable – Undesirable).

Once a participant had completed the three parts of the experiment, he/she was asked to leave the room. The seat mat was turned over and the trial was repeated with the motors on the other side of the seat. The side (left/right) and order of the three parts were counterbalanced for trials 1 and 2. A single trial took approximately 8 min to complete.

3 Results

The results did not show any significant differences between the left and right side of the seat. Therefore, these results has been averaged, and are not reported separately.

Part I. Figure 4 shows the mean and standard deviations of the vibration intensity (percentage of the PWM-duty cycles) that annoyed the participants. For both the ascending and descending method, the annoyance threshold of second trial was lower than the first ($t(11) = 1.96$, $p = 0.076$ and $t(11) = 2.21$, $p = 0.049$, respectively). There was no statistically significant difference between seat bottom and seat back.

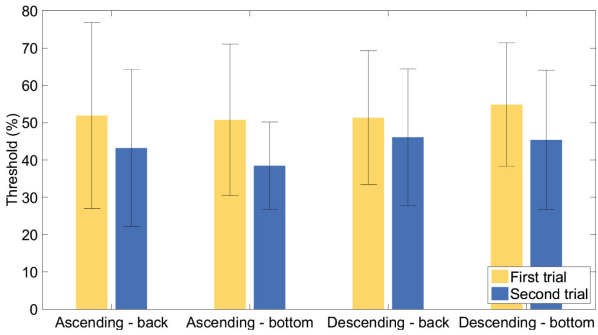


Fig. 4. Means ± standard deviations (N = 12) of the % PWM duty cycles values that annoyed the participants. Ascending = measured when the duty cycle increased; Descending = measured when the duty cycle decreased

Part II. Figure 5 shows the results of the two-point discrimination test. The percentage of correct perceptions (i.e., the participant perceived two separate stimuli) in the seat back increased with inter-motor distance. A two-way ANOVA indicated a significant difference between distances ($F(4, 44) = 30.61$, $p < 0.01$). Post-hoc analysis revealed that all pairs, except 4/8, 12/16, 16/20 cm, were significantly different ($p < 0.05$, after Bonferroni correction).

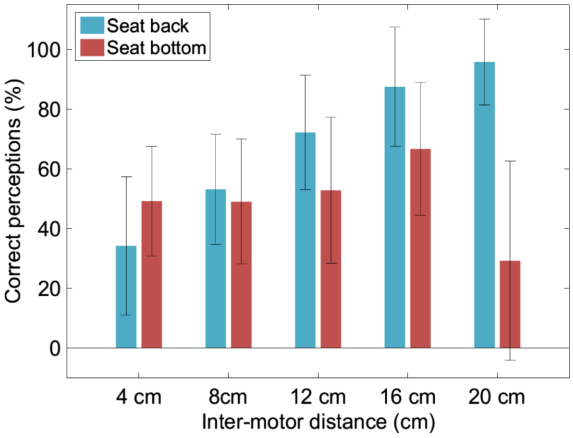


Fig. 5. The mean ± standard deviations (N = 12) of the correct perceptions percentage for two point

The results for the seat back are less clear. The percentages for the 4, 8, 12, 16 cm inter-motor distances are similar, whereas for 20 cm the percentage is considerably lower. A two-way ANOVA indicated there was a significant difference between the groups $F(4,44) = 4.07$, $p = 0.007$). The post-hoc analysis showed that the difference between 16 and 20 cm was statistically significant ($p < 0.05$ after Bonferroni correction).

Part III. Table 1 shows the mean and standard deviation of the satisfaction scores for the 6 dynamic patterns that were presented to the participants. However, there seems to be a trend that the patterns with more overlap received higher satisfaction ratings. A repeated measures ANOVA ($F(5,55) = 2.04$, $p = 0.087$) showed that there were no significant differences between the six patterns.

Table 1. Means and standard deviations (N = 12) of the satisfaction score. Scores were on a 5 point-Likert scale with a range from −2 to +2 as in [34]

	0-overlap	2-overlap	4-overlap
6 motors	0.177 (0.989)	0.542 (0.636)	0.833 (0.611)
12 motors	0.354 (0.732)	0.604 (0.832)	0.615 (0.418)

4 Discussion

In this paper, we presented the design of a tactile display that is able to present vibrotactile stimuli with dynamic location, timing, and intensity in the driver's seat. The seat was developed to investigate the potential to use vibrotactile stimuli in conditionally automated driving. The seat can communicate via standard Ethernet and USB-protocols, making it compatible with various software environments. The Arduino software

enables the control of the individual motors and the seat mat allows for easy modifications of the motor configuration.

The average annoyance threshold across participants was approximately at a 50% duty-cycle, which, according to the motor specifications, resulted in a vibration frequency and amplitude of approximately 100 Hz and 1.3 g, respectively. During the first part of the experiment, single motors were activated. However, to present dynamic patterns, multiple motors will have to be activated simultaneously. Since larger stimulus areas are perceived more intensely [33], designers of vibrotactile displays should adjust the vibration intensity accordingly. Moreover, results show large variance in the annoyance threshold, indicating large differences between participants, despite the fact that the participant group was homogeneous. It is expected that variance will be even higher in a group that is more representative of the general driving population. If a vibrotactile display would be implemented in future vehicles, the intensity may have to be adjustable in order to avoid annoyance or irritation.

The annoyance thresholds seemed to be lower for the second trial as compared to the first. A trial took approximately 8 min, during which participants received many vibrations in quick succession. Thus, vibrations become annoying if people are frequently exposed to them for a prolonged time. Some participants commented after the experiment that they started to dislike the vibrations towards the end of the experiment.

The results in Fig. 5 showed that the percentage of correct perceptions increases as a function of the inter-stimulus location. Ji et al. [35] found similar results and additionally reported that more intense vibrations yielded higher percentages of correct perceptions. When multiple vibrotactile stimuli are presented simultaneously, the distance should be large enough to distinguish separate locations in order to avoid "tactile clutter" [36]. Remarkably, the perception percentage is low for a distance of 20 cm in the seat bottom. This is probably because there were only two motors 20 cm apart in the seat bottom and back respectively. Therefore, the inter-motor distance of 20 cm was presented 4 (out of 60) times to each participant. Alternatively, the result could perhaps be explained by a phenomenon called "apparent location" [20]. This phenomenon occurs when two stimuli are simultaneously presented at different locations, but are perceived as one vibration in one location. Future experiments should investigate if this could potentially be an issue in the design of vibrotactile displays.

The results in Table 1 seemed to indicate that dynamic patterns with more motors simultaneously active (i.e., larger overlap) were rated as more satisfying. However, there were no statistically significant differences between the patterns. A more thorough investigation into dimensions like overlap and waveform should be performed to explore if these can mitigate annoyance.

Note that during the experiment the participants did not actually drive the simulator nor had to monitor a conditionally automated system. Consequently, they could focus all their attention to detecting the vibrotactile stimuli. Future research should investigate how the effectiveness of vibrotactile displays when the driver is engaged in other tasks, like reading a book or playing with a smartphone.

Tactile warnings might be more effective than visual or auditory stimuli, if the driver is engaged in other tasks. Moreover, the vibrotactile seat could perhaps be used to convey

complex information, by directional cues or "tactification" of the stimulus (i.e., the stimulus location is mapped to a physical location in the world). The current design of the seat allows further investigation of such dynamic vibration patterns.

Research is needed on the implementation of multimodal displays in a conditionally automated car. Future experiments should investigate how to complement visual and auditory displays with vibrotactile stimuli, aiming to effectively and safely assisting the driver in the take-over process. Nevertheless, the results of this experiment showed that annoyance could be an issue when vibrations are presented to driver. Interface designers should aim to mitigate annoyance by, for example, making the intensity levels adjustable.

Acknowledgements. Sebastiaan M. Petermeijer, Joost C. F. de Winter, and Klaus J. Bengler are involved in the European Marie Curie ITN project HFauto – Human Factors of Automated Driving (PITN-GA-2013-605817).

References

1. Begg, D.: A 2050 vision for London: what are the implications of driverless transport (2014). http://www.transporttimes.co.uk/Admin/uploads/64165-Transport-Times_A-2050-Vision-for-London_AWWEB-READY.pdf
2. ERTRAC Task Force: Automated driving roadmap. http://www.ertrac.org/uploads/documentsearch/id38/ERTRAC_Automated-Driving-2015.pdf
3. SAE International: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (2014)
4. Kerschbaum, P., Lorenz, L., Bengler, K.: A transforming steering wheel for highly automated cars. In: IEEE Intelligent Vehicle Symposium (IV), pp. 1287–1292 (2015)
5. Gold, C., Damböck, D., Lorenz, L., Bengler, K.: "Take over!" How long does it take to get the driver back into the loop? In: 57th Annual Meeting of the Human Factors and Ergonomics Society, pp. 1938–1942 (2013)
6. Zeeb, K., Buchner, A., Schrauf, M.: What determines the take-over time? An integrated model approach of driver take-over after automated driving. Acc. Anal. Prev. **78**, 212–221 (2015)
7. Petermeijer, S.M., De Winter, J.C.F., Bengler, K.J.: Vibrotactile displays: a survey with a view on highly automated driving. IEEE Trans. Intell. Transp. Syst. **17**, 897–907 (2016)
8. Endsley, M., Kiris, E.: The out-of-the-loop performance problem and level of control in automation. Hum. Fact. J. Hum. Fact. Erg. Soc. **37**, 381–394 (1995)
9. Bainbridge, L.: Ironies of automation. Automatica **19**, 775–779 (1983)
10. Meng, F., Spence, C.: Tactile warning signals for in-vehicle systems. Acc. Anal. Prev. **75**, 333–346 (2015)
11. Sarter, N.B.: The need for multisensory interfaces in support of effective attention allocation in highly dynamic event-driven domains: the case of cockpit automation. Int. J. Aviat. Psychol. **10**, 231–245 (2000)
12. Mohebbi, R., Gray, R., Tan, H.Z., Lafayette, W.: Driver reaction time to tactile and auditory rear-end collision warnings while talking on a cell phone. Hum. Fact. J. Hum. Fact. Erg. Soc. **51**, 102–110 (2009)
13. Gray, R., Ho, C., Spence, C.: A comparison of different informative vibrotactile forward collision warnings: does the warning need to be linked to the collision event? PLoS ONE **9**, e87070 (2014)

14. Chun, J., Han, S.H., Park, G., Seo, J., Choi, S., Lee, I.: Evaluation of vibrotactile feedback for forward collision warning on the steering wheel and seatbelt. Int. J. Ind. Ergon. **42**, 443–448 (2012)
15. Radlmayer, J., Selzer, T., Arcati, A., Bengler, K.: Haptic gear shifting indication: evaluating acceptance and potential fuel consumption. In: 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015), vol. 3, pp. 2746–2752 (2015)
16. Navarro, J., Mars, F., Forzy, J.-F., El-Jaafari, M., Hoc, J.-M.: Objective and subjective evaluation of motor priming and warning systems applied to lateral control assistance. Acc. Anal. Prev. **42**, 904–912 (2010)
17. Llaneras, R.E., Salinger, J., Green, C.A.: Human factors issues associated with limited ability autonomous driving systems: drivers' allocation of visual attention to the forward roadway. In: 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, pp. 92–98 (2013)
18. Omae, M., Hashimoto, N., Sugamoto, T., Shimizu, H.: Measurement of driver's reaction time to failure of steering controller during automatic driving. Rev. Autom. Eng. **26**, 213–215 (2005)
19. Geldard, F.A.: Some neglected possibilities of communication. Science **131**, 1583–1588 (1960)
20. Van Erp, J.B.F.: Guidelines for the use of vibro-tactile displays in human computer interaction. In: Eurohaptics, pp. 18–22 (2002)
21. Ho, C., Tan, H.Z., Spence, C.: Using spatial vibrotactile cues to direct visual attention in driving scenes. Transp. Res. F. Traf. Psych. Beh. **8**, 397–412 (2005)
22. Wertheimer, M.: Experimentelle studien über das sehen von bewegung. Zschr. F. Psych. Phys. Sinnesorgane **61**, 161–265 (1912)
23. Van Erp, J.B.F., Toet, A., Janssen, J.B.: Uni-, bi-, tri-modal warning signals: effects of temporal parameters and sensory modality on perceived urgency. Safe. Sci. **72**, 1–8 (2015)
24. Jones, L.A., Sarter, N.B.: Tactile displays: guidance for their design and application. Hum. Fact. J. Hum. Fact. Erg. Soc. **50**, 90–111 (2008)
25. Prewett, M.S., Elliott, L.R., Walvoord, A.G., Coovert, M.D.: A meta-analysis of vibrotactile and visual information displays for improving task performance. IEEE Trans. Syst. Man Cybern. C Appl. Rev. **42**, 123–132 (2012)
26. Schwalk, M., Kalogerakis, N., Maier, T.: Driver support by a vibrotactile seat matrix – recognition, adequacy and workload or tactile patterns in take-over scenarios during automated driving. In: 6th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences, pp. 1427–1434 (2015)
27. Straughn, S.M., Gray, R., Tan, H.Z.: To go or not to go: Stimulus response compatibility for tactile and auditory pedestrian collision warnings. IEEE Trans. Hap. **2**, 111–117 (2009)
28. Burns, C.M., Ho, G., Arrabito, G.R.: Mapping ecologically to modalities. In: 55th Annual Meeting of the Human Factors and Ergonomics Society, pp. 335–339 (2011)
29. Precision MicrodrivesTM: Pico Vibe 9 mm vibration motor – 25 mm type. Model number: 307–103. https://catalog.precisionmicrodrives.com/order-parts/product/307-103-9mm-vibration-motor-25mm-type
30. Texas Instruments© - TLC5940NT – 16 channel LED driver w/EEprom DOT correction & grayscale PWM control. http://www.ti.com/product/TLC5940
31. Arduino & Genuino Products© – Arduino Mega 2560. https://www.arduino.cc/en/Main/ArduinoBoardMega2560
32. Petermeijer, S.M., Cieler, S., De Winter, J.C.F.: Comparing spatially static and dynamic vibrotactile take-over requests in the driver seat. Acc. Anal. Prev. **99**, 218–227 (2017)

33. Geschneider, G., Bolanowski, S., Verrillo, R.: Some characteristics of tactile channels. Beh. Brain Res. **148**, 35–40 (2004)
34. Van der Laan, J., Heino, A., De Waard, D.: A simple procedure for the assessment of acceptance of advanced transport telematics. Transp. Res. C Emerg. Tech. **5**, 1–10 (1997)
35. Ji, Y.G., Lee, K., Hwang, W.: Haptic perceptions in the vehicle seat. Hum. Fact. Erg. Manuf. Serv. Ind. **21**, 305–325 (2011)
36. Van Erp, J.B.F., Veltman, J.A., Van Veen, H.A.H.C., Oving, A.B.: Tactile torso display as countermeasure to reduce night vision goggles induced drift. In: Symposium on Spatial Disorientation in Military Vehicles: Causes, Consequences and Cues, pp. 15–17 (2002)

Learning from the Best – Naturalistic Arbitration for Cooperative Driving

Gina Weßel[1(✉)], Constanze Schreck[1], Eugen Altendorf[1], Yigiterkut Canpolat[1], and Frank Flemisch[1,2]

[1] Institute of Industrial Engineering and Ergonomics, RWTH Aachen University, Bergdriesch 27, 52062 Aachen, Germany
{g.wessel, c.schreck, e.altendorf, y.canpolat, f.flemisch}@iaw.rwth-aachen.de
[2] Fraunhofer Institute for Communication, Information Processing and Ergonomics, Wachtberg, Germany

Abstract. In cooperative automated driving, the task of lateral and longitudinal vehicle control can be shared by driver and automation. However, conflicting action intentions of the two partners could arise, which need to be resolved within limited time. This can be achieved through structured multimodal negotiation, called arbitration. In order to explore intuitive interaction patterns for arbitration situations, insights from human-human interaction might be transferred. Accordingly, in a field study, couples holding hands or walking arm in arm were videotaped and interviewed when a conflict concerning motion control has been observed. The analysis of the data shows that conflict situations concerning velocity and/or direction of movement occur in natural human-human interaction and that these types of conflict can be dependent on each other. Furthermore, partners use different interaction resources to successfully solve these situations. Results are transferred to cooperative automated driving and an example of an interaction pattern is presented.

Keywords: Arbitration · Human-machine cooperation · Human systems integration · Automated driving · Cooperative movement · Cooperative guidance and control

1 Introduction and Theoretical Background

Increasingly more domains of life can be automated due to technological progress, but interaction between human and automation can be complex (e.g. [1, 2]). For example, in automated driving, the automation and the human driver can share the tasks of lateral and longitudinal guidance and control. Research indicates that the interaction between driver and automation can be improved by a cooperative automation, which supports the human as an intelligent and cooperative team member [3, 4]. A natural example for this type of cooperation is the relationship of a rider and a horse, which has been explored as a design metaphor [5, 6]. Another example is the teaming of humans that was explored using the design metaphor of a copilot [7].

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_61

However, among intelligent partners different action intentions can arise. In cooperative driving there might be a conflict about direction or speed [8]. Through systematic arbitration, these conflicts can be solved. According to [9], arbitration denotes "a structured negotiation between human and automation with the intention to reach a common and unambiguous […] decision on how to act in due course of time."[1] This concept of arbitration is illustrated by Fig. 1 (*left*), which displays the cooperation between driver and automation in cooperative vehicle guidance and control. The *lower* part shows that human and automation both receive and process information from the environment. Based on the processing results and their subsequent action intentions the partners interact with each other and arbitrate conflicts, so that their intentions can be integrated into a mutually accepted decision, which affects the entire vehicle system [3]. On the *right* side of Fig. 1, an arbitration situation is depicted, in which driver and automation have different intentions. In this case, the driver rapidly approaches a lead vehicle, while the automation intends to reduce speed.

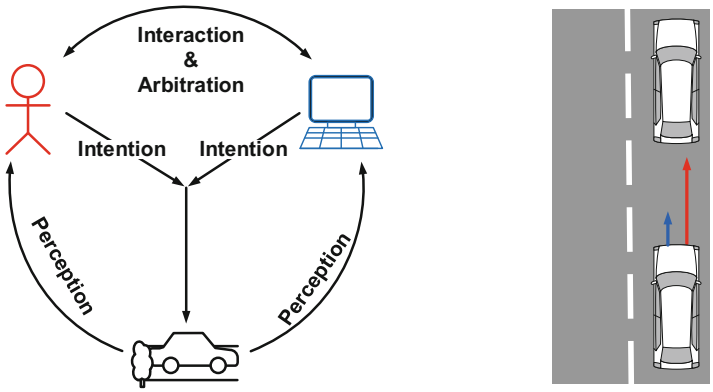


Fig. 1. *Left*: Interaction adapted from [3]. *Right*: Example arbitration situation: automation intends to reduce velocity, driver intends to keep current velocity.

Arbitration for solving this conflict can take place through interaction patterns. Within the present example, the following, escalating interaction pattern could be applied: When the lead vehicle is approached, the automation proposes to slow down the ego vehicle in order to avoid a crash. This can be communicated to the driver by a haptic signal (e.g. a repulsing force on the acceleration pedal) followed by a visual signal in the head-up display (e.g. highlighting the lead vehicle) followed by an acoustic warning and ultimately, by a braking intervention. This escalating sequence of signals and actions is influenced by the driver's actions. If she reduces speed or evades to the other lane, the escalation is stopped.

[1] The definition provided by [9] is written in German. The presented quotation is a translation by the authors of this paper.

This interaction pattern addresses different modalities: haptic, acoustic and visual, which is in line with the proposition by [8] to design multimodal interaction patterns for arbitration situations. While the benefits of multi-modal interaction have been demonstrated by [10, 11], the search for intuitive patterns for arbitration situations continues. [12] provide the following definition of intuition in human-machine interaction "A technical system is intuitively usable if the users'unconscious application of prior knowledge leads to effective interaction".

In order to find intuitive multi-modal interaction patterns for solving arbitration situations, inspiration could come from cooperative human movement. This general approach of deriving information from human-human cooperation for human-machine cooperation has been deployed in research before. E.g. [13] transferred the situation of two humans jointly carrying a heavy object to the situation of driver and vehicle automation cooperating on the driving task. There are several domains, in which humans are interacting on motion control in pairs with a haptic coupling and insights might therefore be transferrable to cooperative driving. Examples are walking with held hands, ballroom dancing, martial arts, carrying objects together etc.

For verifying this proposition, the most common and natural setting of the above-mentioned situations of cooperative movement was chosen – walking with held hands. It was assumed that this delivers most universal results, because walking with held hands is practiced by more different individuals than e.g. dancing. Additionally, it can be assumed that no clear rules for interaction have been deliberately taught, like this might be the case in e.g. martial arts. To this end, a field study was conducted. Couples walking with a haptic connection (holding hands or walking arm in arm) were observed in order to document situations in which different intentions about movement arose and got solved. These conflicts could concern either direction or velocity of movement. It was investigated, whether partners used different modalities for interaction and whether conflicts were solved. Methodology and results will be described in the following. As a next step, use cases of cooperative driving will be proposed, which display similarities to the observed situations and an exemplary interaction pattern will be described.

2 Methodology

In 2016, a field study was conducted at the Christmas Market in Aachen, Germany between November 29 and December 14. Couples holding hands or walking arm in arm were filmed when walking around the stalls. Mostly, they were videotaped from the back in order not to disturb their natural interaction. The Christmas Market was chosen as a setting for several reasons. Firstly, it was assumed that different intentions might easily arise because of the different interests targeted by different stalls and a Christmas Market is an event that is visited by many people. Accordingly, it was assumed that a sufficient amount of the rather specific target situations could be observed. Secondly, it was assumed that couples on a Christmas Market are commonly strolling and then spontaneously stopping to have a look at interesting stall. This mindset may enhance the chances that participants agree on being interviewed. Thirdly,

it was assumed that a rather heterogeneous sample could be observed due to the diverse target group a Christmas Market addresses.

Every time a conflict situation was observed (e.g. one partner displaying the intention to walk at a different speed or in a different direction), the recording was saved and participants were approached for an interview. However, situations in which the partners could not solve the conflict together and dissolved the connection of the hands/arms were excluded and video recordings were deleted, as the focus of this study lies on successful arbitration in which an integrated solution is found for the entire system.

The interview consisted of an explanation of the background of the study and the question for consent of using the video recording for research purposes. In case no permission for further use of the collected data was given, the recordings were deleted immediately. Furthermore, participants were asked to reproduce the situation of the conflict and describe the modalities they used for communicating their intentions.

For evaluating the results, the notes from the interview were analyzed by three trained raters on whether participants were describing different action intentions.

The video recordings were evaluated concerning whether a conflict about direction and/or velocity of movement was displayed. A conflict concerning direction was operationalized as the feet of one partner pointing into a different direction than the feet of the other partner or a different direction of movement of one partner (which does not necessarily involve different position of feet, as humans could walk sideways). Furthermore, a changing lateral distance between the partners could indicate a conflict about direction of movement. A conflict about velocity was operationalized as a longitudinal drift between the partners. Moreover, the video recordings were analyzed on whether the conflict was solved. This was executed through rating whether the partners realigned their movement, either by agreeing on a movement direction and velocity without parting or by both coming to a standstill without parting. Finally, recordings and interviews were assessed for which interaction resources have been used.

3 Results

The analysis included twelve video recordings of couples composed of a male and a female participant (see Fig. 2 for example snapshots from the video recordings). This homogeneity of gender was not intended. Couples of the same gender were observed as well, but they did not meet the criteria for being included in the study or refused to participate. The results of the analysis will be described in detail.

3.1 Action Intentions

Firstly, it was evaluated whether the participants mention different intentions of actions regarding motion control in the interview. As one couple refused to be interviewed, only 11 interviews were analyzed. For this and all subsequent ratings, Fleiss' kappa

Fig. 2. Snapshots from video recordings as examples of a conflict of direction (*left*) and velocity (*right*).

was calculated using MATLAB [14, 15]. Concerning different action intentions, the three raters completely agreed on all 11 interviews. Accordingly, Fleiss kappa equaled 1 (p = 0.000), denoting perfect agreement [16]. Kappa values, confidence intervals and p-values for each test are given in Table 1.

Table 1. Inter-rater reliability. The answering format on all questions was Yes/No. Three raters analyzed all video recordings.

Question	Fleiss' kappa	95% CI lower bound	95%CI upper bound	p-value
Different action intentions?	1	0.9112	1.0888	0.0000
Conflict on direction?	0.7187	0.6338	0.8037	0.0000
Conflict on velocity?	0.6073	0.5223	0.6923	0.0003

In 9 of 11 interviews different action intentions were described. For example, one partner wanted to a have a look at a stall and therefore wanted to move closer to it or reduce speed in several situations, or there was a conflict about which way to go at the end of a row of stalls. This supports the assumption that different intentions can arise in a natural context of two persons moving together with a haptic coupling.

3.2 Conflict Concerning Direction of Movement

Furthermore, it was assessed, whether the video recording demonstrated a conflict concerning the direction of movement (assessed by positioning of the feet, direction of movement of the individual partners and changing lateral distance between them). The overall kappa was 0.7187 (p = 0.0000) (See Table 1). According to [16] this denotes substantial agreement.

Raters agreed that in 10 of 12 video recordings, there was a conflict concerning direction of movement and in one recording there was not. In one recording, raters did not agree on whether there was a conflict. This was due to the turning of the upper body of the person in the video recordings. While her feet and direction of movement were in line with that of her partner, her whole upper body was turned to another side (see Fig. 3 *left*). This behavior could be seen as a divergent intention, but was not included in the operational definition. Moreover, when couples were standing in front of a stall they might be turning their feet in different positions without intending to move in that direction. Using these insights, operational definitions might be further refined for a following study. However, the results indicate that conflicts concerning directions of movement can arise in a natural context.

Fig. 3. *Left:* Feet and direction of movement of the two partners are aligned, but the woman is turning her upper body to the left. *Middle*: Feet of the partners are not aligned, however this does not seem to reflect different intentions of movement (as indicated by the video recording and the interview). *Right*: The couple just turned right, causing a necessity for longitudinal realignment.

3.3 Conflict Concerning Velocity of Movement

In the next step, it was evaluated whether a conflict concerning speed was displayed in the video recordings (assessed by longitudinal offset between persons). The overall kappa was 0.6073 ($p = 0.0003$) (See Table 1). This denotes substantial agreement between raters [16]. The raters agreed that in seven recordings, there was a conflict concerning speed and in two recordings, there was none. This indicates that in a natural context, observable situations arise in which there is a conflict of speed.

In the three video recordings, which raters were undecided about, one of two situations occurred. The first situation has already been mentioned before: the upper body of one partner was turned to a stall and away from the direction of movement. Accordingly, this could be interpreted as an offset between the persons depending on which part of the body is focused on for the evaluation (see Fig. 3 *left*). In the second situation, the couple was turning to one side, which subsequently evoked the necessity for one partner to catch up with the other (see Fig. 3 *right*).

This situation illustrates the point that when a video recording is rated as displaying both types of conflict (velocity and direction), these do not necessarily occur at the same time. One conflict can be a result of the other or different conflicts simply occurred at different points in time during a couple's walk on the market. In Sect. 4, this insight will be transferred to cooperative driving.

3.4 Interaction Resources

Lastly, it was analyzed which interaction resources have been used by the participants with focusing on visual, vocal and haptic interaction. Since most couples were filmed from the back and it cannot be assumed that there was a perfect recall of interaction resources and their sequence in the interviews, no quantitative analysis has been conducted. In most video recordings, forces are applied to the haptic connection of hands or arms, sometimes leading to a change in direction or speed. In some recordings, it can also be seen that participants were using verbal communication. Furthermore, pointing gestures can be observed in some recordings, indicating that one of the partners tries to direct the visual attention of his counterpart to a point of interest. Not all couples used all interaction resources. Regularly, there was haptic or haptic-vocal interaction.

4 Transfer to Automated Vehicles and Cooperative Driving

The analysis has shown that there are conflicts concerning lateral and longitudinal movement occurring between human partners in a natural setting. This generally enables transfer of these situations, interaction between the partners and their solutions to conflicts in cooperative automated driving. Accordingly, the situations in the video recordings were investigated more closely. They will be described exemplarily and comparable use cases of cooperative driving will be proposed. Furthermore, an example of an interaction pattern will be presented.

4.1 Use Cases of Cooperative Driving

At the end of an alley, a couple had to decide to either turn right or left. This can be directly transferred to an use case involving a (T-)junction (see Fig. 4 *left*). For example, the automation might be pursuing the fastest route to a destination, but the driver decides to turn right because she prefers this way. However, the automation intends to stay on the initial route, i.e. to turn left. Here, both cooperation partners are deciding on lateral guidance.

The conflict in the second use case (see Fig. 4 *middle)* occurs at the lateral and longitudinal level. In the situation of the study, one partner becomes interested in a stall and initiates a deviation from the straight line of walking. This can be transferred to the situation in which the automation is aiming to reach a destination, but the driver wants to take a motorway exit e.g. for getting a coffee.

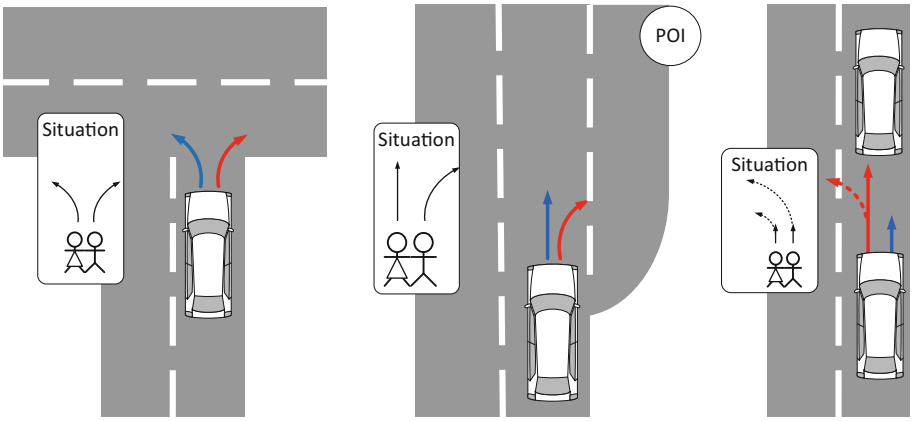


Fig. 4. Use Cases. (T-)junction (*left*), point of interest (*middle*), dependence of conflicts (*right*).

The last use case is derived from the situation, where one partner initiates a turn to one side, which also causes a conflict on velocity (see Fig. 4 *right*). This dependence of the two types of conflict can be transferred to cooperative driving as follows: A driver might want to initiate an overtaking maneuver in the near future (e.g. when the left lane is available), while the automation intends to stay on the right lane and adapting speed according to traffic situations. This denotes a lateral conflict. When in this situation a slowly driving lead vehicle is approached, the automation might want to reduce speed, which is conflicting with the driver's intention to keep speed in order to change lanes in the next moment. This use case demonstrates that the dependence of the two types of conflict can be transferred. Furthermore, it implies that arbitration situations should be analyzed in detail on whether there really is only one type of conflict.

4.2 Exemplary Interaction Pattern for Arbitration

The interaction pattern of a jointly moving couple facing the use case presented in the *right* part of Fig. 4 will be described and transferred to cooperative automated driving. This pattern is depicted in Fig. 5: Firstly, both partners walk and look in the same direction. Then they turn their heads toward a point of interest (POI). While the male partner keeps looking in that direction, the female partner focusses her eyes back to the walking direction. This might be the first step towards an evolving conflict, as attention of both partners is not focused in the same direction. In the next sequence, the male partner starts walking to the right. This causes a pulling motion on the connected hands. In reaction to that the partners look at each other and the male partner says something. Finally, they both look at the POI and move toward it. Due to the different actions of the partners concerning different modalities, the conflict evolves over time and finally gets solved. In Fig. 5, the levels of the escalation are color coded.

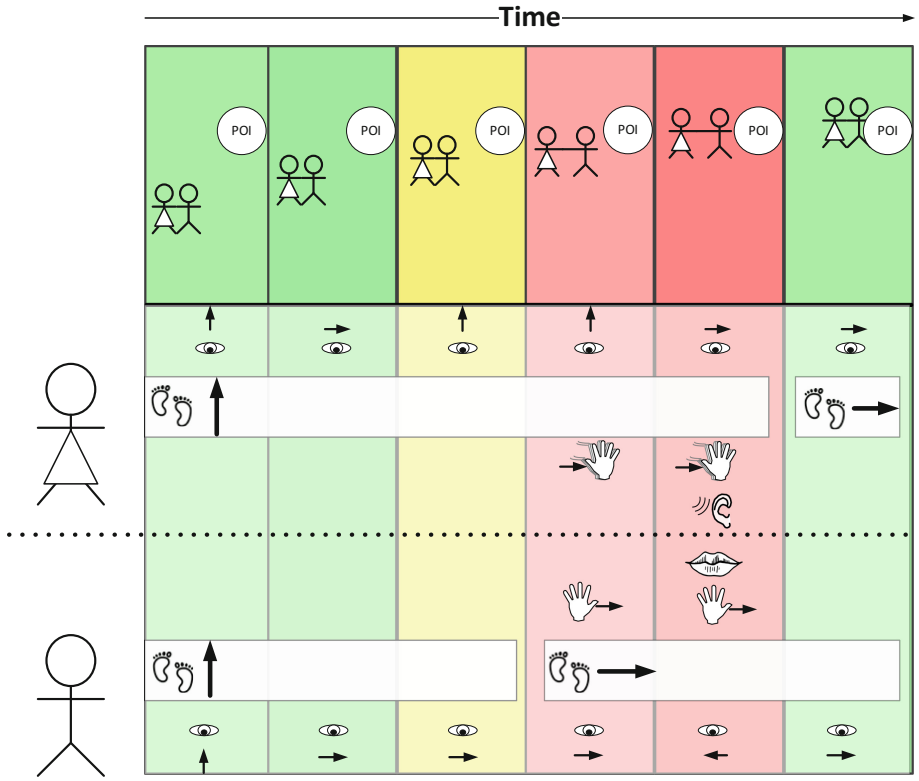


Fig. 5. Interaction pattern of a couple facing conflicting intentions for motion control. *Upper part:* development of the situation over time; *Lower part:* actions of the participants.

This pattern can be transferred to cooperative automated driving (see Fig. 6), if one partner is replaced by an automation. As driver and automation move together within one vehicle, the feet from Fig. 5 have been replaced by a steering wheel for indicating motion control. Moreover, as the automation has no focus of attention, the eyes on the *upper* side of the figure have been replaced by the goal the automation pursues. In the case of the first columns this equals reaching a specified destination).

The transfer of the pattern yielded the following results: In the beginning, driver and automation pursue the same goal of reaching a destination, which is in line with the driver's focus of attention. However, then he notices the POI, which might be a café or a sign for a café. As he decides to have a coffee, he turns the steering wheel to the right. Using drive-by wire technology, the vehicle dynamics would not necessarily be directly influenced by the inputs on the steering wheel. Therefore, the interaction could be designed accordingly and only if a lane change maneuver is intended it would be conducted. Otherwise, the steering wheel could have been touched by accident or the driver might have wanted to change the offset from the lane only. Hence, a dialog targeting a different interaction resource is executed to confirm that the corresponding lane change maneuver is intended. This parallels the auditory component of the

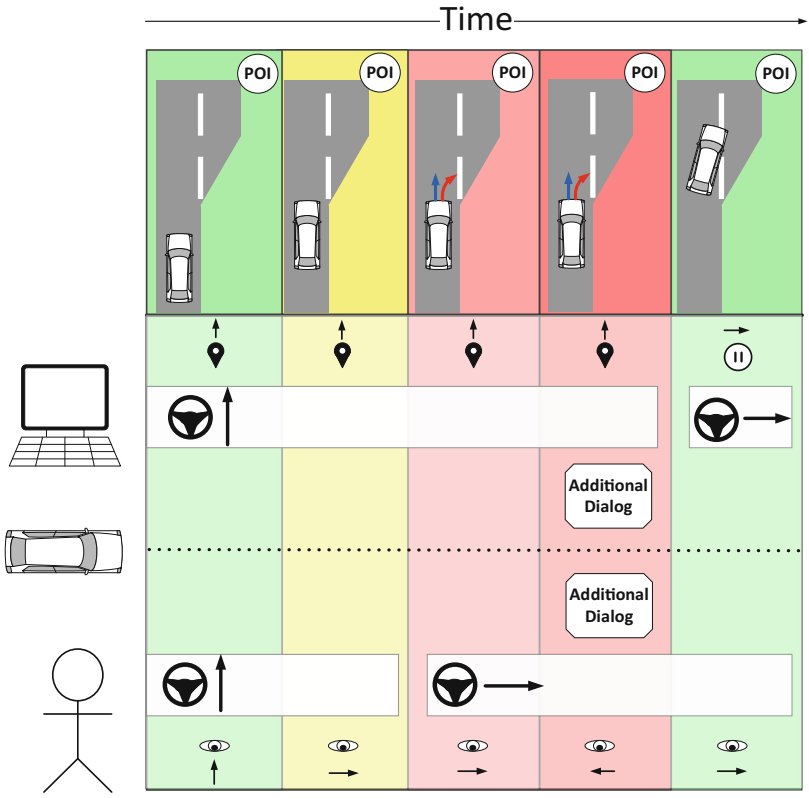


Fig. 6. Interaction pattern for cooperative automated driving.

human-human interaction pattern. After the change of direction has been confirmed, the automation changes the current goal to "break". Because both partners now agree on the new direction, the vehicle approaches the POI.

5 Discussion and Conclusion

In cooperative human-machine interaction, conflicts may arise, which could be solved by interaction patterns. The present study demonstrated that arbitration situations occur in the natural context of two humans moving together in a leisure context with haptic coupling (holding hands or walking arm in arm). For this purpose, a video analysis of 12 video recordings was conducted. It was shown that conflicts occur naturally concerning speed as well as direction of movement. Despite most video recordings being rated as displaying lateral and longitudinal conflict, these two types of conflict did not always occur at the same time, but one type of conflict can be caused by the other. For example, when a couple turns right the partner walking left needs to catch up with the other. While it might be discussed, whether that expresses a conflict concerning speed

of movement per se, it was evaluated as such according to the operational definitions. However, it could be argued that this is legitimate, because there really was an offset, which needed to be solved by either one partner waiting for the other, or the other accelerating to catch up.

For communicating their own intentions to their partner, participants used different interaction resources: haptic signs (e.g. pulling or pushing with hands, arms or upper body), verbal communication and directing visual attention by pointing to or touching objects. Accordingly, it can be assumed that it is intuitive to use all interaction resources for solving conflicts in cooperative driving as proposed by [8]. However, the results indicate that not all resources are used in all situations. It was found that participants sometimes only used haptic or haptic-vocal communication. In line with escalating interaction schemes, in a natural setting people might only use all three resources when the situation is critical or urgent.

However, when analyzing the interviews and investigating the corresponding video recordings, it was noticed that couples sometimes had difficulties to correctly reproduce their use of interaction resources. This indicates that some interaction might be executed subconsciously. Some couples even denied the existence of a conflict, but nevertheless reported different intentions or pulling actions. This emphasizes the importance of using behavioral data for gaining insight in this field of research.

In this study, only couples that did not give up their haptic connection have been approached for an interview. This approach was chosen in order to focus on successful arbitration situations. In future studies failed arbitration could also be investigated to separate effective from non-effective interaction patterns. However, the results of this study indicate that naturally occurring conflicting situations can be solved naturally through multimodal interaction.

These findings have been transferred to cooperative driving. Analogous use cases to the filmed situations have been proposed and it was demonstrated that the dependence of the types of conflict can be transferred to cooperative driving. Moreover, a first interaction pattern has been reported and transferred. Accordingly, this study demonstrated that there is a high potential of transferring insights from human—human to human-machine cooperation.

Acknowledgments. The research conducted was partly funded by the Deutsche Forschungsgemeinschaft (DFG) within the project "Arbitration of cooperative movement for highly-automated human machine systems" and partly by RWTH Aachen University.

References

1. Hoc, J.-M.: From human-machine interaction to human-machine cooperation. Ergonomics **43**, 833–843 (2000)
2. Pacaux-Lemoine, M.P., Debernard, S.: Common work space for human-machine cooperation in air traffic control. Control Eng. Pract. **10**, 571–576 (2002)
3. Flemisch, F.O., Bengler, K., Bubb, H., Winner, H., Bruder, R.: Towards cooperative guidance and control of highly automated vehicles: H-Mode and Conduct-by-Wire. Ergonomics **57**, 343–360 (2014)

4. Flemisch, F., Kelsch, J., Löper, C., Schieben, A., Schindler, J., Heesen, M.: Cooperative control and active interfaces for vehicle assitsance and automation. In: FISITA World Automotive Congress, München. München (2008)
5. Flemisch, F.O., Adams, C.A., Conway, S.R., Goodrich, K.H., Palmer, M.T., Schutte, P.C.: The H-Metaphor as a guideline for vehicle automation and interaction (2003)
6. Altendorf, E., Baltzer, M., Kienle, M., Meier, S., Weißgerber, T., Heesen, M., Flemisch, F.: H-Mode 2D. In: Handbuch Fahrerassistenzsysteme, pp. 1123–1138. Springer (2015)
7. Holzmann, F., Flemisch, F.O., Siegwart, R., Bubb, H.: From aviation down to vehicles-integration of a motions-envelope as safety technology (2006)
8. Baltzer, M., Flemisch, F., Altendorf, E., Meier, S.: Mediating the interaction between human and automation during the arbitration processes in cooperative guidance and control of highly automated vehicles. In: Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE (2014)
9. Kelsch, J., Flemisch, F., Löper, C., Schieben, A., Schindler, J.: Links oder rechts, schneller oder langsamer? Grundlegende Fragestellungen beim Cognitive Systems Engineering von hochautomatisierter Fahrzeugführung. In: Cognitive Systems Engineering i.d Fahrzeug- und Prozessführung; 48. FAS Anthropotechnik. Karlsruhe (2006)
10. Oviatt, S.: Advances in robust multimodal interface design. IEEE Comput. Gr. Appl. **23**, 62–68 (2003)
11. Dumas, B., Lalanne, D., Oviatt, S.: Multimodal interfaces: a survey of principles, models and frameworks. In: Human machine interaction, pp. 3–26. Springer (2009)
12. Mohs, C., Hurtienne, J., Israel, J.H., Naumann, A., Kindsmüller, M.C., Meyer, H.A., Pohlmeyer, A.: IUUI – intuitive use of user interfaces. Usabil. Professionals **6**, 130–133 (2006)
13. Flemisch, F., Abbink, D., Itoh, M., Pacaux-Lemoine, M.-P., WEßEL, G.: Shared control is the sharp end of cooperation: towards a common framework of joint action, shared control and human machine cooperation. IFAC-PapersOnLine 49, 72–77 (2016)
14. Fleiss, J.L.: Measuring nominal scale agreement among many raters. Psychol. Bull. **76**, 378 (1971)
15. Cardillo, G.: Fleiss' es kappa: compute the Fleiss' es kappa for multiple raters. MATLAB MathWorks (2007). http://www.mathworks.com/matlabcentral/fileexchange/15426
16. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. Biometrics **33**(1), 159–174 (1977)

Modelling the Dynamics of Driver Situation Awareness in Automated Driving

Zhenji Lu(✉), Riender Happee, and Joost de Winter

BioMechanical Engineering Department, Mechanical, Maritime and Materials Engineering Faculty, Delft University of Technology, Delft, The Netherlands
Z.lu@tudelft.nl

Abstract. This study presents a numerical model that describes the dynamic process of building situation awareness after an automation-initiated transition. The model predicts the level of situation awareness as a function of elapsed time since the transition, and is verified using data from an experiment in which participants watched animated video clips of automated driving scenarios. Additionally, the 'number of fixations per second' is suggested for real-time monitoring of situation awareness in automated driving.

Keywords: Automated driving · Situation awareness · Monitoring/control transition · Visual attention

1 Introduction

Automated driving systems are being developed with the aim to reduce driver workload and to increase safety. However, these systems may create dangerous situations if the technology fails while the driver is not attentive to the driving task. The current solution to such dangers is to try to redirect the driver's attention in order to get the driver back into the loop. Indeed, as pointed out by Bainbridge (1983) [1], to 'take over control' from automation is a primary task for the human supervisor of automation.

How to safely transfer control from the automation to a (situationally unaware) driver is a topic that has drawn much attention from human factors researchers (see Lu et al. [2], for a review). Human factors researchers have focused on improving the human machine interface (HMI) and on creating driver monitoring systems for real-time feedback or adaptive automation.

Lu et al. [2] distinguished between monitoring transitions and control transitions. A control transition refers to a transition that involves a reallocation of the longitudinal and/or lateral control task between the driver and the automation, whereas a monitoring transition involves the driver's reallocation of attention between the driving task and a non-driving task. Because being in control usually requires one to monitor the environment [3], a control transition usually coincides with a monitoring transition [2]. In this paper, we are not concerned with the control task itself, and we use the term 'transition' to represent both a control transition and a monitoring transition. We also assume that an HMI produces a warning that signals the driver that he/she has to redirect attention from an out-of-the-loop situation back to the driving task.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_62

Research in the aviation domain has found that pilots sometimes make wrong decisions with hazardous consequences [4, 5]. Situation awareness (SA) is regarded as one of the fundamental precursors of decision making [3]. A computational model to assess the operators' SA during transitions may be an asset for refining the design of automated driving systems and HMIs.

This paper aims to model the dynamic process of building SA as a function of time after a transition. We assume a driver who allocates his visual attention either to the traffic environment or to a non-driving task. Additionally, in this paper, we establish links between the process of building SA and eye movements, in order to offer new ideas for the development of driver monitoring systems.

2 Dynamics of Building Situation Awareness

The topic of SA has been studied extensively in the last decades [6] and may be of key importance for decision making in time-critical situations such as transitions in automated driving. In other words, a quantification of the level of SA could be used to derive probabilities of human decisions. In the following section, we describe the proposed dynamic model of SA.

2.1 Model Architecture

To build the SA model, several elements need to be represented: (1) the information density of the environment, (2) the human's attentional resources, and (3) the time-dependency of SA. In our model, we assume that the environment (information density) is constant. The model is a convergent function that describes the level of driver SA of the environment as a function of elapsed time since the transition.

$$S(t, t_1) = A \frac{E_{\max}}{E_{required,t_1}} \left(1 - e^{-B \frac{E_{\max}}{E_{required,t_1}} (t - t_1)} \right) \tag{1}$$

Here, t_1 is the moment of an automation-initiated transition that requires driver attention; $E_{required,t_1}$ is the effort level that is required by the driver to acquire all necessary awareness of the environment at t_1; E_{max} is the maximum visual effort the driver can exert (representing human limitations); A and B are fitted parameters of the exponential function. The required effort $E_{required,t_1}$ represents the information density of the environment as well as the required spatial movement of the eyes that scan and sample the environment. In a particular scenario, A, B and E_{max} are constants. As can be seen in Eq. (1), $E_{required,t_1}$ does not only affect the stationary SA level, but also the time to achieve said stationary SA level.

Figure 1 shows a description of the process according to functions $S(t, t_1)$ and $S(t, t_1)'$, where the required effort in $S(t, t_1)'$ is smaller than the required effort in $S(t, t_1)$.

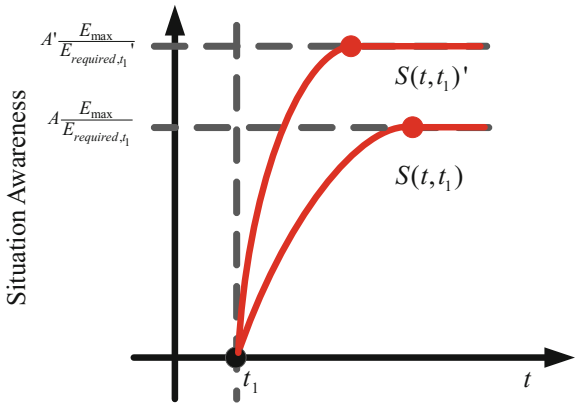


Fig. 1. The dynamic process of driver SA after a warning that signals an automation-initiated transition

2.2 Model Verification

In a study by Lu et al. [11], SA assessments in automated driving were conducted. To ensure that the environment was identical for all participants, the study used animated video clips from an ego-car perspective. The scenarios were presented to the participants for 1, 3, 7, 9, 12 or 20 s with either 4 or 6 surrounding cars. After each video, the participants were asked to position the vehicles in a top-down view, placing a minimum of 1 and a maximum of 8 vehicles without time restriction. Participants also indicated the speeds of the placed cars in relation to the ego-vehicle. This allowed the authors to calculate participants' error scores regarding the number of cars, the relative positions of the cars, and the relative speed of the cars with respect to the true values at the end of the video. A SmartEye DR120 remote eye-tracker was used to record the eye movements. Screenshots of a video and the questionnaire interface are shown in Figs. 2 and 3.

Fig. 2. Screenshot of a video that includes six surrounding cars

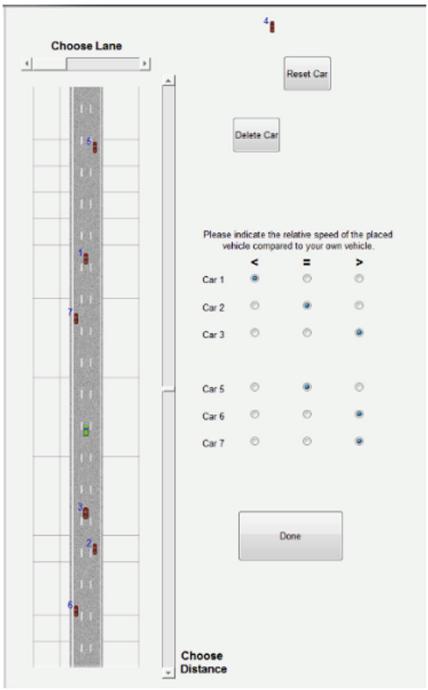


Fig. 3. Reproduction of a traffic situation. *Red cars* represent the surrounding cars. The *green car* represents the ego-vehicle. The *slider bars* were used to set the positions of surrounding vehicles.

Thirty-four persons (5 female, 29 male) participated in the study. The participants all had a valid driver's license, and were aged between 20 and 31 years ($M = 24.6$, $SD = 2.6$ years). To verify the dynamic model proposed above, we use the absolute percentage error of the number of placed cars as a measure of SA. The absolute percentage error of the number of placed cars is the absolute difference between the number of placed cars in the questionnaire and the number of actual cars in the video, divided by the number of actual cars in the video. The results (Fig. 4) show a good fit for an exponential function.

3 SA Real-Time Monitoring with Eye-Tracking Data

The model can be further developed in several ways. For example, one may collect as much data as possible in simulators or in real traffic in order to define the parameters in the model for different traffic scenarios. Accordingly, a scheduling function can be implemented that selects the appropriate parameters for each type of transition. Ideally, the driver's SA level should be recorded online (i.e., while driving). However, it is usually not possible to ask the drivers to answer whether they are ready to perform a

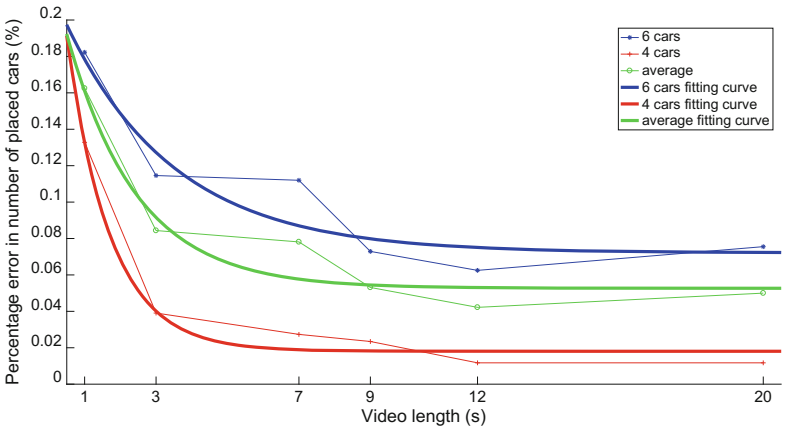


Fig. 4. Mean absolute percentage error of the number of placed cars in videos with 4 surrounding cars, in videos with 6 surrounding cars, and in all videos. The corresponding fitting functions are $f_6(t) = 0.1475e(-0.3280t) + 0.07214$; $f_4(t) = 0.2611e(-0.8235t) + 0.01807$; $f_{all}(t) = 0.1801e(-0.5103t) + 0.05266$

transition, or what they are aware of in real traffic. Therefore, the goal should be to find an online measure that can reflects a driver's momentary SA level.

According to Endsley's three stages of SA, the first stage is perception [6]. In our future studies, we will particularly focus on visual scanning, which is a known index of attention allocation and cognitive strategies [7–12]. Even though Corbetta and Shulman [10] pointed out that different parts of the brain may carry different attentional functions (e.g., goal-directed vs. stimulus-driven attention), we do not distinguish between these different types of attention. In our model, the information is presented when the driver comes back from an out-of-the-loop situation triggered by a warning signal.

The number of fixations per second is a common measure in visual search and visual attention studies [13]. Eye-tracking research in which observers viewed static scenes [14] or dynamic events [15] show a similar pattern as the SA dynamics in Fig. 4. In [14], the fixation duration (i.e., closely related to the reciprocal of the number of fixations per second) as a function of time was also found to conform to an exponential function: $f = be^{(-a/t)}$.

4 Discussion

SA is a concept that has been extensively studied in the last several decades. However, the concept of SA has also been criticized [16]. Even so, a quantification of SA may prove useful in analyzing a driver's readiness to engage an automation-to-manual transition. The results of the numerical model that we presented could also be applied to predict the required time budget between the warning signal (take-over request) and the moment of taking of control. However, further verification of our model will be required.

The architecture of the proposed model is based on an exponential function. Historically, the exponential function has been widely used to quantify some measurements in psychology, such as learning curves [17], information extraction rate [18], and spatial memory decay [19]. Therefore, we used it to model the dynamics of SA. Some researchers have argued that the power law is more suitable for describing learning, memory, etc. [20]. Whether exponential functions or power functions are more suitable is a topic that can be explored in the future studies.

Driver monitoring systems have been developed that measure distraction and fatigue [21]. For future vehicles, a monitoring system may compute a 'level of SA' instead of computing a discrete value (e.g., attentive vs. inattentive). Only 'duration of fixation' was proposed in our study as a possible indicator of SA. In future studies, multiple measures could be adopted for SA assessment.

Acknowledgments. The authors are involved in the Marie Curie Initial Training Network (ITN) HFAuto – Human Factors of Automated Driving (PITN-GA-2013-605817).

References

1. Bainbridge, L.: Ironies of automation. Automatica **19**, 775–779 (1983)
2. Lu, Z., Happee, R., Cabrall, C.D.D., Kyriakidis, M., De Winter, J.C.F.: Human factors of transitions in automated driving: a general framework and literature survey. Transp. Res. Part F **43**, 183–198 (2016)
3. Flemisch, F., Heesen, M., Hesse, T., Kelsch, J., Schieben, A., Beller, J.: Towards a dynamic balance between humans and automation : authority, ability, responsibility and control in shared and cooperative control situations. Cognit. Technol. Work **14**, 3–18 (2012)
4. Wickens, C.D., McCarley, J.S., Alexander, A.L., Thomas, L.C., Zheng, S.: Attention-situation awareness model of pilot error. In: Foyle, D., Hooey, B. (eds.) Pilot error models. Erlbaum, Mahwah (2008)
5. Wiegmann, D.A., Shappell, S.A.: Human factors analysis of post accident data: applying theoretical taxonomies of human error. Int. J. Aviat. Psychol. **7**, 67–82 (1997)
6. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Fact. **37**, 85–104 (1995)
7. Bellenkes, A.H., Wickens, C.D., Kramer, A.F.: Visual scanning and pilot expertise: the role of attentional flexibility and mental model development. Aviat. Space Environ. Med. **68**, 569–579 (1997)
8. Yarbus, A.J.: Eye Movements and Vision. Plenum Press, New York (1967)
9. Cohen, K.M.: The development of strategies of visual search. In: Fisher, D.F., Monty, R.A., Senders, J.W. (eds.) Eye movements: cognition and visual perception. Erbaum, Hillsdale (1981)
10. Corbetta, M., Shulman, G.: Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. **3**, 201–215 (2002)
11. Lu, Z., Coster, X., De Winter, J.C.F.: How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving. Appl. Ergono. **60**, 293–304 (2017)
12. Rayner, K.: Eye movements and attention in reading, scene perception, and visual search. Q. J. Exp. Psychol. **62**, 1457–1506 (2009)

13. Borji, A., Itti, L.: State-of-the-art in visual attention modelling. IEEE Trans. Pattern Anal. Mach. Intell. **35**, 185–207 (2013)
14. Unema, P.J., Pannasch, S., Joos, M., Velichkovsky, B.M.: Time course of information processing during scene perception: the relationship between saccade amplitude and fixation duration. Vis. Cognit. **12**, 473–494 (2005)
15. Velichkovsky, B.M., Rothert, A., Kopf, M., Dornhoefer, S.M., Joos, M.: Towards an express diagnostics for level of processing and hazard perception. Transp. Res. Part F **5**, 145–156 (2002)
16. Flach, J.M.: Situation awareness: context matters! A commentary on Endsley. J. Cognit. Eng. Decis. Mak. **9**, 59–71 (2015)
17. Anzanello, M.J., Fogliatto, F.S.: Learning curve models and applications: literature review and research directions. Int. J. Ind. Ergon. **41**, 573–583 (2011)
18. Loftus, G.R.: Picture perception: effects of luminance on available information and information-extraction rate. J. Exp. Psychol. **114**, 342–356 (1985)
19. Atkinson, R.C., Shiffrin, R.M.: Human memory: a proposed system and its control processes. In: Spence, K.W. (ed.) The psychology of learning and motivation: advances in research and theory, 2, pp. 89–195. Academic Press, New York (1968)
20. Anderson, R.B.: The power law as an emergent property. Mem. Cognit. **29**, 1061–1068 (2001)
21. Dong, Y., Hu, Z., Uchimura, K., Murayama, N.: Driver inattention monitoring system for intelligent vehicles: a review. IEEE Trans. Intell. Transp. Syst. **12**, 596–614 (2011)

Evaluation of an Autonomous Vehicle External Communication System Concept: A Survey Study

Jingyi Zhang(✉), Erik Vinkhuyzen, and Melissa Cefkin

Nissan Research Center Silicon Valley, 1215 Bordeaux Dr, Sunnyvale, CA 94089, USA
jingyi.zhang@nissan-usa.com

Abstract. Road users communicate by reading the intention, and demonstrating their own intention to others. The drivers of Autonomous Vehicles (AVs) may not be engaged in controlling the vehicle and therefore unavailable for interaction. Based on empirical research on road use interactions, a new form of vehicle lighting signal was developed to indicate AV's intention. In this study, we used a web-based survey to collect public feedback on a number of different light configurations to communicate five vehicle motions: Slowing down, Waiting, Planning to go, Starting to go, and Going. The results validated the understandability of the concept, but reversed our initial design of the display option. Respondents associated green with a moving vehicle, and red a slowing down or stopped vehicle; and associated a forward moving light with a car that was starting to go or going and a backward flowing light with a vehicle slowing down or stopping.

Keywords: Autonomous vehicles · Vehicle external communication · Human-machine interaction (HMI) · Other road users · Social interactions

1 Introduction

In traffic, road users communicate in a variety of ways, through their movement, their positioning, their body language, and sometimes through interactions, that is, by establishing eye contact and gesturing [1, 2]. Such interactions, though not regulated, play a critical role in making traffic flow smoothly. With the rapid development of autonomous technology, it is anticipated that Autonomous Vehicles (AVs), or Highly Automated Vehicles, would soon be deployed on the road [3, 4]. However, how AVs can interact with other road users still remains as a crucial challenge to AV developers [5].

In this study, we are looking at the interactions between AVs and other road users. We believe that, if not achieving mutual interactions, AVs should at least display its own internal states to other people on the road, and as such make clear what its own intentions are. Based on our extensive video recordings of interactions between road users at intersections, we have discovered two critical findings: (1) Interactions were mostly between drivers and other road users in the front and the side of the car; and (2) Interactions were brief, sometimes just one turn ("go ahead"), occasionally two

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_63

("go ahead," "no you go ahead"), followed by thank you in the form of a wave, a head nod or the like. With everything in mind, we have developed the Intention Indicator, a new form of lighting signal to indicate AV's intent to other road users.

The Intention Indicator would indicate 5 modes, based on observable behavior of vehicles at four-way stop intersections: (1) Slowing down, when cars approach the intersection, apply the breaks to come to a stop; (2) Waiting, when cars are stopped to let other road users go ahead; (3) Planning to go, when cars have anticipated the gap in which it will go; (4) Starting to go, when cars 'creep' into the intersection to show that they will go next; and (5) Going, when cars accelerate to take their spot to go through the intersection.

The descriptions of these modes are focused on the behavior of the AV. However, traffic is social and actions by one road user engender reactions and of course the whole point of the intention indicator is to communicate, so the AV's action of "slowing down and waiting" has as its natural response "going" on the part of other road users. And the modes of "planning to go," "starting to go," and "going" has as the natural response "yielding" (or simply 'not going') behaviors on the part of the other road users. The open question is whether other road users would interpret the Intention Indicator as a communication about the AV's internal state or a communication about the expected response on the part of the other road users.

Since the Intention Indicator's functionality is aimed at facilitating the ordering in which road users can cross an intersection, we thought it would be quite natural to use the colors green and red that all experienced road users are familiar with. However, given the contrast between what the intention indicator communicates about itself and the actions it thus enables on the part of other road users, the decision of which color the intention indicator should display in different modes can be done in two diametrically opposed ways: it could show red when it is slowing down or waiting, and green when it is planning to go and going, or do the direct opposite. It all depends on whether one sees the intention indicator as a way to display what the AV is doing, or as communicating to other road users what actions they can take. In the default configurations, it was decided to use color red to indicate the vehicle starting to go, therefore communicate to other road users that they should stop, and to use color green to indicate vehicle slowing down or stopping to communicate that other road users can go.

The intention indicator is a new way for a car to signal its intention, and therefore introducing such a system means a new learning for people on the road. We are currently testing how such a system can work most effectively, and how it most naturally taps into expectations people have about cars and the signals that indicate the driver (or its) intention. In part the study was aimed to shed some lights on two questions: (1) if the five modes can be identified and understood; and (2) what are the configurations of each mode that were found more intuitive than others.

2 Method

2.1 Intention Indicator

The prototype of the Intention Indicator being tested in the study consists of an approximate of 6-meter long RGB LED strip that covers the front doors and the hood

(Fig. 1). The strip is connected to an Arduino board and is manually controlled for switching among the mode configurations. The total luminous intensity of the LED is 1.63 candela, with Red being 0.65 candela, Green 0.52 candela, and Blue 0.46 candela.

Five modes were tested in the survey: Slowing Down, Waiting, Planning to Go, Starting to Go, and Going. The default configurations of the modes are illustrated in Table 1.

Fig. 1. Intention indicator concept car, front and side

2.2 Survey Design

Each configuration was pre-filmed and edited into a 6-second video clip. In order to include and use video clips in the instructions, we have decided to develop the survey using the online survey program SurveyMonkey. Additionally, web-based surveys also have certain advantages in terms of time and cost [6], and can prevent questions from being skipped. A total of 13 questions were included in the survey. The questions were divided into 4 sections: Concept Identification (CI) (Q1–Q5), Concept Comparison

Table 1. Testing modes and default configurations

Mode	Default configuration (color, motion)
Slowing down	Green, moving forward
Waiting	Green, static
Planning to go	White, flashing
Starting to go	Red, moving backward (slow)
Going	Red, moving backward (fast)

(CC) (Q6–Q10), comments (Q11), and demographic section (Q12–Q13). All the questions, except for the questions in the demographic section, were set as required. In CI and CC sections, participants were asked to watch a video clip that contains either one or more configurations before answering each question.

At the beginning of the survey, an introduction was provided. A video of the concept car coming to a stop sign with all different modes was shown. The modes in the video appeared in the order of Slowing down, Waiting, Planning to go, Starting to go, and Going. Participants were suggested to watch the introductory video first before continuing taking the survey. Although the concept is inspired by the idea of improving the communications between AVs and other road users, we did not specifically address this to be an AV scenario for two reasons: (1) since AVs do not yet exist on the road, it is difficult for the participants to imagine an AV scenario which they don't have any experience with; and (2) The concept could also potentially be added to current in-market vehicles to improve the vehicle – road user communications.

In the CI section, five questions were designed to examine if participants were able to identify each mode as we defined. In each question, participants were asked to watch one default configuration and complete the sentence "The car communicates that …". The vehicle in the videos was stationary and parked in a parking spot.

In the CC section, five questions with 2 sub-questions in each were designed to examine if participants agreed that the default configurations could best represent their associated modes. In each question, participants were asked to watch a video before answering the two sub-questions. Besides the default configuration, each video also contained 2–3 additional configurations which all represented one mode (Table 2). In the first question, participants were asked to choose one among the 5 options – "Slowing down", "Waiting", "Planning to go", "Starting to go", and "Going" – that best described what the vehicle in the video was communicating. In the second question, participants were asked to, based on their previous selection, rank the

Table 2. Concept comparison default and additional configurations of each mode

Mode	Additional configurations (color, motion)		
Slowing down	Red, moving forward	Blue, moving forward	Breaking green, static
Waiting	Red, static	Blue, static	
Planning to go	Blue, flashing	Red, static	
Starting to go	Breaking red, static	Green, moving backward	Blue, moving backward
Going	Red, static	Green, moving backward	Blue, moving backward

configurations shown in the video according to their liking, where 1 meant the best, and 3 (or 4 in 4-configuration comparison) meant the worst.

2.3 Participants

The study was first piloted with 6 employees within the research facility using a mix of random sampling and convenience sampling. The pilot feedback was used to assess the validity and modify the survey accordingly. The recruiting of the final survey was arranged by SurveyMonkey. A total of 145 participants accessed the survey, among which 92 responses were gathered.

3 Results

3.1 Demographics

A total of 74 respondents completed the demographic section, results are shown in Table 3. 63.5% of the respondents were female, 77% of the respondents were 30 years old and older, 67.6% reported to commute by car, and 52.7% reported annual income of more than $50,000.

Table 3. Characteristics of survey participants (N = 74)

Characteristic	N (%)
Gender	
Male	27 (36.5%)
Female	47 (63.5%)
Age	
18–29	17 (23.0%)
30–44	26 (35.1%)
45–59	21 (28.4%)
60+	10 (13.5%)
Way of commute	
Drive	50 (67.6%)
Walk	3 (4.1%)
Bike	1 (1.4%)
Public transportation	7 (9.5%)
Carpool	1 (1.4%)
Other	12 (16.2%)
Annual household income	
< $25,000	13 (17.6%)
$25,000–$49,999	17 (23.0%)
$50,000–$99,999	18 (24.3%)
$100,000–$149,999	11 (14.9%)
≥ $150,000	10 (13.5%)
Prefer not to answer	5 (6.8%)

3.2 Concept Identification

In the CI section, each question contained a video of one mode with its default configuration. Participants were asked to finish the sentence "The car communicates that…" Due to video quality and other technical issues, the video showing "Starting to go" was not properly played, thus its responses were not analyzed. For other questions, a content analysis was conducted in order to determine if any meaningful results existed. After examining participants' responses, a taxonomy (Table 4) was created based on our testing hypotheses. Since the questions were set as required, there were a few invalid responses (e.g. symbols, irrelevant answers) that were not included in the analysis.

Table 4. Concept identification questions taxonomy

Coding category	Descriptions and keyword examples
Indication of vehicle stopping	"Stopping", "Parked", "Stopped", "Waiting"
Indication of vehicle movement	"Ready to go", "Leaving", "Moving", "Turning"
Indication of vehicle status	"Vehicle…", "On", "Off", "I'm looking", etc
Indication to other road users	"It's safe to cross"
Not sure	Not sure what the vehicle is indicating

Figure 2 shows the number of participants' responses of each question based on the coding categories. Chi Square statistics was utilized to examine the differences among the categories in each question. For all of the modes, the differences of number of responses among the categories were significant: Slowing down $\chi^2(4, N = 75) = 31.20, p < .001$; Waiting $\chi^2(4, N = 65) = 13.69, p = .008$; Planning to go $\chi^2(4, N = 62) = 15.90, p = .003$; and Going $\chi^2(4, N = 55) = 12.91, p = .012$. Post hoc analysis of using Binomial test with Bonferroni correction was conducted afterwards to further explore the differences between the response categories (Table 5).

3.3 Concept Comparison

In the CC section, five questions with two sub-questions in each were designed. In each question, participants were asked to watch a video clip before answering the two sub-questions. In the first sub-question, participants were asked to choose one in the 5 options "Slowing down", "Waiting", "Planning to go", "Starting to go", and "Going" to best describe what the vehicle in the video was communicating. In the second sub-question, participants were asked to, based on their previous selection, rank the configurations shown in the video according to their liking.

For multiple-choice questions, Chi Square statistics was performed to explore the differences in each question, and significances were found in all the modes: for Slowing down $\chi^2(4, N = 76) = 90.05, p < 0.001$; Waiting $\chi^2(4, N = 74) = 121.27, p < 0.001$; Planning to go $\chi^2(4, N = 85) = 53.06, p < 0.001$; Starting to go $\chi^2(4, N = 78) = 57.26, p < 0.001$; and Going $\chi^2(4, N = 87) = 42.83, p < 0.001$. Response distributions are shown in Fig. 3.

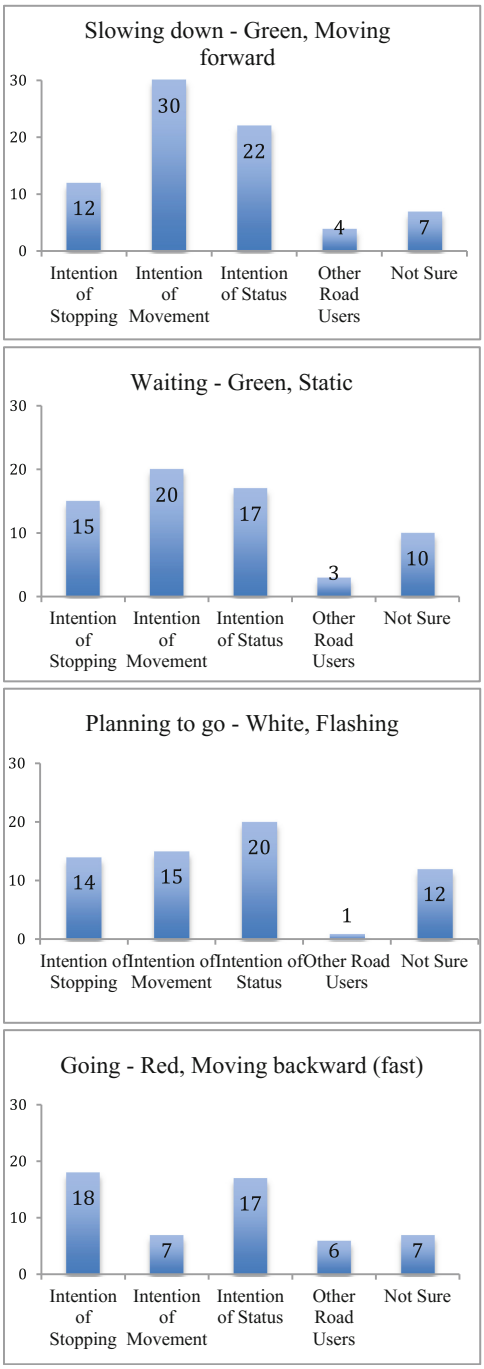


Fig. 2. Concept identification – participant responses

Table 5. Concept identification – binomial test results between the coding categories

	Intention of movement	Intention of status	Other road users	Not sure
Slowing down				
Intention of stopping	*no sig.*	*no sig.*	*no sig.*	*no sig.*
Intention of movement		*no sig.*	$p < .001$	$p < .001$
Intention of status			$p = .001$	*no sig.*
Other road users				*no sig.*
Waiting				
Intention of stopping	*no sig.*	*no sig.*	*no sig.*	*no sig.*
Intention of movement		*no sig.*	$p < .001$	*no sig.*
Intention of status			$p = .003$	*no sig.*
Other road users				*no sig.*
Planning to go				
Intention of stopping	*no sig.*	*no sig.*	$p = .001$	*no sig.*
Intention of movement		*no sig.*	$p = .001$	*no sig.*
Intention of status			$p < .001$	*no sig.*
Other road users				$p = .003$
Going				
Intention of stopping	*no sig.*	*no sig.*	*no sig.*	*no sig.*
Intention of movement		*no sig.*	*no sig.*	*no sig.*
Intention of status			*no sig.*	*no sig.*
Other road users				*no sig.*

Post hoc analysis of using Binomial test with Bonferroni correction was conducted to compare the differences between the selections of the default option and the other options respectively.

For "Slowing down" mode testing, the number of selecting "Slowing down" ($N = 47$) was significantly more then the number of selecting "Waiting" ($N = 3$, $p < .001$), "Planning to go" ($N = 5$, $p < .001$), "Starting to go" ($N = 5$, $p < .001$), and "Going" ($N = 16$, $p < .001$).

Similarly, for "Waiting" mode testing, the number of selecting "Waiting" ($N = 52$) was significantly more then the number of selecting "Slowing down" ($N = 12$, $p < .001$), "Planning to go" ($N = 4$, $p < .001$), "Starting to go" ($N = 1$, $p < .001$), and "Going" ($N = 5$, $p < .001$).

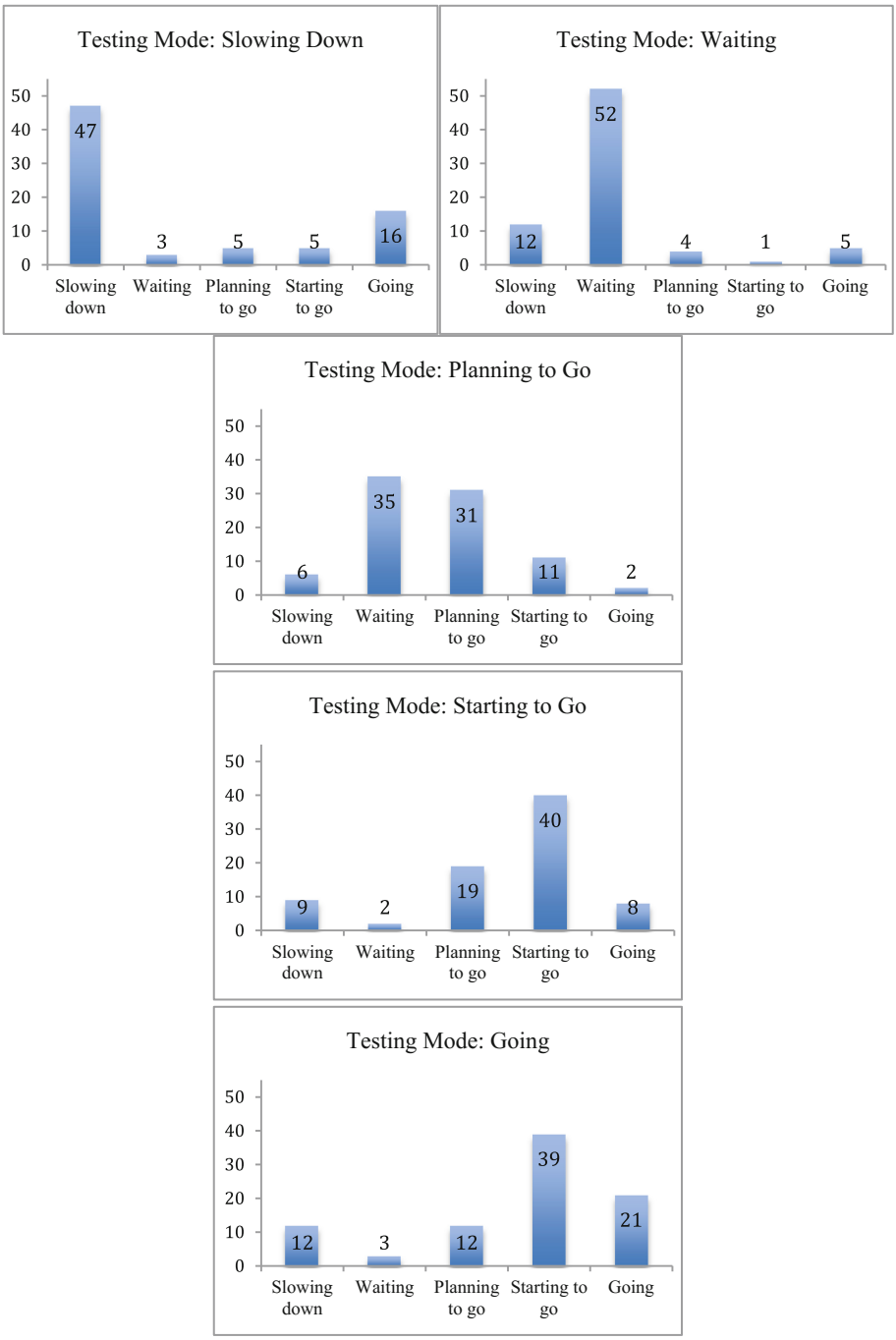


Fig. 3. Concept comparison – participant responses on multiple-choice questions

For "Planning to go" mode testing, the number of selecting "Planning to go" (N = 31) was significantly more than the number of selecting "Slowing down" (N = 6, p < .001), "Starting to go" (N = 11, p = .003), and "Going" (N = 2, p < .001).

For "Starting to go" mode testing, the number of selecting "Starting to go" (N = 40) was significantly more than the number of selecting "Slowing down" (N = 9, p < .001), "Waiting" (N = 2, p < .001), "Going" (N = 8, p < .001); and was marginally significantly more than the number of selecting "Planning to go" (N = 19, p = .009).

For "Going" mode testing, the number of selecting "Going" (N = 21) was only significantly more than the number of selecting "Waiting" (N = 3, p < .001). No other significances were found.

For questions of ranking the given configurations, only the selections of the default options in each question were analyzed. A non-parametric Friedman test was performed to examine the differences among the configurations. Significant differences were found in the "Slowing down" configurations ($\chi^2(3, N = 47) = 26.87$, $p < 0.001$), and "Planning to go" configurations ($\chi^2(3, N = 47) = 26.87$, $p < 0.001$). For "Waiting", "Starting to go", and "Going", no significant differences were found.

Post hoc analysis of using Wilcoxon Signed-rank test with Bonferroni correction was conducted to explore the differences between the selections of the default configuration and the other configurations respectively. For the given "Slowing down" configurations, respondents preferred "Red, Moving forward" than the default configuration "Green, Moving forward" (Z = 3.327, p = .001). For the given "Planning to go" configurations, respondents preferred the default configuration "White, Flashing" over the configuration "Blue, Flashing" (Z = 2.936, p = .003), and the configuration "Red, Static" (Z = −4.059, p < .001).

4 Conclusion and Discussion

This study was exploratory in nature with the objective of using survey to identify public feedback on an early lighting signal concept of communicating five vehicle intentions to other road users. Specifically, the study was conducted with two questions in mind: (1) If the five modes can be identified and understood, and (2) if the default configurations of the five modes were intuitive and accepted. The five modes being examined were: Slowing Down, Waiting, Planning to Go, Starting to Go, and Going.

The results of the Concept Identification section have indicated that participants did not consider the Intention Indicator as an indication of instructions to other road users in all five modes. The results have supported our design of the Intention Indicator communicating vehicle's own intents. This is also consistent with all the existing lighting equipment on the vehicles, which are indications of either vehicle status or movement.

In terms of colors and pattern, participants had various interpretations. Although no significance was found, a trend has indicated that participants associated with color green with a moving vehicle or a functioning vehicle, and color red with a stopping vehicle or a malfunctioning vehicle. This reversed our color design, as we used red for vehicle acceleration, and green for vehicle deceleration (Table 1). One explanation could be since participants interpreted the Intention Indicator as a communication of

the AV's internal state, or AV's status, not a communication about the expected response on the part of other road users, participants would expect both the color and the motion to represent AV's behaviors or status as well.

The results of the Concept Comparison section have indicated that participants were able to recognize Slowing Down, Waiting to go, and Starting to Go; but confused Planning to Go with Waiting, and Going with Starting to Go.

The results have also suggested that participants has agreed with our design in terms of the movement of the light, but again disagreed with our design in terms of color: participants preferred color red with Slowing Down and Waiting, white with Planning to Go, green with Starting to Go and Going. This is consistent with the results shown in the Concept Identification section.

The study has provided meaningful information regarding to our questions. It has confirmed that the Intention Indicator should display AV's intents, rather than an instruction to other road users; it has indicated that colors are more difficult to interpret than the motion of the lighting; and it has showed that color green is commonly associated with a moving/accelerating vehicle, and color red with a stopping/decelerating vehicle. However the question of whether the design of the five modes is intuitive or not still remains unclear. It is suggested in another study that when encountering such novel communication methods used for AVs and other road user communication, pedestrians are more likely to understand the messages when training is provided [7]. Other modalities such as icons, text displays and sound are also being considered, however, Clamann et al. has suggested that the type of display may be less relevant when pedestrians are making decisions of crossing the road [8].

Based on the results, several suggestions have been made: (1) three to fours modes, instead of five modes, may be more appropriate, with the possibility of combing Starting to Go with Going, and Slowing Down with Waiting; (2) a single neutral color combined with various motions to represent different modes might be sufficient enough, as movement seems clear enough, and color creates confusions. The results of the study will inform the design of the Intention Indicator in future testing and research.

References

1. Haddington, P., Rauniomaa, M.: Interaction between road users: offering space in traffic. Space Cult. **17**(2), 176–190 (2013)
2. Jensen, O.B.: Negotiation in motion: unpacking a geography of mobility. Space Cult. **13**(4), 389–402 (2010)
3. Nissan's Autonomous Driver Car. https://www.nissanusa.com/blog/autonomous-drive-car
4. Corporations Working on Autonomous Vehicles. https://www.cbinsights.com/blog/autonomous-driverless-vehicles-corporations-list/
5. Vinkhuyzen, E., Cefkin, M.: Developing socially acceptable autonomous vehicles. In: 2016 Ethnographic Praxis in Industry Proceedings, pp. 522–534 (2016)
6. Hayslett, M.M., Wildemuth, B.M.: Pixels or pencils? The relative effectiveness of web-based versus paper surveys. Libr. Inf. Sci. Res. **26**(1), 73–93 (2004)

7. Langstrom, T., Lundgren, V.M.: AVIP—Autonomous vehicles interaction with pedestrians: an investigation of pedestrian-driver communication and development of a vehicle external interface. Master Thesis, Chalmers University of Technology, Gothenborg, Sweden (2015)
8. Clamann, M., Aubert, M., Cummings, M.L.: Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles. In: 96th Annual Transportation Research Board Meeting, Washington, DC (2017)

Machine Learning and Big Data Analytics in Support of Fleet Safety During Severe Weather

Zachary Spielman[1(✉)], David I. Gertman[2], Haoran Liu[3], Ira Pray[1], Justin Traiteur[3], Scott Wold[1], and Steven Wysmuller[3]

[1] Idaho National Laboratory, Human Factors, Idaho Falls, ID 83415, USA
zachary.spielman@inl.gov
[2] Idaho National Laboratory, [Retired], Idaho Falls, ID 83415, USA
[3] IBM Global Business Services, Advanced Analytics Center of Competency (CoC), Armonk, NY, USA

Abstract. The US DoT estimates 22% of the 5.7 million vehicle crashes a year are weather related. At Idaho National Laboratories, home of the DOE's largest transit, heavy and light vehicle fleet in the nation, weather is a constant challenge for the 4000 employees traveling the 45 to 65 mile stretch of road. Driving conditions can vary immensely; micro-climate conditions at INL site locations highways go unmonitored and causing severe challenges. INL has taken the initiative to review applicable technologies determining that addressing severe weather and road conditions through the application of advanced modeling methods holds promise for enhancing driver safety and dispatch planning. INL engaged IBM Global Business Services Advanced Analytics Center of Competency (CoC) Team for support in this effort. This presentation reviews the benefits expected, data surveyed, and how to use integrated sources and cognitive analytics to improve real-time weather forecasting and INL site fleet and operations planning.

Keywords: Human factors · Machine learning · Weather prediction · Transportation

1 Introduction

Weather has a detrimental role in road safety and efficiency. In fact, the US Department of Transportation (DoT) estimates 22% of the 5.7 million vehicle crashes a year are weather-related [1]. Weather can also cause unexpected road closures and delays keeping people from work and reducing productivity. From an economic perspective, weather delays caused approximately 3.4 billion dollars a year in freight costs [2]. These numbers are not changing. Over the past decade, the number of deaths related to vehicular accidents has remained almost constant, and the US DoT has been investigating means by which to reduce this danger to the public. The US DoT Federal Highway Administration (FHWA) emphasizes the national need to improve road condition and weather data saying, "Today's weather data is generally insufficient to support transportation operations and safer travel decisions - forecasts are not specific to roadway segments and current data does not characterize roadway conditions". Development of

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_64

a technical basis for improving safe driving under varied weather conditions requires two improvements. From a meteorological perspective, increased weather predictability, focusing on forecast accuracy and precision, is necessary and accomplishable by incorporating additional high-resolution weather data in weather models. Aggregating and integrating currently available data sources is one way to obtain the necessary data well as identifying or creating new sources for analysis. Examples of data sources include:

- Local, high resolution, geographically specific sources
- Private weather collection systems to increase data resolution of a geographic area
- Public sources such as government entities and historical weather databases
- Stationary sensors such as road side weather stations
- Mobile data sources, a concept in development that helps integrate information from vehicles instrumented to gather road condition information.

A second option is to increase communication of weather and road conditions to drivers and vehicles. As vehicle safety autonomy increases, the advantages of signaling the vehicle directly to automatically ensure proper drive modes can be engaged (or disengaged i.e. cruise control) to accommodate road conditions should be considered. Providing access to accurate and timely alerts to daily commuters helps vehicle operators avoid road hazards and increase driver awareness of current road conditions. Design principles must ensure driver safety is a priority. That is, messages should not take too long to read and understand as well as not distract the driver when sent. Also being considered is driver workload with concurrent tasks. It is likely driver situation awareness is already compromised by hazardous weather when the information arrives thus careful delivery and display design must occur to avoid incurring further driver risk. Communication to the driver may assist in improving decisions as well as promote safe driving behavior. Increasing driver awareness by providing the appropriate road and weather information is expected to reduce weather-related crashes, improve commute times, and lessen the number of emergency response incidents during hazardous conditions.

Increasing the level of detail and accuracy in road condition reporting is a complex but valuable investment, as it can significantly improve the caliber of site-specific, road-condition forecasts for benefiting the daily commuter, mass transportation agencies and road management agencies. Federal and state transportation agencies require detailed, accurate, and geo-specific weather data to anticipate, observe, and manage hazardous conditions. Improving the data quality and model capability available helps government entities manage and mitigate hazardous road conditions caused by weather more efficiently. Areas in need can be precisely targeted and thus avoid wasting resources and time on unaffected areas. Some road hazard mitigation techniques such as road salt/liquid applications can have damaging effects on the surrounding environment as well as the vehicles driving across it [3].

The US DoT FHWA further notes that road weather sensors are implemented by a variety of organizations:

- State agencies
- Local agencies
- Weather monitoring

– Private sector firms

The end result of having multiple organizations involved is little aggregation, integration, or data exchange to supply the detail needed to support safer travel choices at either broader, regional levels or specific, local levels. The research reported in this paper, seeks to improve the integration of various sources including historical data to produce safety messaging and weather and road status that can be used on the local level [3].

Challenges to collaboration among the listed owners of road weather sensors noted by the FHWA include:

– Data formats differ across sensor types (environmental, imaging, mobile, and remote)
– Data formats differ across organizations and borders
– Data are typically not readily available to forecasters
– Existing data tends to address only general atmospheric conditions rather than specific roadway information
– The public does not have sufficient or necessary road weather awareness and tools to be able to access information they can use
– Complexities in data curation, data integration and addressing gaps in data can be challenging and time consuming
– Inherent systemic biases between disparate data sources are difficult to measure or account for across sensor or weather station types
– There is not an agreed upon repository nor format for the display of this information to users

As a result, it is difficult to validate the quality of the data to ensure reliability and accuracy. As long as forecasts depend on separate rather than integrated sources, different information is reported by individual sources and this places the burden on the user to decide which the most accurate and useful source [4].

INL Mission Support Services' approach to this problem is to develop an integrated network of local weather and road condition sensors, both mobile and stationary, to boost the resolution of existing meteorological sources. Then, using the integrated network and the power of big data combined with advanced analytical modeling, create a real-time weather prediction system with acuity at the microclimate level. Specifically, ahistorical model is trained based on historical route performance (i.e. how weather impacts driver speed) and the top weather variables/predictors which impact vehicular performance are identified. Using these weather drivers, a forecast model is being created to predict future route delays due to weather factors. The future goals of this project include real time forecasting to government and private dispatch units as well as communicating the relevant weather and road information to vehicles approaching hazardous road conditions.

2 Idaho National Laboratory Mission Support Services

The US Department of Energy (DOE) owns a large research facility in eastern Idaho called the Idaho National Laboratory (INL) and operated under contract with Battelle

Energy Alliance (BEA) employing approximately 4000 people at INL's site, with an additional 2000 employees working in Idaho Falls. The research facility covers approximately 890 square miles in a remote part of eastern Idaho, away from major population areas, requiring a large portion of the workforce to commute to the facility on a daily basis. BEA provides transportation services for their employees through a fleet of transit buses, vanpools, and car pools. Most of the employees live along the I-15 and US-20 corridors in eastern Idaho, from Pocatello to Rexburg, and travel west or northwest to the INL site. The Inland the Idaho Transportation Department (ITD) have been working together for many years in dealing with the road weather issues common in this part of the state, including snow, ice and visibility limitations caused by fog, smoke, blowing dust and blowing snow. These conditions contribute to crashes, sudden road closures, mandate rescheduling of or adjustment to INL work hours and in extreme conditions the closing of INL facilities. In the past, hazardous weather has stranded employees at INL facilities due to unsafe travel conditions. The ability to predict with greater acuity the near future highway conditions, employees could be evacuated earlier during safer conditions to prevent stranding.

The INL's Mission Support Services (MSS) manages the fleet of motor coaches used to transport INL employees from in town park-and-rides to the INL site on a route 45–65 miles each way. MSS monitors and maintains highway and service roads INL Mission Support Services. There is a joint effort with ITD to create a network of both mobile and stationary intelligent sensors with V2X communication capabilities to support development of an intelligent weather forecast system able to communicate meaningful information to vehicles approaching hazardous conditions. The implementation will provide the capability of either communicating directly with vehicles on the road, or broadcasting messages to vehicles approaching particularly hazardous areas using their location to trigger warnings. Difficulties in developing a location sensitive, data rich, secure communications network abound. However, INL is unique in having access to controlled roads, roadside weather information systems (RWIS), vehicles specially equipped to collect road condition data, access to a range of heavy and light vehicles, diverse and hazardous weather conditions each year, and an isolated but highly trafficked portion of highway year round. Part of this approach entails making available state of the art weather and road condition information.

3 Near-Real Time Weather Forecasting

3.1 Increased Resolution

Collaborating with academia and industry, MSS is employing an integrated systems approach to couple stationary roadside weather sensors with mobile road condition sensors mounted on various INL vehicles such as trucks, plows, and motor coaches. Adding more roadside units will increase the resolution of weather forecasting by collecting more data at closer intervals than is currently available. Adding information from mobile units provides near continues road condition data across the entire route. The outcome from using machine learning techniques to data mine both mobile and stationary information source is an effective algorithm that can interpolate the stationary

weather information into road condition information and vice versa creating redundant tools for forecasting information relevant to creating safer, more manageable roads.

Parallel activities to real time weather reporting may include the development and field test of an analytically driven microclimate prediction system. Having better detail in weather reporting systems is important to drivers currently on the road or entering the road. However, as road conditions are likely to change and operations need to make informed plans it is important to have an accurate future model of how the road conditions may change. Using the same tools but applying machine learning in a different fashion, an algorithm can be created to support prediction of road and weather conditions at intervals into the future, likely with varying levels of confidence. Though such forecasts seem to exist now, the difference here is the acuity and customized forecast for the area using higher resolution data.

3.2 Increased Accessibility

Along with increased resolution is the potential to establish an invisible boundary at each RWIS. Drivers crossing the boundary could receive current information regarding the road conditions such as accidents, hazards, or bad weather along their assumed route. This concept combined with V2X communication sets the potential for cars to pre-empt drive settings to compensate for specific road conditions while informing the driver of situation ahead.

3.3 Driver Benefit

U.S. driver efficiency and safety will benefit from successful implementation of a road and weather condition communication system. Success at INL creates a roadmap for implementation across DOE sites and areas where weather has proven itself a driving hazard. Furthermore, as advances in autonomous vehicle functionality increase, the road and weather condition communication system can easily be adapted to provide the on-board vehicle intelligence with weather information that is shareable with the human operator.

3.4 Communication Technology

Plans also include development of a dashboard and streaming of real time information to drivers and dispatch centers on site. In 2014, ITD and INL executed a Cooperative Research and Development Agreement (CRADA) to develop connected vehicle technology for the ITD fleet vehicles and INL fleet vehicles. The progress to date includes deployment of four Vaisala DSP310 mobile road weather sensor sets, data integration within the Vaisala Navigator™ website, plus deployment of three Savari on-board units (OBU) and one Savari road-side unit (RSU). The CRADA research was enhanced to include the application of big data analytics and machine learning to the issue of microclimate prediction for hazardous weather conditions on INL highways.

4 Method

The basic ideas of numerical forecasting and climate modeling were developed about a century ago, long before the first electronic computer was constructed [5]. As Lynch notes weather forecasting was a haphazard process, imprecise and unreliable. That has changed with the introduction of General Circulation Models (GMC) based upon the upper air and air over the oceans. Scientists now have access to terrestrial, -sensor combinations and historical data to guide the analysis to near term conditions and for very specific locations. It should be noted that specific locations tend to be larger than the ones under investigation in the present study. Additionally; advances in computation have made short range, short duration forecasting possible [5].

The National Observation and Atmospheric Agency (NOAA) support a global forecast model (GFM) providing forecasts 6–8 days out worldwide. Our research effort is for a limited geographical area (<1 km) and for near term duration. The value of this approach lies in the ability to handle huge volumes of data and to do so at fast speed, allowing for near real time information processing. Artificial neural networks (ANN) are a subset of big data approaches used in meteorological studies. Shrivastava et al. [6] present a comprehensive review of the application of ANN model to meteorology and oceanography. Aggarwal et al. and others [7] have used the ANN model to predict maximum and minimum temperatures. Rani et al. report using the ANN approach employing a training algorithm to establish ANN weights used in predicting rainfall and drought for parts of India [8]. Huang and Ran used a back propagation method to train neural networks in the prediction of traffic speed in adverse weather conditions [9].

A major focus of the present research effort was to extend a historical weather and road condition model into a forecasting model. The results to date are substantiating that success. Three models were developed: a route performance model, road closure probability model, and a weather-derived travel index model. One of the novel aspects of this research has been development of two indexes: a Black Ice Index (including freezing fog) and a Snow Drift Index (SDI). Both are based upon a number of data sources including: RWIS, NWS, mobile data, historical crash data, route data collected via Zonar^{TM} tablets onboard all INL motor coaches, local geographic data, and bus speed data for the varied INL routes. Ambient air temperature, solar radiation, and wind were then integrated into the model to support algorithm tuning and validation. A prototype visualization dashboard, which integrates disparate data sources, analytical models, and ITD traffic cameras has also been developed. Black Ice and Snow Drift Indexes are inferred from other measures.

Figure 1 below presents the (SDI) model plotted against wind speed and snowfall. Typical values range from zero (no drifting snow) to one (ground blizzard conditions) with SDI $\geq$ 0.5 indicating severe driving conditions. While not shown, temperatures below 40 °F are required to cause blowing and drifting (Table 2).

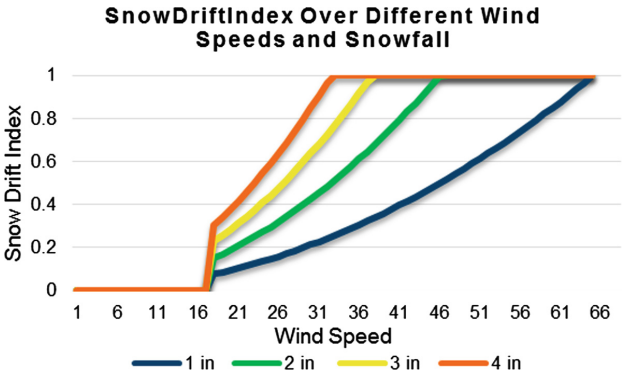


Fig. 1. Typical Snow Drift Index versus Wind Speed for multiple snowfall rates. While not shown, temperatures below 40 °F are required to cause blowing and drifting

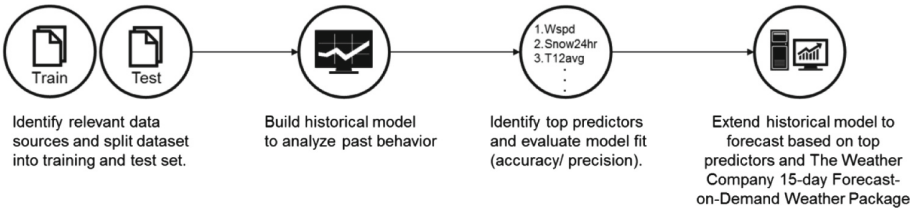


Fig. 2. Generalized high-level approach used for analytical modeling. This approach was used for creation of both route performance model and road closure probability model.

Because road closures on INL highways do not occur at a high frequency, historical closures of nearby interstates were used in support of model development. Zonar tracking systems onboard INL motor coaches were used as a source of driving speed for different road conditions. The speed data are available to INL dispatch at the end of each shift and can also be acquired near real time via back haul cellular.

5 Research Goals

Hazardous weather conditions remain a threat to driver safety as well national productivity, especially in the case of shipping [2]. To mitigate the damage hazardous weather can cause, INL in collaboration with IBM, is employing data-mining and machine learning methods to integrate a weather sensor network. The integrated network will provide accurate, real-time data sources to allow for near-real time weather forecasting as well as reliable and accurate near-term weather prediction with high-spatial resolution. That is, areas currently with low-spatial resolutions in current weather forecasts will be identified and accounted for in future weather forecasts, especially problematic areas determined by historical data.

Once high-resolution, reliable, and accurate forecasting capabilities are created and tested, a delivery platform will be employed to bring information to all entities who may

benefit from such information such as commuters, long-haul drivers, mass transit and road management agencies. The delivery platform will involve multiple communication avenues to include dashboard analytics to dispatch and management for big picture decisions. Safety alerts will target vehicles approaching hazardous weather. Targeted vehicles will receive driver alerts as well vehicle input to properly engage (or disengage) driving systems to mitigate risk on upcoming road conditions.

Both the dashboard display and the vehicle safety alerts will be evaluated to ensure each supports their intended audience as best as possible. General usability evaluations will suffice for the dashboard engagement. However, the driver safety alerts will require more rigorous testing to ensure driver safety is not compromised by incoming alerts. INL design of the in-vehicle display will comply with the Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices [10, 11]. Design guidelines for safety and information presentation will encompass various HFE concerns including: driver distraction, workload, consideration of secondary tasking, and general ergonomics such as the display distance and placement in relation to the driver position. Testing of the prototype design is planned for July 2017 and will be conducted in the INL heavy vehicle simulator. Standards such as the NHTSA eye glance-driving simulator protocol for glance (<2 s) will be followed.

6 Summary and Conclusion

Maintaining traffic flow and safety during severe weather is a national and world-wide concern. As part of the Idaho National Laboratory's goal to promote a safe work environment Mission Support Services is committed to establishing the capability to improve traveler safety along roads susceptible to hazardous weather. Partnering with IBM, MSS is part of a project to increase the reliability, the resolution, and the communication of road and weather conditions along popular routes. Success at INL means safer roads for local personnel and, in the grander scheme, provides a roadmap for any institution or government agency to develop the capability to improve road and weather condition forecasting and communication by informing communication centers and drivers and creating safer roads anywhere there is potential for hazardous weather.

This research represents an application of big data analytics to severe weather and road conditions that will further enable and improve fleet supervisor planning and decision making including: road closure decisions, fleet scheduling, and improving INL site traffic flow. It does so based on innovative features for snow drift and black ice prediction.

The approach taken integrates in near real-time, weather information from local sources including mobile data streaming sources and RWIS collection and integrates these data with other publically and privately available weather data sources.

The predictive analytics forecasting model being developed at INL in conjunction with IBM Global Business Services enhances the forecasting currently available through the National Weather Service and National Oceanographic and Atmospheric Administration and can be implemented in other parts of the country where severe weather is an ongoing concern.

The machine learning model developed in support of this research and the use of available historical weather and road condition data will be useful in predicting black ice conditions that pose considerable to INL fleet and private vehicles. The result of this year's testing and refinement is expected provide a scalable model allowing for the prediction of road and weather conditions for small areas less than one kilometer, most affected by severe weather conditions.

In conjunction with this work, INL is developing information presentation strategies for motor coaches and INL dispatch.

Acknowledgements. The authors acknowledge Daniel Yawitz and Reid Mechanick for their knowledge, technical expertise, and invaluable contributions to the creation of the analytical models and visualization platform.

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, nor any contractor performing in support of the U.S. Government, nor any employees or personnel of such contractor, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights, including, but not limited to copyright, patent or trademark rights. References herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

The information in this report is as accurate as possible within the limitations of the uncertainties of the basic data and methods used. The potential quantities presented in the report were determined analytically. Users need to ensure that the information in this report is adequate for their intended use. Battelle Energy Alliance, LLC and International Business Machines Corporation make no representation or warranty, expressed or implied, as to the completeness, accuracy, or usability of the data or information contained in this report, or results obtained.

References

1. Federal Highway Administration: How Do Weather Events Impact Roads? https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm (2017). Accessed 16 Mar 2017
2. Texas A&M Technology Institute: 2015 Urban Mobility Scorecard. https://mobility.tamu.edu/ums/report/ (2015)
3. Federal Highway Administration: Road Weather Information System Environmental Sensor Station Siting Guidelines Version 1.0April 2005. https://ops.fhwa.dot.gov/publications/ess05/index.htm (2005). Accessed 16 Mar 2017
4. Pisano, P., Murphy, R.: Road Weather Management—Weather-Savvy Roads. https://www.fhwa.dot.gov/innovation/everydaycounts/edc_4/roadweather.cfm (2016). Accessed 16 Mar 2017
5. Lynch, P.: The origins of computer weather prediction and climate modeling. J. Comput. Phys. **227**(7), 3431–3444 (2008)
6. Shrivastava, G., Karmakar, S., Kowar, M.K., Guhathakurta, P.: Application of artificial neural networks in weather forecasting: a comprehensive literature review. Int. J. Comput. Appl. **51**(18), 17–29 (2012)

7. Aggarwal, A., Kumar, V., Pandey, A., Khan, I.: An application of time series analysis for weather forecasting. Int. J. Eng. Res. Appl. ISSN: 2248–9622, **2**(2), 974–980 (2012)
8. Rani, B.K., Srinivas, K., Gavardhan, A.: Rainfall prediction with optimized TLBO applied to ANN. J. Sci. Ind. Res. **73**, 643–647 (2014)
9. Huang, S.H., Ran, B.: An application of Neural Network on Traffic Speed in Adverse weather conditions, in Transportation Research Board Proceedings. Annual Meeting, Washington, DC (2003)
10. National Highway and Safety Administration: Regulations.gov. https://www.regulations.gov/documentD=NHTSA-2010-0053-0009 (2010). Accessed 16 Mar 2017
11. Campbell, J.L., Brown. J.L., Graving, J.S., Richard, C.M., Lichty, M.G., Sanquist, T., Morgan, J.L. (2016). Human factors design guidance for driver-vehicle interfaces (Report No. DOT HS 812 360). Washington, DC: National Highway Traffic Safety Administration

A Study on the Human and the Automation in Automated Driving: Getting to Know Each Other

Eugen Altendorf(✉), Raphael Schütz, Yigiterkut Canpolat, Gina Weßel, and Frank Flemisch

Institute of Industrial Engineering and Ergonomics, RWTH Aachen University, Bergdriesch 27, 52062 Aachen, Germany
e.altendorf@iaw.rwth-aachen.de

Abstract. In recent years, advanced driver assistant systems (ADAS) and solutions for automated driving have been introduced by several automotive original equipment manufacturers (OEMs) and suppliers. Currently, these types of automation are designed for partially automated driving, but the step towards higher levels of automation can be expected to be made soon. One of the most commonly addressed use cases is driving on a highway such as the German Autobahn. In this paper, we propose an approach for adapting the automation's behavior to the human's driving preferences, providing a cognitive automation system with a machine-learning algorithm. This system has been implemented in a simulator for automated driving and has been used in a study addressing conditional automation. Within the presented experiment, typical situations for automated driving under varying conditions have been tested in the driving simulator. During cooperative human-machine driving, the automation can learn the human's preferences regarding relevant driving states.

Keywords: Human-systems integration · Automated driving · Human-machine systems · Driver-vehicle interaction · Cooperative driving · Neural networks · Machine learning · Cooperative automation · Cooperative guidance and control

1 Introduction

Advanced driver assistant systems (ADAS) and solutions for partially automated driving are being introduced into the automotive market. Moreover, customers are shifting their attention towards these technologies, which is reflected in higher take rates. Thus, partially automated driving is becoming a reality in today's, traffic, especially in certain driving situations. Although such systems have mainly been reserved for premium and luxury vehicles in the past, the increasing availability of feasible and cost-efficient information and communication technology facilitates a distribution of ADAS into mid-size and compact vehicles. At the same time, even more sophisticated automation systems are introduced to the automotive market's high end range, making automated driving not just a future trend anymore. Typically, such systems are mainly designed for the use in more standardized environments, such as

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_65

highway driving. Usually, these driving situations yield reduced complexity in comparison to urban scenarios due to traffic laws and restrictions. Nevertheless, vehicles equipped with advanced assistance and automation technology are not the only agents in today's and tomorrow's traffic. A diverse mix of substantially different and technically capable systems combined with manually driven cars and other often vulnerable traffic participants such as pedestrians can be expected on the roads. With more or less technologically advanced vehicles and humans on bikes and on foot interfering with each other, highly complex driving situations will arise. In such ambiguous situations, the capabilities of a technical system can be exceeded. Therefore, the human cognition or flexibility might be needed as a part of the solution. Thus, in assisted and automated driving, the human driver and the technical system form a joint, cooperative human-machine system.

2 Automated and Cooperative Driving

2.1 Vehicle Guidance

Vehicle guidance and control conducted by human drivers, consists of all processes which are required for operating a vehicle in traffic. According to the definitions of [1] these processes include the detection, recognition and classification of objects and events, appropriate responses to these objects and events, maneuver planning, control of longitudinal and lateral vehicle dynamics, and enhancing conspicuity.

[2] introduces a definition of vehicle guidance and control and describes it as a target-oriented sensumotoric activity in a three-layered model consisting of the layers navigation, guidance, and control. These layers relate to different subtasks, which occur during the driving task and can be categorized in different levels of human behavior [3], introduced by the model of [4]. This model differentiates between skill-based behavior, which is conducted automatically and makes the execution of a secondary activity affordable, and rule-based behavior, which is performed during frequently occurring situations. Additionally, knowledge-based behavior is defined as the actions, applied for complex and unknown situations.

The insights of this model can be transferred to the model by [2]. The first layer of the cognitive model introduced by [2] is the navigational level. The process on this level can be considered as knowledge-based behavior and is not part of the dynamic driving task, but can be described as a planning process within the driving task. The dynamic driving task also includes guidance and control levels. The guidance level comprises the selection of maneuvers, whereas the control level includes the control of longitudinal and lateral vehicle dynamics. On both levels, ruled-based or skill-based behavior can be required, depending on the experience of the driver, her condition, and the driving situation.

[5] describes a model related to the model by [2], organizing the levels along strategic, tactical and operational driver efforts. The strategic effort consists of planning destination and routes, whereas the tactical effort describes driving maneuvers. Operational effort includes reactions to the lateral and longitudinal dynamics for controlling the vehicle.

[6] extends these models to a four-layered model, where the guidance level is divided into a maneuver planning and a trajectory planning level. In the context of cooperative guidance and control of a vehicle, the automation has to consider this cognitive architecture defined by the described four-layered model of the human driver, in order to achieve high compatibility. Thus, the automation has to be designed using similar levels.

2.2 Levels of Assistance and Automation

In human machine interaction, different levels of automation can be depicted as a spectrum or continuum ranging from manual to full automation control [7, 8]. This concept has also been applied to the automotive domain, where the allocation of vehicle control between the human driver and the automation may vary. However, the notion of a two-dimensional spectrum encompasses a simplification for modelling the complex reality in terms of human automation cooperation. In regard to this, [9] discusses the relations between the dimensions of control, ability, authority and responsibility on each level of the automation spectrum.

[10] introduces a definition of automation degrees, which combines the levels of automation and stages of human information processing. The four stages of information processing include sensory processing, perception/working memory, decision making and response selection [11]. Thus, when combining the two models in order to develop degrees of automation, a process is incorporated and therefore a time-related component is introduced to levels of automation [8, 12, 13]. Additionally, [9] extends this concept by articulating the relations between the dimensions control, ability, authority and responsibility on each level of the automation spectrum.

The dimensions of automation degrees for automated driving consist of the automation level, responsibility, authority for transitions, capability for autonomy, and stages of information processing. Since this definition probably does not cover all dimensions of automation, in many circumstances the introduced dimensions are cumulated to unidimensional models for reducing complexity [8, 13].

The execution of the dynamic driving task and the responsibility are also considered as one single dimension by the definition of the German Federal Highway Research Institute (BASt) [14]. As described in the SAE taxonomy [1], in the partial automation level (level 2), the automation controls the longitudinal and lateral vehicle dynamics, whereas the driver has to monitor the environment and has to be able to take over the vehicle control directly, as it is required in case of automation failure or limitations.

In this paper, the scope ranges from assisted to partially automated driving, thus the driver and the automation are conducting parts of the driving task simultaneously.

2.3 Cooperative Guidance and Control

In terms of human-machine systems, cooperativeness describes the compatibility of the partners' behaviors and interactions [15]. Therefore, the behavior and architecture of

the automation should be complementary to the needs of the human [16]. To achieve this kind of cooperativeness, a minimum level of interaction between the partners must be enabled for e.g. resolving conflicts. Furthermore, the intentions and capabilities of the partners have to be predictable and traceable [17]. For this purpose, a complementary perception of the environment and complementary representations of information and action are required. Additionally, the partners should share a common understanding of the current situation, since they are not able to cooperate appropriately without the same interpretation of the present situation and future scenarios [18].

A concept related to cooperativeness is the concept of shared control. Shared control denotes operational actions, which have immediate effect on the mutual task of the partners. The longitudinal dynamics of a vehicle can be controlled by the driver and the automation collectively, while the two partners are connected to each other by haptic interfaces (e.g. accelerator pedal) [15, 17, 19]. In line with this, shared control has no impact on the tasks on the tactical and strategic/navigational level [20]. Hereby, shared control can be considered to be a "sharp end" of cooperation [21].

3 Addressed Use Cases and Experimental Data

3.1 Driving Situation

The considered driving situations are depicted in Figs. 1 and 2. These use cases are framed on a highway such as the German Autobahn. Here, at least two ongoing lanes can be used by the vehicles. All lanes are strictly separated from oncoming traffic. The experimental data has been created after introducing the participants to certain regulations. Hereby, we refer to German traffic law, and especially the *Rechtsfahrgebot*. This obligation requires traffic participants to use the right-hand lane of the road, whenever possible. When driving on multiple lanes, a vehicle is supposed to change to the furthermost right lane. Thus, when overtaking another vehicle, the overtaking car should change back to the right lane after the other vehicle has been passed.

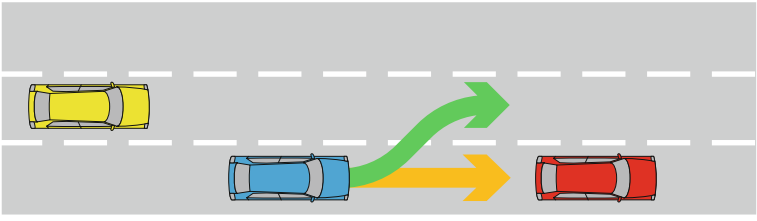


Fig. 1. Considered driving situation 1: The upcoming slower vehicle (right side, red) can be overtaken by changing to the left lane with faster vehicles (left side, yellow). Alternatively, the ego-vehicle (center, blue) can stay on its lane, adjust the velocity and follow the leading car

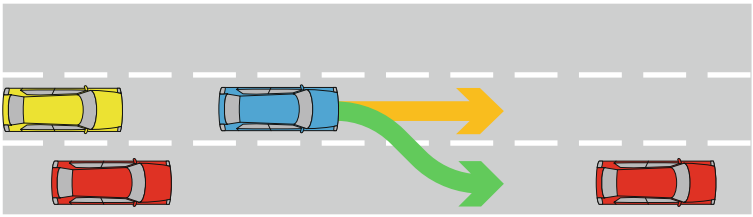


Fig. 2. Considered driving situation 2: A faster vehicle (left side, yellow) is approaching the ego-vehicle (center, blue). Possible maneuvers are a lane change to the right lane or staying on the current lane. The right lane is occupied by slower vehicles (red).

3.2 Experimental Data

For training the artificial neural networks, data from a simulator study was used (cf. [22]). The experiment with 33 participants was conducted using a static driving simulator at the Institute of Industrial Engineering and Ergonomics (IAW) at RWTH Aachen University. In this study, the participants were supported by an automation system and drove in a conditionally automated driving mode. They could initiate driving maneuvers using an active side stick, while the automation would execute the driving task on the stabilization and trajectory planning level. In the described driving situation with a slower car in front (Fig. 1), the vehicle would, if possible, automatically change the lane if and when the corresponding maneuver was initiated by the human driver. Otherwise, the ego-vehicle would reduce its velocity and follow a leading vehicle, if no other maneuver was chosen. During the second driving situation (Fig. 2), the automation would stay on the chosen lane until the driver would initiate a lane change maneuver. The participants could set a speed guide for the automation, comparable to an adaptive cruise control system (ACC). All driving data was logged with a frequency of 10 Hz. The behavioral model for the ongoing traffic was self-implemented and is based on existing car-following models (such as [23, 24]).

4 Assessing the Driver's Intention Using Neural Networks

4.1 Theoretical Background

A neural network, sometimes also called artificial neural network, is a structure of processing elements which are connected using unidirectional channels [25]. Such a network can consist of several layers. These layers are made of neurons and arcs. A simple network usually consists of three layers: an input layer linking the data to the neurons, one hidden layer and an output layer. The latter links the neurons to the outputs. While the neurons can be compared to nerve cells in a biological, the arcs represent the axons and therefore the connection to other cells [26].

Neural networks are widely used in pattern recognition applications as they can handle a high number of inputs and samples. In most cases, a neural network is pretrained on a specific data set and can then be utilized in the current state or can be further adapted to each new data sample it receives. This gives the possibility to

increase robustness as well as adapt the network to one person, for example such that the network learns about the person's behavior [27].

Applied to lane change maneuvers in manual driving, [28] uses the neural network approach to model the decision process of individual drivers. In this study, a neural network with a single hidden layer and back propagation training has been applied. Hereby, information about the environment was fed to the neural network as input. The network predicted the new lane and position of the ego-vehicle as outputs. An interactive driving simulator was used to collect data for the training process of the network.

4.2 A Neural Network for Cooperative Guidance and Control

The developed neural network consists of an input layer accepting 27 to 35 values depending on the number of lanes. Hereby, the network's inputs contain various parameters such as time-to-collision and time headway as well as the current lane and surrounding vehicles' states (relative position, velocity).

For reasons of simplicity, the network features a single hidden layer with a variable number of neurons. For achieving a high fit of the network, the exact number of neurons is determined by an optimization algorithm. The output layer can support three to five outputs, representing the decisions of the driver-vehicle system. The three main and two additional outputs are shown in Table 1.

Table 1. Output states of the neural network.

Output index	Driving action	State type
1	Initiating change to the left lane	Main
2	Proceeding lane change (to the left)	Additional
3	Initiating change to the right lane	Main
4	Proceeding lane change (to the right)	Additional
5	Keeping the current lane	Main

All outputs are states which are defined using a numeric region of 9 logging steps such that the network predicts the driver's actions within an interval of roughly 1 s.

The network is designed to be pre-trained on an asset of 33 participants delivering around 1,450 overtaking maneuvers. For the training process, a total of 500,000 logging entries were used. This pre-trained network can be further adapted to new drivers, which can be done in real-time, for example during an experiment in the same driving simulator.

In order to accommodate for the significantly higher amount of "state 5" logging entries, i.e. the driver stays on his lane; only ten state 5 logging steps were used between every lane. Hereby, the steps were chosen with a uniform distribution regarding time.

4.3 Finding an Optimal Network

The network parameters are optimized by using a genetic algorithm provided by the MATLAB optimization toolbox. The chosen algorithm optimizes the number of neurons of the hidden layer and the maximum validation fails which is a key parameter for the training process and prevents overtraining if set correctly. However, this parameter is difficult to fix for networks with a large training data set as in the presented case. If the value is too low, the training is stopped early which results in poor accuracy. If the value is too high, this leads to overtraining and poor behavior on sets other than the training set. Training set and validation set are set with 80% training, 20% validation data, and are randomly drawn from the full data set. This implies a different set of data for validation and training on every function call. Hereby, a function call done by the genetic algorithm is a full-scale training process of a newly created network. During the early development of this algorithm it became obvious that the choice of validation and training data set has significant influence on the network's performance. While for some function calls, and thus, trainings, the performance of recognizing state one and five is close to 90%, for other calls it could remain as low as 60%. This applies to validation and training data set so that overtraining can be excluded as a reason for this phenomenon.

In order to find the optimal parameterization, the training algorithm was called five times for each parameter set. Within the optimization, two different approaches of defining the objective function's value, and therefore the genetic algorithm's fit, were implemented. Either the minimum performance or the mean performance was used for the optimization. Hereby, performance is defined as the cross entropy of the trained network, and this value is the lower, the better the network performs.

The training and optimization of the neural network (Fig. 3) was performed in MATLAB, which offers several neural network structures and training methods in the neural networks toolbox and the global optimization toolbox [29, 30]. In order to speed up optimization, the calculation was performed on the high-performance cluster of

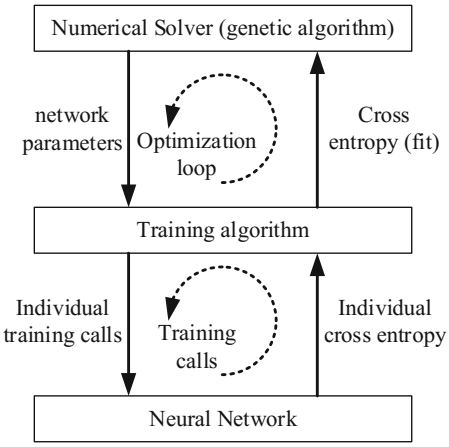


Fig. 3. Optimization approach.

RWTH Aachen, which features batch systems with several hundred cores. For each optimization run, 12 CPU cores and 8 GB RAM have been used.

Live learning and adaptation to a new driver was implemented in MATLAB as well using the Robotics Toolbox which allows the receiving of ROS messages. As the simulator employs ROS messages for internal communication, this provides a simple way of live learning. Figure 3 illustrates the optimization procedure.

5 Results

The optimization shows that an optimal network in the presented case has between 70 and 100 hidden neurons. Figure 4 shows that the minimum cross entropy decreases to 70 neurons. For a higher number of hidden neurons, performance does not improve. However, calculation time rises significantly as the networks grow and therefore takes more time for training. This effect is highly contra-productive for the purpose of real-time adaption. Furthermore, the amount of redundancies will increase as multiple neurons work similarly. The worst performance after training is not influenced by the number of neurons and stays on the same level for the whole range of network sizes. Only for very high numbers of neurons an even higher level of high cross entropy can be observed. Compared to this, the number of maximum validation fails does not significantly influence network performance.

Fig. 4. Cross entropy over the number hidden neurons. Dots represent individual training results.

Figure 5 shows, that below 700 maximum validation fails there are three performance levels while above 700 fails a fourth level is reached. As stated before, these performance levels are due to the random selection of training data set and validation

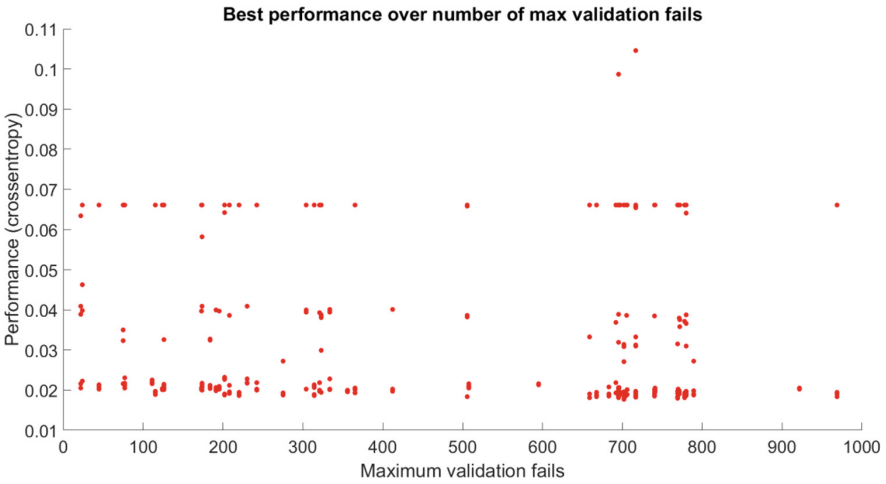


Fig. 5. Cross entropy over the number of maximum validation fails in the training process. Dots represent individual training results.

data set. Performance always indicates the overall performance, but depending on which subset of the entire set was employed for training, the training either succeeds or fails independent of the network parameters.

Accordingly, for reasons of simplicity, a network with 70 hidden neurons and training with 100 as number of maximum validation fails was employed for further calculations.

Figure 6 (left) shows the confusion matrix of the network trained on data from all subjects. Maximum performance of recognition rates can be seen: 89% for state 1;

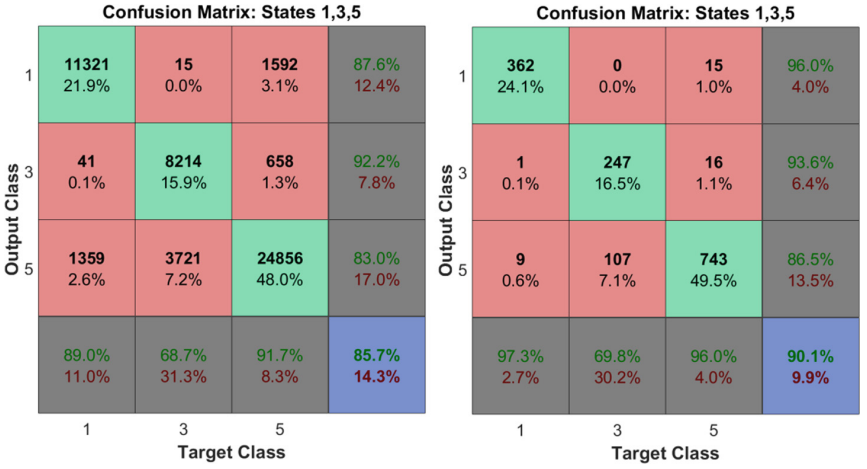


Fig. 6. *Left*: Confusion matrix (cf. [29]) over all subjects showing the recognition of state 1; 3 and 5. *Right*: Confusion matrix for subject 23 showing the recognition of state 1; 3 and 5.

68.7% for state 3 and 91.7% for state 5. This means that a general neural network is able to predict the start of an overtaking maneuver within one second in nine out of ten occasions. However, this is the result for 33 study participants without any distinction regarding the individual driving style. In case of a single driver, the network can fully adapt to individual behavior and can therefore achieve even higher accuracies. For the tested data set, accurate predictions of more than 97% for the initiating process of an overtaking maneuver could be achieved. Moreover, the overall accuracy regarding all states defined in Table 1 was higher than 90% (Fig. 6, right). The mean accuracies for each of the 33 participants are 97.5% for driving action 1, 98.4% for driving action 2, 89.8% for driving action 3, and 90.8% overall.

In automated driving, the human driver and the automation form a joint human-machine system. This driver-vehicle system disposes of a joint decision making process, which is based on the situation assessment of both, the driver and the technical system [18]. By using typical patterns, the automation can learn the driver's preferences regarding certain aspects of the dynamic driving task. Thus, a system that knows its driver can derive decisions which are compatible to the human's situation assessment.

6 Discussion and Outlook

In automated driving, a spectrum of different automation levels can be used. Currently, different implementations of partially automated vehicles are being developed and offered. Typically, these automations are capable of conducting the driving task for certain use cases, such as highway driving. These scenarios yield reduced complexity in comparison to other scenarios due to traffic laws and restrictions. Nevertheless, vehicles equipped with ADAS or automation technology are not the only participants in today's and tomorrow's traffic. A diverse mix of substantially different and technically capable systems combined with manually driven cars and other sometimes vulnerable traffic participants such as pedestrians can be expected to form a highly complex traffic. In ambiguous situations, the capabilities of a technical system can be exceeded. Therefore, the human cognition or flexibility might be needed to compensate for technical limits. Thus, in assisted and automated driving, the human driver and the technical system form a joint, cooperative human-machine system. When analyzing the decision-making process of a human-machine system, not only the technical system, but also its human counterpart play a role.

In this paper, we propose an approach for adapting the automation's behavior to the human's driving preferences, by augmenting a cognitive automation system with a machine-learning algorithm. This system has been implemented in a simulator for automated driving and has been used in a study addressing partial and high automation modes. Within the presented experiment, typical situations for automated driving under varying conditions have been tested in the driving simulator. During cooperative human-machine driving, the automation learned the human's preferences regarding driving states such as distance, velocity, and headway and adapted its behavior accordingly. Using a machine-learning approach, the human's decisions regarding driving maneuvers can be learned and therefore anticipated.

Acknowledgments. The research conducted was partly funded by the Deutsche Forschungsgemeinschaft (DFG) within the projects "*Arbitration of cooperative movement for highly-automated human machine systems*", "*Systemergonomics for cooperative interacting vehicles*" and partly by *RWTH Aachen University*.

References

1. SAE International: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. SAE International (2016)
2. Donges, E.: Aspekte der aktiven Sicherheit bei der Führung von Personenkraftwagen. AUTOMOB-IND 27 (1982)
3. Donges, E.: A conceptual framework for active safety in road traffic. Veh. Syst. Dyn. **32**, 113–128 (1999)
4. Rasmussen, J.: Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Trans. Syst. Man Cybern. **3**, 257–266 (1983)
5. Michon, J.A.: A critical view of driver behavior models: what do we know, what should we do? In: Human Behavior and Traffic Safety, pp. 485–524. Springer, New York (1985)
6. Loper, C., Schomerus, J., Brandt, T., Flemisch, F., Sattel, T.: Bahnplanung, Bahnfuehrung und haptische Interaktion fuer ein Fahrerassistenzsystem zur Querfuehrung/Path planning, vehicle guidance and haptic interaction for a driver assistance system for lateral guidance. VDI-Berichte (2006)
7. Sheridan, T.B., Parasuraman, R.: Human versus automation in responding to failures: an expected-value analysis. Hum. Factors **42**, 403–407 (2000)
8. Sheridan, T.B., Verplank, W.L.: Human and computer control of undersea teleoperators (1978)
9. Flemisch, F., Heesen, M., Hesse, T., Kelsch, J., Schieben, A., Beller, J.: Towards a dynamic balance between humans and automation: authority, ability, responsibility and control in shared and cooperative control situations. Cogn. Technol. Work **14**, 3–18 (2012)
10. Wickens, C.D., Li, H., Santamaria, A., Sebok, A., Sarter, N.B.: Stages and levels of automation: an integrated meta-analysis. Proc. Hum. Factors Ergon. Soc. Annu. Meet. **54**, 389–393 (2010)
11. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. IEEE Trans. Syst. Man Cybern. Part A **30**, 286–297 (2000)
12. Manzey, D., Reichenbach, J., Onnasch, L.: Human performance consequences of automated decision aids: the impact of degree of automation and system experience. J. Cogn. Eng. Decis. Mak. **6**, 57–87 (2012)
13. Onnasch, L., Wickens, C.D., Li, H., Manzey, D.: Human performance consequences of stages and levels of automation: an integrated meta-analysis. Hum. Factors **56**, 476–488 (2014)
14. Gasser, T.M., Arzt, C., Ayoubi, M., Bartels, A., Bürkle, L., Eier, J., Flemisch, F., Häcker, D., Hesse, T., Huber, W., et al.: Rechtsfolgen zunehmender Fahrzeugautomatisierung. Berichte der Bundesanstalt für Straßenwesen, Unterreihe Fahrzeugtechnik (2012)
15. Flemisch, F.O., Adams, C.A., Conway, S.R., Goodrich, K.H., Palmer, M.T., Schutte, P.C.: The H-Metaphor as a guideline for vehicle automation and interaction (2003)
16. Bubb, H.: Ergonomie. 3. Aufl., Kap. 5.3 Systemergonomische Gestaltung, S. 390–420. Carl Hanser Verlag, München (1993)

17. Flemisch, F.O., Bengler, K., Bubb, H., Winner, H., Bruder, R.: Towards cooperative guidance and control of highly automated vehicles: H-Mode and Conduct-by-Wire. Ergonomics **57**, 343–360 (2014)
18. Altendorf, E., Weßel, G., Baltzer, M., Canpolat, Y., Flemisch, F. Joint decision making and cooperative driver-vehicle interaction during critical driving situations. i-com **15**, 265–281 (2016)
19. Altendorf, E., Baltzer, M., Heesen, M., Kienle, M., Weißgerber, T., Flemisch, F.: H-Mode. In: Winner, H., Hakuli, S., Lotz, F., Singer, C. (eds.) Handbook of Driver Assistance Systems: Basic Information, Components and Systems for Active Safety and Comfort, pp. 1499–1518. Springer International Publishing, Cham (2016)
20. Abbink, D.A.: Neuromuscular analysis of haptic gas pedal feedback during car following. Behaviour **65**, 1 (2006)
21. Flemisch, F., Abbink, D., Itoh, M., Pacaux-Lemoine, M.-P., Weßel, G.: Shared control is the sharp end of cooperation: towards a common framework of joint action, shared control and human machine cooperation. In: 13th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems. Kyoto, Japan (2016)
22. Altendorf, E., Schreck, C., Flemisch, F.: A new method and results for analyzing decision-making processes in automated driving on highways. In: Advances in Human Aspects of Transportation, pp. 571–583. Springer, New York (2016)
23. Treiber, M., Hennecke, A., Helbing, D.: Congested traffic states in empirical observations and microscopic simulations. Phys. Rev. E **62**, 1805 (2000)
24. Wiedemann, R.: Simulation des Straßenverkehrflusses. Institut für Verkehrswesen Univ. Karlsruhe, Karlsruhe (1974)
25. Hecht-Nielsen, R. (ed.): Theory of the Backpropagation Neural Network. IEEE Press, New York (1989)
26. Hopfield, J.J.: Artificial neural networks. IEEE Circuits Devices Mag. **4**, 3–10 (1988)
27. Kumagai, T., Sakaguchi, Y., Okuwa, M., Akamatsu, M.: Prediction of driving behavior through probabilistic inference. In: Proceedings of the 8th International Conference on Engineering Applications of Neural Networks, pp. 117–123 (2003)
28. Hunt, J.G., Lyons, G.D.: Modelling dual carriageway lane changing using neural networks. Transp. Res. Part C **2**, 231–245 (1994)
29. MATLAB: Global Optimization Toolbox User's Guide Version 3.4.2. The MathWorks, Inc. (2017)
30. Beale, M., Hagan, M., Demuth, H.: Neural Network Toolbox™ User's Guide (2017)

How Automation Level and System Reliability Influence Driver Performance in a Cut-In Situation

Jonas Radlmayr[1(✉)], Veronika Weinbeer[2], Cathrine Löber[2], Mehdi Farid[2], and Klaus Bengler[1]

[1] Chair of Ergonomics, Technical University of Munich, Boltzmannstraße 15, 85748 Garching, Germany
jonas.radlmayr@tum.de
[2] BMW Group, 80788 Munich, Germany

Abstract. This study evaluated the effect of different levels of automation and system reliability on driver performance in a cut-in motorway situation. The study was conducted in the dynamic driving simulator of the BMW Group with 65 participants. Driving data and an electrocardiogram (ECG) were recorded. The recorded driving data confirmed an automation effect. Concerning the type of automation, drivers' response was found to be significantly faster in the manual driving condition than in the partial automated driving condition. Moreover, significant differences were found for the minimum time to collision (TTC) and the maximum deceleration in favor of the manual driving condition. In addition, it was confirmed that system reliability does influence the driver performance in a critical situation. The findings indicate that higher system reliability causes a delay in the brake response time, higher maximum deceleration and lower minimum TTC. Additionally, a close correlation could be seen between the TTC and the subjective criticality. The ECG data did not show effects.

Keywords: Automation level · Vehicle automation · System reliability · Cut-in situation · Driver performance

1 Introduction

Research on automated vehicles and related human factors questions concerning safety and comfort have seen a high increase in the past years. This work focuses on partly automated driving. Partly automated driving constitutes the system taking over longitudinal and lateral control of the driving task while the driver has to monitor the system. In case of system limits, the driver has to regain control of the manual driving task within a short period of time. Consequently, the driver should be alert and vigilant during the partly automated drive. The definition of partly automated driving as Level 2 (L2) (Level 0: manual driving, Level 1: assisted driving) is taken from the SAE [1].

This work focuses on the effect of automation and system reliability on driver performance in highly critical situations. Cut-in situations represent a scenario that can lead to highly critical situations in case drivers notice the merging vehicle too late. Also, current sensor technology typically has problems noticing a lane change from a merging

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_66

vehicle in due time for vehicles in the near vicinity of the ego-vehicle [2]. Figure 1 shows the cut-in situation as depicted in the simulation. The picture is taken from the control booth of the simulator and does not show the field of view of the participants.

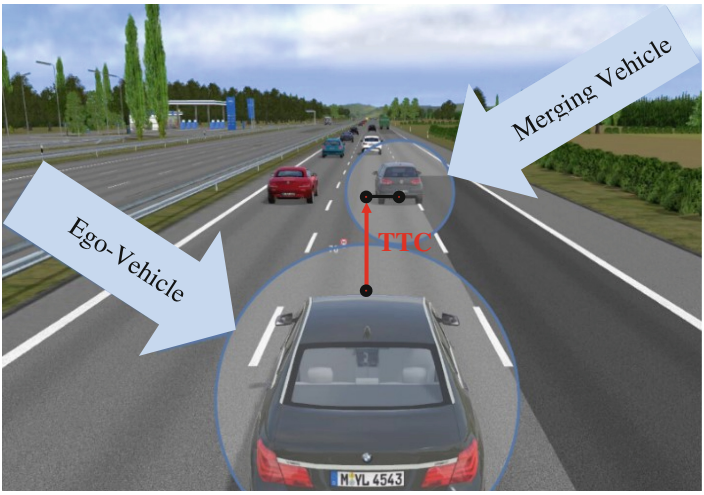


Fig. 1. Cut-in situation including the ego-vehicle and the other vehicle engaging in the lane change. The picture is taken from the control screen of the simulator. The TTC analyzed in the result chapter is displayed.

Manual drivers have to react to vehicles cutting into their lane appropriately in order to avoid accidents or critical situations. In case of partly automated driving, drivers have to be alert and potentially take over control in order to avoid critical situations or accidents as well. These reasons emphasize the choice of the cut-in situation as critical scenario suited for evaluating the influence of automation and system reliability on driver performance. The cut-in situation as shown in Fig. 1 including the merging vehicle and adjusting various parameters of the situation allows for a wide range of possible scenarios. The parameter automation is operationalized as different levels of automation using the SAE definition [1]. In case of partly automated driving, the driver has to monitor the system. [3] showed that monitoring tasks are significantly influenced by the system reliability.

Consequently, differences in driver performance at the same level of automation (L2) could occur due to a change in system reliability. Thus, system reliability was chosen to be the second factor for this study.

This work investigates the influence of automation and system reliability on driver performance by analyzing a number of metrics that will be explained in more detail in the hypothesis and parameters chapter. Since the influence of automation and system reliability have been studied extensively in a wide field, major models and methods concerning human behavior and information processing will be presented. Apart from having positive effects concerning workload, automation has been known to negatively affect human performance. Overreliance on automated systems including a loss of skills

due to automation have been shown to be significant [4, 5]. In the case of partly automated driving, these findings are highly relevant, since drivers relying on automation will possibly develop a temporary overreliance and might lose their manual driving skills. In case of long monitoring tasks, studies have shown a highly significant decrease in vigilance and arousal [6]. In case of partly automated driving where the driving task is executed by the automation but drivers have to monitor and take-over with little time in case a system limit is reached, this decrease in vigilance and arousal could account for a decrease in driver performance in a take-over situation. Consequently, this could lead to an increase in criticality of the situation. [7] showed the link between low vigilance and an increase in fatigue adding to the deterioration of driver performance.

In case monitoring performance keeps decreasing due to low vigilance and arousal, a loss of situation awareness [8] can be witnessed. In combination with the model of human information processing [9] and related resources necessary to react appropriately, a take-over necessary due to a merging vehicle cutting constitutes a highly critical scenario. The effect of automation operationalized by using partly automated driving provides a valid situation to analyze related effects.

The presented literature so far focuses on the variable automation. As mentioned above, system reliability has been known to affect monitoring since a higher system reliability has been shown to increase trust [3].

Complacency depends on different factors. In this context the relation between reliability and the comprehensibility of the automation is of a great importance. Referring to partly automated driving and the necessary monitoring, a change in system reliability could affect driver performance in cut-in situations. In addition, [10] looked at different automation levels and results showed significant differences between manual and partly automated drivers concerning minimum time headway and brake pedal reaction time when a merging vehicle moved into the ego lane. Participants were additionally warned acoustically. Results underline the feasibility of this work's design of experiment.

Eye-Tracking was used in this work to allow for an analysis of visual attention and the allocation of gazes throughout manual driving and during monitoring phases. Results will not be presented in this paper due to data quality. In order to assess changes in driver state objectively, data from an electrocardiogram (ECG) was recorded. While various studies assess changes in ECG parameters differently considering vigilance, arousal, workload or stress, [11] showed significant changes before and after a 90-minute drive in ECG parameters. They credited these changes to a growing mental fatigue due to the simulated driving.

The design of experiment considered the models, methods and findings described in the introduction in order to allow for valid findings concerning the effect of automation and system reliability on driver performance in cut-in situations.

2 Method

The experiment was conducted in the high-fidelity driving simulator of the BMW Group in the research and development center in Munich. A full vehicle mockup is moving inside a dome that is mounted on a hexapod allowing for the dynamic simulation. Seven

projectors provide an almost fully immersive field of view including front, side, back, side and interior mirror. The system allows for precise recording of all relevant driving parameters that are needed for analysis. In addition, the head-mounted eye-tracking system *Dikablis* is used in the study. The eye-tracking results will not be presented in this publication due to data quality.

For recording of the ECG parameters the system *eMotion-Faros* (180°) was used. Sampling frequency was 500 Hz.

In accordance to the choice of two independent variables, level of automation and system reliability constitute a two-factor design of experiment and a between-subjects design. Table 1 shows the design of experiment constituting the three groups. The independent variable Level of automation features manual drivers and partly automated driving (Level 2). The independent variable reliability constitutes the automation (L2) controlling the vehicle to differ between a reliable and an unreliable specification.

Table 1. Design of experiment with the three groups. Unreliable, manual drivers were considered to be not feasible.

Design of experiment		System reliability	
		Reliable	Unreliable
Level of automation	Level 0	Manuel drivers	X
	Level 2	Group 2	Group 1

The condition manual drivers and an unreliable test condition was determined to be not feasible since manual drivers in simulator studies typically try to drive as reliable as possible if not instructed otherwise.

Since driver performance was evaluated in one critical cut-in scenario, a power analysis projected very little significant effects with a hypothesized mid-range effect size. Consequently, the number of relevant scenarios for analysis was increased. Since time in the simulator was a limiting factor, more participants would not have helped significantly. Participants drove the same track twice with different characteristics instead, leading to a total of 6 different groups. Table 2 shows the distribution of participants to the different study groups. Using that design of experiment, a total of 120 scenarios can be used for analysis, with 40 scenarios for the three relevant conditions manual drivers, unreliable automation (Group 1) and reliable automation (Group 2).

Table 2. Study groups that drove the same track with different characteristics. For each group and track section the number of participants was $n = 10$.

Study group	Track section 1	Track section 2
1	Group 1	Manual drivers
2	Manual drivers	Group 1
3	Group 2	Manual drivers
4	Manual drivers	Group 2
5	Group 1	Group 2
6	Group 2	Group 1

For the complete track, participants were told to follow a lead vehicle in the middle lane and keep a constant speed of 90 km/h. In order to allow for comparable scenarios between manual and partly automated drivers, the lead vehicle was programmed to keep a 4 second time gap to the ego vehicle. Every participant experienced a total of six cut-in situations during the experiment. The first five were uncritical in comparison to the last one. Merging vehicles cutting into the middle lane all used their indicators and concerning the duration of the lane change and their relative velocity to the ego vehicle were classified as unmotivated following the definition of [12]. All cut-in situations happened on a straight part of the test track. The critical last cut-in situation was defined to have a duration of the lane-change of seven seconds with a relative velocity of 30 km/h. Obviously, the merging vehicle cutting in was slower than the ego vehicle. The left lane was blocked by other vehicles so drivers could not avoid a potential accident in the critical scenarios by changing lane but had to brake. The controller in the partly automated drive was identical, only the system reliability changed. Participants were free to drive hands-on or hands-off. No take-over requests were issued.

In Group 2 with the reliable system, the participants experienced a total of five scenarios, where the automation was able to cope with the situation, followed by the critical scenario. In Group 1, the unreliable condition, participants first experienced a situation in where the system detected the merging vehicle in time and adjusted the speed accordingly. After that, participants experienced a situation in which a reaction from the driver was necessary to prevent an accident and so forth. Thus, the five situations before the critical last scenario included two system failure or limits, where participants had to react. In addition, participants in Group 1 experienced a steering moment to the left in order to manipulate the whole system reliability and not just rely on cut-in situations where participants had to react as well. It can be questioned, whether this manipulation of the general system reliability was successful. The procedure was chosen for reasons of operationalization. The objective system reliability was defined as the percentage of relevant situations (e.g. total number of cut-in situations, take-over situations, etc.) where no system failures or limits were met.

$$\text{System Reliability} = n_{\text{no failures/limits}} * \frac{100}{n_{\text{total}}}(\%) \tag{1}$$

The formula shows the objective system reliability as described. This definition was used to change the system reliability between unreliable and reliable in the experiment. Consequently, before the critical last scenario, participants in Group 1 experienced a total of five cut-in situations with two system failures and one out-of-order steering moment in L2 driving. This leads to an objective system reliability of 50%. In Group 2, all cut-in situations were handled by the system, leading to a perceived system reliability of 100% before the final critical scenario. Throughout the experiment, the system issued no warning or offered automatic braking. The participants had to detect all system limits and/or failures by themselves.

2.1 Experimental Setup and Realization

Participants completed a demographic questionnaire before being introduced to the study and the simulator. Since study groups 5 and 6 (Table 2) experienced the track twice and the factor system reliability was changed, participants were not told that the system reliability will be changed but that they would experience two generally different conditions of partly automated driving. It was hypothesized, that without mentioning the system reliability participants would not just focus on that aspect but rather perceive a changing system reliability more realistically.

Participants were briefed on the questionnaire that were presented to them during the experiment, the ECG system, the eye-tracking system and the simulator were explained. They were also instructed that they were at all times responsible for the safety of the drive. Secondary tasks during the study were forbidden. Participants were told to park the car on a parking lot, close their eyes and relax. The baseline ECG was recorded. After that, a familiarization drive of five minutes of manual driving was instructed. Different participants from study groups 1 – 6 were permutated to avoid an influence of the circadian cycle and different arousal levels throughout the day on the ECG data. The partly automated driving was introduced to the relevant participants after the familiarization drive including the first scenario which was handled by the system for both groups. After each critical scenario at the end of one track, participants were asked to park the car briefly and assess the criticality of the last situation. Also, they were asked to report their workload using a scale similar to the Rating Scale Mental Effort (RSME) [13]. Questions on the perceived system reliability were instructed to relate only to the last section, without the critical situation. After that, participants experienced the track a second with an identical procedure. The whole experiment lasted roughly 70 min for every participant.

2.2 Hypothesis and Parameters

The independent variables are the level of automation and the system reliability. The level of automation was altered between manual driving and partly automated driving. The system reliability was changed by including system limits (two cut-in situations that could not be handled) and system failures (steering moment during partly automated driving). Before the critical last scenario, depending on the condition, an objective system reliability of 50% and 100% could be realized.

The dependent variables were maximum negative, longitudinal acceleration, or maximum deceleration, maximum lateral acceleration, brake pedal reaction time and time to collision. Since the merging vehicle cutting in from the right lane would present no possibility to calculate a TTC in the beginning of the lane change maneuver, the TTC in this study was defined as the time needed to close the distance projected between the front bumper of the ego vehicle and the rear bumper of the merging vehicle as depicted in Fig. 1. Also, subjective ratings of criticality and more subjective ratings along with eye-tracking and ECG data (Heart Rate, RMSSD and pNN50) was recorded. The eye-tracking data was not analyzed for this paper and results from the ECG data will be reported qualitatively. The ECG data was checked visually for artefacts.

Since the first 5 cut-in scenarios in combination with the steering wheel moment were used to manipulate different system reliabilities, results will mainly cover data from the last critical scenarios. The subjective reliability rating applies to all scenarios.

The hypotheses can be divided into three major parts considering the dependent variables: objective parameters, subjective parameters and ECG data. The null hypothesis will always constitute no change from different levels of automation and different system reliabilities. Consequently, the alternative hypotheses will constitute an effect observable in the dependent variables by changing the level of automation and the system reliability. The alternative hypothesis are non-directional. All hypothesis address the last critical scenario. The alternative hypothesis are:

1. The minimum TTC is influenced by the level of automation and the system reliability.
2. The brake pedal reaction time is influenced by the level of automation and the system reliability.
3. The subjective criticality rating is influenced by the level of automation and the system reliability.
4. There is a correlation between the minimum TTC and the subjective criticality rating.

3 Results

A total of 65 participants (all employees of the BMW Group) took place in the study. Three were excluded from analysis due to an additional professional driver's license. Two were excluded due to technical issues during the experiment. One dataset in the reliable system condition was excluded because an accident happened due to the participant not reacting even though video recordings showed he was monitoring. The data of 45 men and 15 women, with a mean age of 30.58 years ($SD = 7.6$) was analyzed. More than half (58.3%) had no experience using adaptive cruise control, while only 18,3% had previous experience with a traffic jam assistance, resembling a L2 system. Participants were distributed equally throughout the study groups.

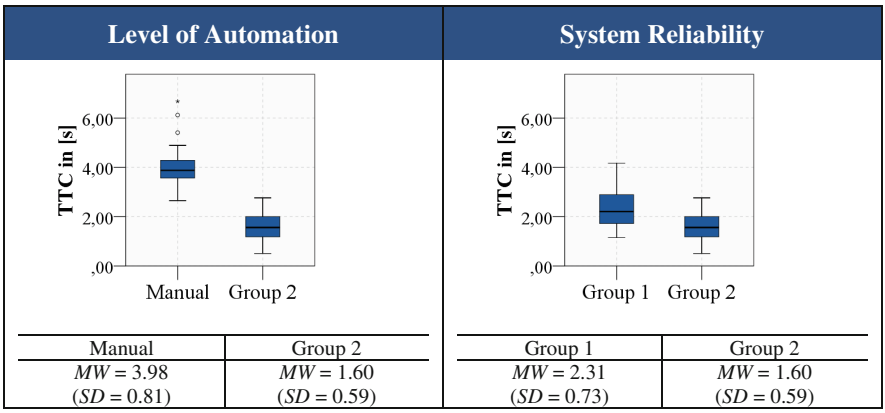
All data was tested considering normal distribution using the Kolmogorov-Smirnov-Test. In case a normal distribution could be assumed, a t-test for independent samples was used. In case no normal distribution could be assumed, the nonparametric Mann-Whitney-U-test was used. The p-value for all tests was set to .05. Effect sizes were calculated using Hedges-g with $|g| = .2$ for small effects, $|g| = .5$ for medium effects and $|g| = .8$ for large effects.

For analysis of the reliable system condition, participants in the reliable study groups were excluded that deactivated the system by braking in one of the previous five situations prior to the critical scenario, since it was hypothesized, that an objective system reliability of 100% could only be experienced by not taking over, leaving 29 participants. In the manual driving group, the sample size was 41, in the unreliable condition, sample size was 38.

The subjective rating of the system reliability confirmed the chosen procedure of manipulating the system reliability. While results did not represent the objective reliability of 50 and 100 % and it cannot be assumed that different system failures would lead

to equal effects on the perceived reliability, the manipulation allows for a valid interpretation of the objective parameters.

Figure 2 shows the analysis of the TTC for level of automation and system reliability. A Mann-Whitney-U-test showed a significant difference between the two levels of automation, $z = -7.06$, $p < .001$, $|g| = 3.27$. Data for system reliability was normally distributed, a t-test showed a significant difference, $t(65) = 4.28$, $p < .001$, $|g| = 1.06$. Thus, alternative hypothesis 1 was accepted.



Manual	Group 2
$MW = 3.98$ ($SD = 0.81$)	$MW = 1.60$ ($SD = 0.59$)

Group 1	Group 2
$MW = 2.31$ ($SD = 0.73$)	$MW = 1.60$ ($SD = 0.59$)

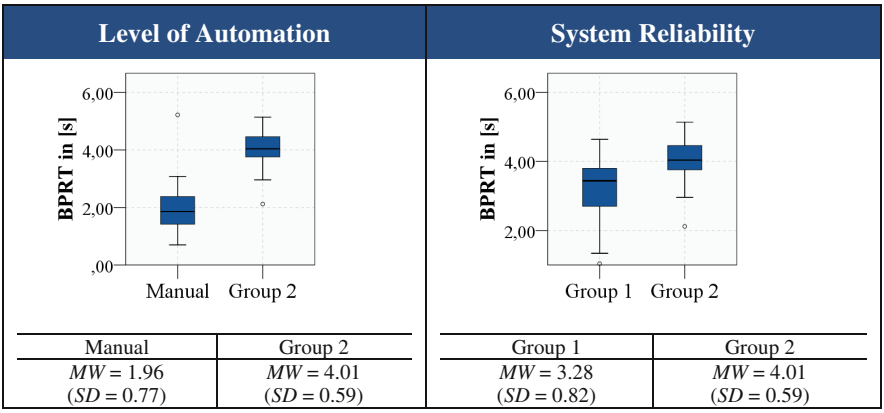
Fig. 2. TTC results. Both level of automation and system reliability show significant results. Group 1 represents the unreliable condition, Group 2 the reliable one.

In addition, the maximum deceleration as a measure for the criticality of a situation was analyzed. Results showed significant differences for level of automation and system reliability, with manual drivers showing lower decelerations, $z = -6.52$, $p < .001$, $|g| = 2.86$ and participants in the unreliable condition showing lower decelerations than in the reliable condition, $z = -3.59$, $p < .001$, $|g| = 1.04$.

Figure 3 shows the results of the brake pedal reaction time. There was a significant difference for level of automation, $t(68) = -11.73$, $p < .001$, $|g| = 2.92$.

Results for the system reliability also showed significant differences, $t(65) = -3.95$, $p < .001$, $|g| = 1.01$. Consequently, the alternative hypothesis 2 was also accepted.

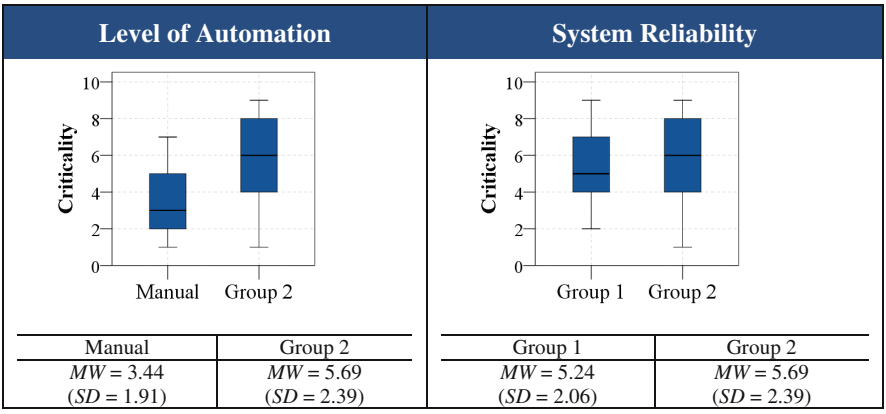
Results for the subjective criticality rating in Fig. 4 revealed strong differences concerning subjective rating scores between the L2 and the manual driving condition, as rating scores for the reliable condition depict a more critical evaluation compared to manual driving. Statistical significance between the two levels of automation was confirmed by a Mann-Whitney-U-Test, $z = -3.73$, $p < .001$, $|g| = 1.06$. Results concerning the system reliability showed no significant differences.



Manual	Group 2
MW = 1.96 (*SD* = 0.77)	*MW* = 4.01 (*SD* = 0.59)

Group 1	Group 2
MW = 3.28 (*SD* = 0.82)	*MW* = 4.01 (*SD* = 0.59)

Fig. 3. Brake pedal reaction time. Group 1 represents the unreliable condition, Group 2 the reliable one. Drivers experiencing the reliable condition reacted significantly later for both levels of automation and system reliability.



Manual	Group 2
MW = 3.44 (*SD* = 1.91)	*MW* = 5.69 (*SD* = 2.39)

Group 1	Group 2
MW = 5.24 (*SD* = 2.06)	*MW* = 5.69 (*SD* = 2.39)

Fig. 4. Subjective rating of the perceived criticality. A significant difference can be seen for level of automation but not for system reliability.

In order to evaluate a connection between TTC and subjective criticality rating, a nonparametric Spearman correlation was calculated. Results revealed a significant correlation between TTC and subjective criticality rating, $r = -.492$, $p < .001$.

ECG data was analyzed but showed no relevant or significant differences concerning the two factors automation and system reliability. Since the focus of this work was not put on differences in ECG recordings, no detailed information will be provided in this paper.

4 Discussion

Results concerning the level of automation showed significant differences in TTC and maximum deceleration. The last scenario was more critical for drivers who were driving partly automated (lower TTC's and higher decelerations). Also, drivers in the L2 condition reacted significantly later to the critical scenario than did manual drivers. In addition, the results from the subjective rating of criticality showed significant higher values for the partly automated driving condition. This result goes in line with models of information processing, as [9] point towards a reduction in vigilance due to decreasing resources of attention, as it would be the case in an automated driving condition. Consequently, objective as well as subjective results are in line with general principles concerning an increase in automation. Known issues and limitations of increasing the level of automation could be used to allow for an appropriate system design in the automotive sector.

The system reliability had a similar effect: TTC and maximum deceleration both showed significant differences in favor for the unreliable condition. Also, the brake pedal reaction time was significantly higher in the reliable condition compared to the unreliable one. This could be explained by a higher trust in the system leading to participants waiting longer for a reaction from the system in the reliable condition than people who experienced the unreliable system and already had to react to system limits or failures. Contrary, participants did not report a higher subjective criticality between the two different system reliabilities. Both groups equally considered the last scenario to be disturbing.

In case findings concerning system reliability and level of automation are put into a wider spectrum, participants showed an increase in manual driving skills if automation and system reliability decreased. While this seems perfectly in line with known effects, it rises questions concerning an even higher increase in automation (and most likely system reliability) on the way to highly automated driving (HAD): in case of a take-over in HAD, will the driver still have enough manual driving skills in order to carry the responsibility of being the fallback level in case a system limit is reached?

Future research work should especially address cut-in situations during partially and highly automated driving. Concerning safety issues, cut-in situations need to be analyzed in detail, as an appropriate response to vehicles cutting into the ego lane can be a great challenge for partly or highly automated vehicles including.

Acknowledgments. We would like to thank Patrick Galaske for his contribution to this work.

References

1. Committee, S.O.-R.A.V.S.: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Technical Report J3016_201401 (2014)
2. Winner, H., Schopper, M.: Adaptive cruise control. In: Winner, H., Hakuli, S., Lotz, F., Singer, C. (eds.) Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort, pp. 851–891. Springer Fachmedien Wiesbaden, Wiesbaden (2015)

3. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse, abuse. Hum. Factors **39**, 230–253 (1997)
4. Bainbridge, L.: Ironies of automation. Automatica **19**, 775–779 (1983)
5. Norman, D.A.: The 'problem' with automation: inappropriate feedback and interaction, not 'over-automation'. Philos. Trans. R. Soc. Lond. B Biol. Sci. **327**, 585–593 (1990)
6. Warm, J.S., Parasuraman, R., Matthews, G.: Vigilance requires hard mental work and is stressful. Hum. Factors **50**, 433–441 (2008)
7. Thiffault, P., Bergeron, J.: Monotony of road environment and driver fatigue: a simulator study. Accid. Anal. Prev. **35**, 381–391 (2003)
8. Endsley, M.R., Kiris, E.O.: The out-of-the-loop performance problem and level of control in automation. Hum. Factors **37**, 381–394 (1995)
9. Wickens, C.D., Carswell, C.M.: Information processing. In: Salvendy, G. (ed.) Handbook of Human Factors and Ergonomics, pp. 111–149. Wiley, Hoboken (2006)
10. Merat, N., Jamson, A.H.: How do drivers behave in a highly automated car. In: Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 514–521 (2009)
11. Zhao, C., Zhao, M., Liu, J., Zheng, C.: Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. Accid. Anal. Prev. **45**, 83–90 (2012)
12. Freyer, J.: Vernetzung von Fahrerassistenzsystemen zur Verbesserung des Spurwechselverhaltens von ACC. Cuvillier Verlag, Uitgever (2008)
13. Neukum, A., Lübbeke, T., Krüger, H., Mayser, C., Steinle, J.: ACC-Stop&Go: Fahrerverhalten an funktionalen Systemgrenzen. In: 5 Workshop Fahrerassistenzsysteme-FAS, pp. 141–150 (2008)

Looking at Drivers and Passengers to Inform Automated Driver State Monitoring of In and Out of the Loop

Christopher D.D. Cabrall[1(✉)], Veronika Petrovych[2], and Riender Happee[1]

[1] Intelligent Vehicles, Department of Cognitive Robotics, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands
{c.d.d.cabrall, r.happee}@tudelft.nl
[2] Driver and Vehicle, Department of Computer and Information Science (IDA), Swedish National Road and Transport Research Institute (VTI), Linkoping University, Olaus Magnus vag 35, Linkoping, Sweden
veronika.petrovych@vti.se

Abstract. The "drivenger" aim of the current study was to investigate attentional differentiation of drivers (who are in control) from passengers (who have no control) to inform future driver-in-the-loop monitoring/detection systems and facilitate multiple levels of manual/automated driving. Eye-tracking glasses were worn simultaneously by the driver and front seat passenger on 32 on road trips. Halfway en-route, the passenger was tasked with pretending with their eyes to be driving. Converging with a recent and independent drivenger study, our results found differences of higher probabilities of small saccades and significantly shorter blinks from our drivers and pseudo-drivers. Additionally, a new measure of eye eccentricity differentiated between driver/passenger roles. While naturalistic attentional manipulations may not be appropriately safe/available with actual automated vehicles, future studies might aim to further use the eye behavior of passengers to refine robust measures of driver (in)attention with increasing reductions in measurement intrusiveness and data filtering/processing overhead requirements.

Keywords: Human-systems integration · Driver state monitoring · Eye-tracking · Passenger · Driver · Drivenger · Automated driving

1 Introduction

The era of automated driving is upon us as recognized recently by both human factors experts [1] and public opinion [2]. Moreover, just as manual driving comprises of a multitude of activities in varying situations/scenarios, automated driving spans a wide range of functional allocation frameworks for the division of responsibilities between human and vehicle [3]. Far from entirely removing human factors that are often lamented as culprits for crash causation [4], the introduction/evolution of automation in driving vehicles continues to present ironies of automation [5] and rather than replacing human performance issues, it only changes them [6].

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_67

Human-systems integration benefits are expected from the monitoring of in-vehicle human occupants, especially the driver (or would-be driver), particularly in lower levels of automated driving that require his/her mental engagement and/or readiness to drive. But even in full automation, where pedals/wheel are removed entirely, the behavior of on-board occupants might serve as a communicative feedback input to the automated driving system to better adapt its performance. For example, if the automation is driving in an uncomfortable way human occupants, they may naturally exhibit their worry with greater attention to the driving scene whereas increased trust might be seen from reduced attention [7]. In lower levels of automated driving, humans are needed to periodically become available to drive as conditions may develop beyond operational boundaries of the automation. Under periods or manual, assistive, or supervisory driving control, responsible human beings must continuously monitor the driving activity but unfortunately humans are known to be susceptible to decrements in sustained attention tasks [8] particularly when not in control, lacking prediction/feedback, and forced to discriminate between subtle differences over prolonged periods of time between target signal events [9]. In all of the above cases, it would be useful for a driving system to know if and by how much a human is currently ready (and/or becoming ready) in terms of attention to the task of driving.

Many driving attention studies have been conducted in driving simulators for both good and bad reasons thus resulting in both high quality research and results that are open to criticism [10]. Especially regarding eye-tracking, use of real life visuals is certainly compelling regarding face-validity, and many studies exist employing previously recorded videos of driving rather than artificially generated graphics [11]. Moreover, with the advent of augmented reality, simulated items such as virtual hazards may soon combine on/through the windshields of real life vehicles for new innovative approaches to driver training and research [12] e.g., exploiting real world dynamics/aspects while retaining more benefits of experimental exposure, control, and repeatability than previously practically available.

Furthermore, catalyzed by exponential gains in computational power and reductions in size and cost of measurement equipment (especially cameras), on-road field operational and naturalistic studies have been growing in the last few decades [13] and benefit not only from real life visuals but all the pertinent interactive behavior that comes with being transported in a real vehicle with real life traffic complexities, social interactions and driving difficulties/dangers. Furthermore, alongside the rapid market release of living laboratory beta-testing of automated driving systems (i.e., Tesla Autopilot) and other commercially available features from traditional automobile manufacturers, many on-road studies are being conducted specifically regarding ADAS and automated system interaction and often include eye-tracking and attention components in the research.

Considering all the recent technology developments and advances in conducting driving research and especially the reduction of cost of measurement equipment, another source of data worth considering is attentional behavior from passengers to complement that collected from drivers. Under normal driving, presumably, a driver completes a traditional closed loop control-feedback system whereby consequences of exerted actions are compared for deviations from desired reference goals/targets which in turn influence/determine future driving control inputs. A passenger, by definition is

out of that control loop in terms of lacking physical input devices, yet retains (mostly) the full range of possible visual feedback afforded to drivers and hence may at times show similar and sometimes disparate eye behaviors. Presumably, then passenger eye behavior data may reflect that of a driver who is not actuating on driving input devices (i.e., various levels of automated driving). For example, an assumption of SAE level 2 driving automation systems [3], is that only the hands and feet activity of a driver should be relieved through automated lateral and longitudinal control, whereas the eye and mind responsibilities should remain with the driver. How in- or out-of-the-loop such a driver might be, might overlap with or be informed by measurable behavior from passengers in manually driven cars, without risking safety of experimenting with less than 100% attentive drivers of supervisory control.

A unique benefit to approaches utilizing passengers paired with drivers in the same vehicle, is that both are driving by definition subject to the same external factors, e.g., weather, traffic, local road infrastructure, time of day, etc. whereby they may serve as pseudo controls of one another in ipso facto manner. While not yet widely discussed or mainstream in driving safety research, such an innovative approach utilizing passengers is emerging. Through the power of suggestion and suspension of disbelief, Wizard-of-Oz procedures [14] essentially treated on-road passengers as drivers/supervisors of automated driving systems wherein the vehicle was actually driven by another human hidden behind a partition. Most relevant to the current study, [15] took simultaneous measurements from both the driver and passenger together via contact physiological instruments (i.e., EOG, EEG) to assess eye and attentional behavior of paired participants. The underlying motivation of [15] is an assumption shared by our present study, in which we believe the differentiation in the eye data from such participants may inform the development of detection systems monitoring the visual behavior of drivers with automated systems (i.e., where there are breaks/disconnects in driving physical control but not necessarily so in visual/mental attention).

2 Methods

2.1 Participants

Participants (16 pairs, 78% male, 22% female, mean age = 27.3, SD age = 2.4) were recruited from Delft University of Technology, and the study was run after obtaining written informed consent under the approval of the Human Research Ethics Committee "On Road In Vehicle Eye Tracking: Drivers and Passengers" (26-09-2016). Pairs were formed around a quasi-experimental variable of familiarity, such that half of the pairs knew one another well, whereas with the other half, participants were not known to one another in advance. However, analyses pertaining to this aspect have been reserved for adjunct publications. Each participant had normal or corrected-to-normal vision and reported having capability/comfort in driving from having obtained their initial driver's license for at least more than one year prior to the experiment.

2.2 Driving Route

The driving route began from the rear parking lot of the 3 mE Faculty building of Delft University of Technology (Leeghwaterstraat) and proceeded across campus and southbound on Schoemakerstraat, westward along Kruithuisweg to join the A4 highway northbound until exit 12 for route N 211, at which point the route crossed over the highway to return back along the same roads in reverse direction (Fig. 1). The full route was completed as one trip of about 20.0 km and around 30 min on average, and repeated per pair with a switching of driver/passenger role, for a total of 32 trips.

Fig. 1. The driving route in the experiment covering mixed urban and highway roads. Representative screen shots are provided for various route segments at labeled points on the map.

2.3 Procedures

Drivers were always given no other instructions than to drive as they normally would in a safe manner. Passengers were initially given no instruction other than to allow for them to do whatever they wanted and normally might do as a passenger, but from the

turnaround point (along the on-ramp to re-join the A4 highway southbound), the passengers were given a piece of paper to covertly assign an additional experimental instruction to "*Please **imagine** that you are doing the driving. So try to pay attention and **behave with your eyes** as if you are currently driving. You do not need to move your hand/feet like a driver.*" In lieu of naturalistic motivations, no restrictions were expressed/placed on conversation, use of electronic devices, etc.

2.4 Apparatus

Both passenger and driver participants wore UV shielded eye-tracking glasses from SensoMotoric Instruments (SMI) coupled by a single USB cable each to their own dedicated Samsung Galaxy smartphone running only the eye-tracking software and were held by ride-along experimenters in the backseats (Fig. 2). The car driven was a 2014 Toyota Prius Hybrid passenger vehicle with automatic transmission without use of any cruise control. Subsequent publications are planned to detail additional equipment and data collected during this experiment from concurrent on-board telemetry of the research vehicle including forward radar, GPS, steering, and pedal inputs, etc.

Fig. 2. Passenger and driver equipped with minimally invasive eye-tracking glasses.

2.5 Eye-Tracking Data Measurement/Analysis

The SMI glasses recorded eye measurement samples at 60 Hz. The gaze eye data was indexed along a 960 wide × 780 tall pixel coordinate location system (upper left corner origin) in respect of the image plane of the integrated forward facing camera. It should be noted that the coordinate frame moved as the participant moved his/her head and it remains for us later outside the scope of the present study to further (re)analyze our data in a rectified/resolved 3D world model as needed. Eye data was classified via the automatic categories provided from SMI BeGaze with the default settings as follows. The "low speed event detection" algorithm derived saccades between fixations that were identified from a minimum duration setting of 80 ms and a maximum dispersion area of 100 px. Blink events registered as special fixation cases, whenever the pupil diameter was less than 1 pixel and/or the horizontal and vertical gaze position equaled 0, while discarding durations shorter than 70 ms.

Eccentricity scores were computed as a product of the duration and distance of raw gaze samples away from an individually calibrated central point in head coordinates. First, a diagonal Euclidean distance was computed from a gaze sample location to the central X/Y coordinate point (480,360), and then divided by 600 (and multiplied by 100) to result in a percentile of distance from the center to the corner of the coordinate grid. Next, such relative distances were rounded into bins of size 10 and over an entire period of analysis interest, a modal percentage bin was selected as representative of a new calibrated average center gaze location (i.e., allowing for this to be somewhere off the coordinate grid center point). Finally, for a given analysis period of interest, for every new sample that is outside of that central modal bin (and the previous was not), the duration was recorded until the next gaze sample returns to that central modal area.

In accordance with the analysis procedures used in the previous driver and eye-tracking differentiation study [15], all replicated eye measurement dependent variable data was individually normalized according to the equation: $V_{norm} = (V - V_{pLow})/(V_{pHigh} - V_{pLow})$ where V indicates the non-normalized variable, and V_{pLow} (V_{pHigh}) indicates the bottom (top) 20-percentile of the variable data population of each participant. Subsequently, bandwidth cutoff thresholds of 0 and 1 were applied to the resultant V_{norm} values. In the current study, eccentricity scores were not yet normalized, however, to explore the potentially meaningfulness of off-nominal behavior captured in the data as a more robust differentiating measure of in/out-of-the-loop.

The current study did not yet sectionalize different portions of the route into specific urban and highway scenario portions, but instead examined the three following time/event window periods of interest. Additionally, for the first two periods of interest with shorter time windows, eye tracking data was analyzed with half of the data (i.e. those from the first trip the paired participants drove) in order to help mitigate potential learning/bias effects of participant attentional/visual behavior.

(1a)/(1b) **"Post/Pre Task"** = 120 s

The portion of time for the few minutes immediately before compared against the time period immediately after the passenger task instruction presentation (±120 s). Here the data was taken from passengers only and the "post-task" data (1b) are regarded as "pseudo drivers" attempting to represent visual control whereas the "pre-task" data (1a) are regarded still as natural (untasked) freely varying passenger eye data.

(2) **"Entering A4"** ∼around 45 s;

The first highway on-ramp and merging period where it was assumed a driver would be likely to prioritize and evidence high levels of dedicated driving control visual behavior (about 45 s). Both driver and passenger eye data are included.

(3) **"Gate to Task"** ∼around 900 s;

From the start of the trip (leaving the parking lot gate) up until the start of the passenger task manipulation (about 900 s). Both driver and passenger eye data are included.

3 Results

The result sections are first organized by dependent eye measure: saccade amplitude, blinks, and eccentricity, and within each divided by sectionalized periods of interest (pre/post passenger task instructions, entering A4 highway, and gate-to-task).

3.1 Saccade Amplitude

Figure 3 suggests that across the three analysis periods of interest, smaller saccades appeared more frequently in drivers (or pseudo drivers) than in passengers, whereas the reverse was true of larger saccades. This relational difference was expressed most apparent within the period of entering the A4 highway and least when taken across the full period leading up until the passenger instructional task. In interest of replicating the results from similar previous work [15], the same proportional divisions for "small" and "large" saccades were taken and compared, see dashed vertical boundary lines in Fig. 3 spanning from 0.05 to 0.2 and from 0.2 to 0.8 of the normalized range respectively.

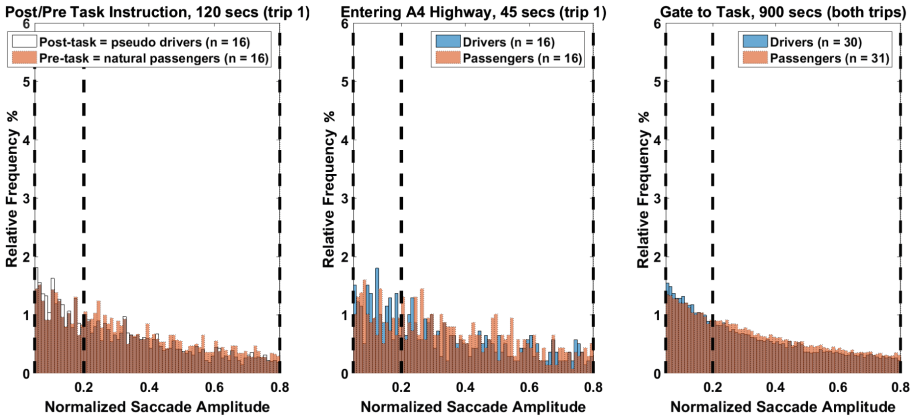


Fig. 3. Relative frequency of saccade amplitude in each analysis period of interest.

A two-way analysis of variance ANOVA for role ("driver", passenger) and analysis period (around passenger task change, enter A4 highway, gate to task) was conducted separately for the number of normalized small and large saccades (presented per second and re-scaled to absolute units in Fig. 4). For small saccades, the hypothesized directional difference was found in higher rates of small saccades per second for driver roles (m = 0.418, SD = 0.38) than passengers (m = 0.366, SD = 0.37), F (1,117) but failed to obtain significance in the present analysis, (p = 0.50). A large increase in rate of small saccades was observed in the period surrounding the passenger task change relative to the other two periods of analysis, but also failed to obtain significance, (p = 0.16). Lastly, the interaction effect also did not reach significance levels

(p = 0.90). Similarly, for rates of large saccades, the role appeared in the expected direction with higher probabilities of larger saccades for passengers (m = 0.754) than for drivers (m = 0.673) but did not reach significance (p = 0.49), nor for analysis period window (p = 0.26), nor for significant interaction effect (p = 0.79).

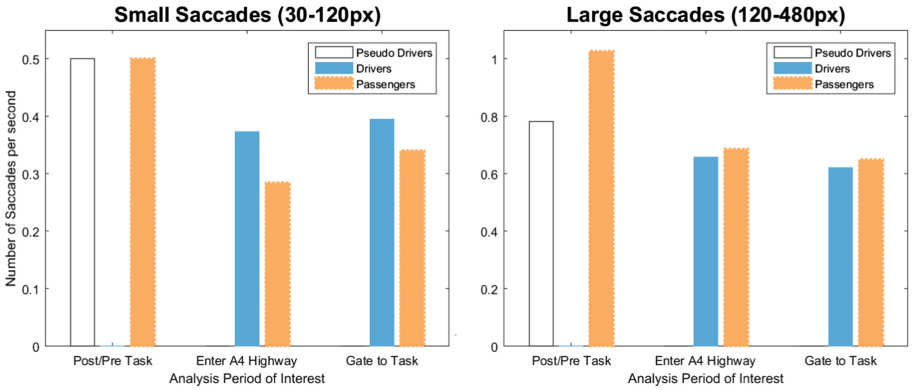


Fig. 4. Number of small (left panel) and large (right panel) saccadic eye-movements.

3.2 Eye Blinks

A two–way analysis of variance ANOVA for role ("driver", passenger) and analysis period (around passenger task change, enter A4 highway, gate to task) was conducted for eye blink duration (Fig. 5). Significant results showed for both main effects with a significant interaction. Overall, drivers evidenced shorter blinks than passengers, $F(1,15922) = 12.832$, $p < 0.001$, $\eta_p^2 .= 0.001$ and significant differences were found between the analysis period windows, $F(2,15922) = 6.15$, $p < 0.01$, $\eta_p^2 .= 0.001$.

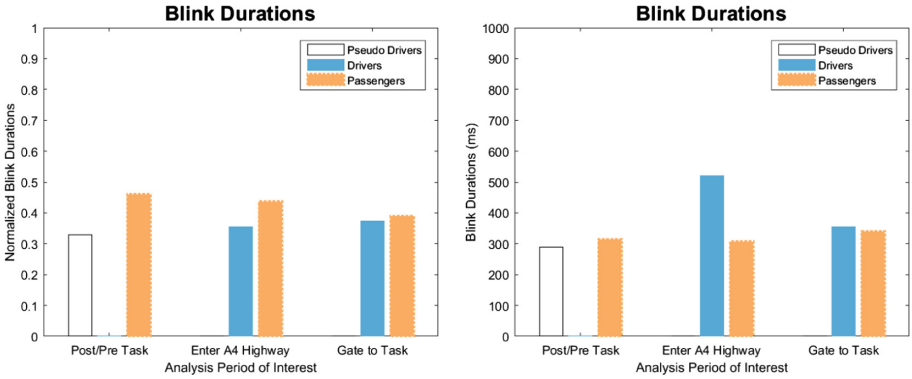


Fig. 5. Mean duration of blinks, normalized (left panel) and in milliseconds (right panel).

Tukey's post hoc analysis of period window showed average blink durations significantly lower in the full 900 s period from gate to task compared to the ±120 s window periods around the passenger task instruction ($p < 0.01$). The interaction effect showed an increasing differentiation between drivers and passengers in higher specified/controlled contexts ($p < 0.01$), $\eta_p^2 .= 0.001$. Differences did not reach levels of significance regarding a main effect of analysis window period ($p = 0.36$).

3.3 Eccentricity

A two–way analysis of variance ANOVA for role ("driver", passenger) and analysis period (around passenger task change, enter A4 highway, gate to task) was conducted for the raw non-normalized scores of eccentricity (i.e., a product of off-center distance multiplied by off-center duration). Significant results were obtained for the main effect of role (Fig. 6) with passengers evidencing higher average eccentricity scores ($m = 20.766$, $SD = 8.97$) over drivers ($m = 15.508$, $SD = 8.81$), $F(1,117)$, $p < 0.005$, $\eta_p^2 .= 0.079$. Differences did not reach levels of significance regarding a main effect of analysis window period ($p = 0.36$), and the directional main effect of the driver vs. passenger role evidenced parallel differentiation without interaction effect (Fig. 6).

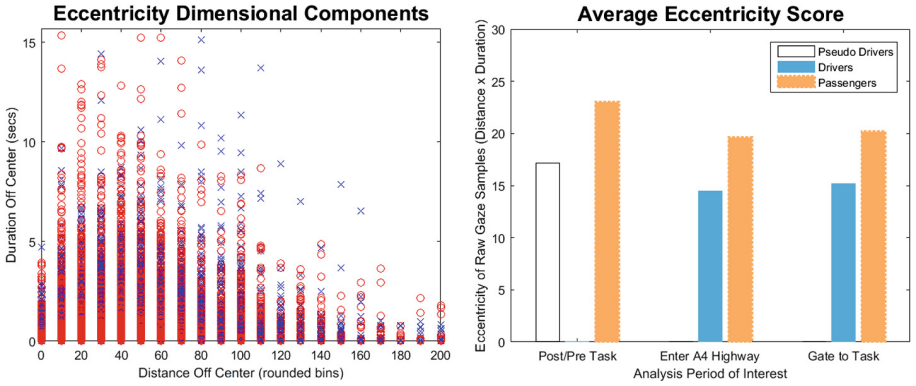


Fig. 6. Eccentricity dimensional distribution (left) and average scores (i.e., the functional product of off-modal-center gaze sample distances multiplied by duration of time until returning to that center) across driver vs. passenger role and analysis period of interest (right).

4 Discussion/Conclusion

In regards to the recent and innovative driver/passenger study [15], the present study complements by replication and extension of both procedural methods and results. In both cases, simultaneous eye-tracking information was extracted from both drivers and front passengers in the same vehicle on the same trip, and with a reversal of role assignment, the trip was repeated. Relative to active in-control drivers, the major eye-tracking results of [15] found passengers to exhibit lower amounts of smaller saccades and higher amounts of larger saccades, found lower visual processing loads of large saccades, and longer blink durations indicative of reduced arousal and collectively reflective of a decrease of attention in passengers. Comparatively, the present study also sought and found such a relational reduction in attention to driving control from drivers to passengers but employed less invasive and demanding eye measurement devices (i.e., eye-tracking glasses similar in form factor to sun-glasses attached by a single USB cable to a smart-phone), as well as direct instruction to task the passenger and his/her attention at a specific point in the drive. Additionally, the added location information from camera based eye-tracking glasses relative to the EOG electrodes of [15], allowed for direct distance measurements. Such eye distance movement data was used in combination with time to derive a measure of eccentricity (i.e., product of off-center samples multiplied by duration of time off-center) to more robustly/meaningfully differentiate between the attentional behavior of in-control drivers and passengers who have no direct vehicular control.

For driver state monitoring aims, it is not the direct reliable classification of passenger eye behavior per se that is most valuable, but what is particularly productive from our results are the main trends (e.g., >85%) evidenced in self-to-self comparisons (Fig. 7) of an increasing eccentricity from an in-the-loop role driver to an either in-/out-of-the-loop more freely varying attention of the passenger. For example, individualized momentary proportionate increases in eccentricity when comparing the same person when tasked as passenger (i.e., to be acting with their eyes like a driver) back to him/her-self as an untasked passenger role (m = +78%, SD = 0.95) and when comparing a longer extended role change of the same individual from serving as a veridical driver vs. a natural passenger (m = +53%, SD = 0.39) may serve as informative to potential threshold ranges to trigger upon within a driver state monitoring system. For example, such thresholds might be built into a system to evaluate if the driving attention of a human supervisor of automated driving is lapsing into a more passive passenger role or in another scenario serve as a check if the human has reached the required level of attentional performance when returning to driving after a period of being away from it (i.e., automated-to-manual transition of control).

Furthermore, such results were obtained from raw gaze samples in spite of inherent noise and lack of head tracking and/or sophisticated techniques to resolve/orient together within a 3D world model. Within our present simplistic eccentricity analysis, saccadic amplitude measures were most likely impacted by any co-occurrence of head movement with a saccade (e.g., in the same direction) that would effectively mathematically reduce, cancel out, or otherwise confound the saccade amplitude measure compared to that same size saccade made without or with less head movement.

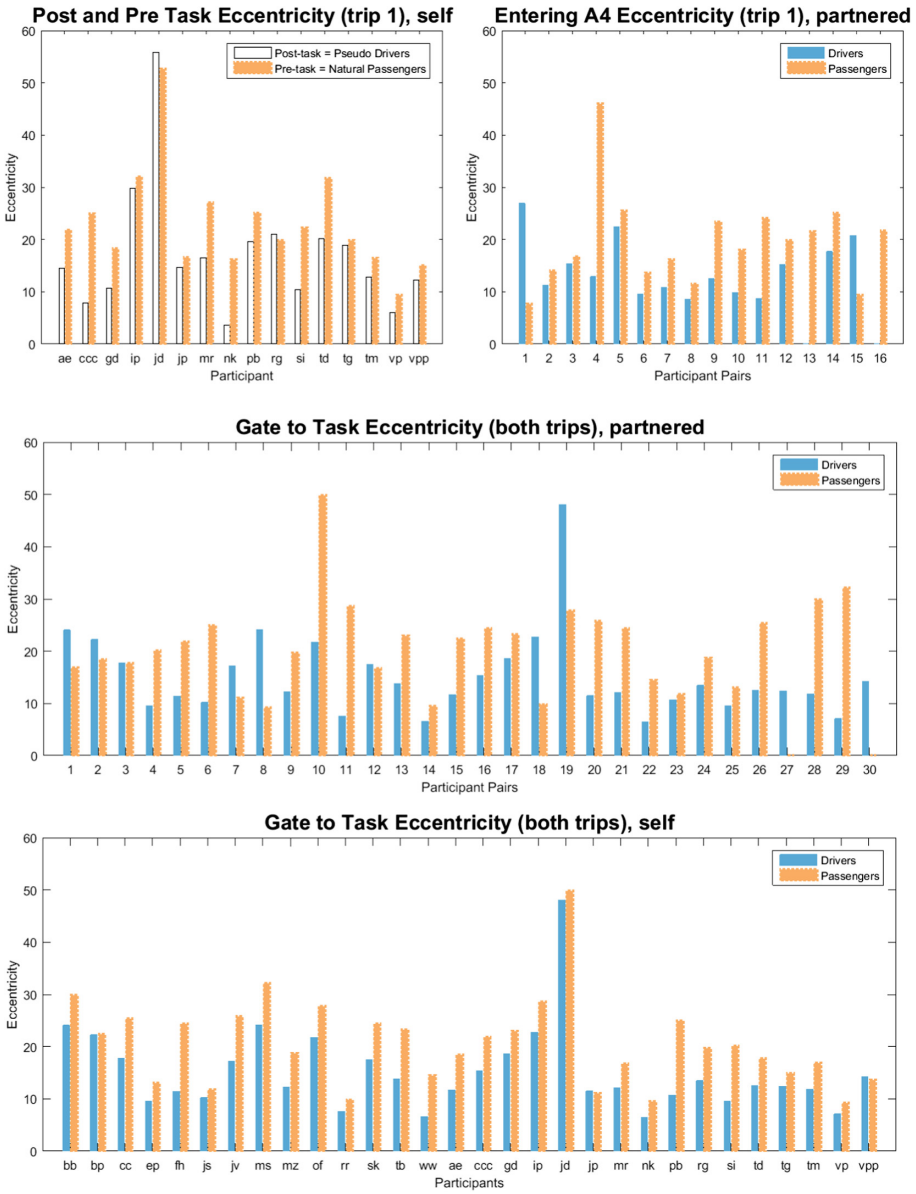


Fig. 7. Eccentricity across analysis periods of interest and across participants both in self and partnered comparisons. Note: passenger eccentricity exceeded drivers in 14 of 16 self- cases (88%, upper left); in 14 of 16 cases (88%, upper right); in 20 of 28 cases (71%, middle); and in 28 of 30 cases (93%, bottom).

While this is not of practical concern for blinks (generally) or for saccade amplitude (in [15]) because measurement is at an origin regardless of head orientation, it is interesting that our results show that a measure of modal eccentricity is still able to differentiate in control drivers from lack of control passengers in spite of such head movement confounds. In general, such practical methodological measures and benefits are in line with approaches of both the raw gaze-based "unfiltered" percentage road center (PRC) and central gaze-based PRC approaches described and validated within [16].

Our results regarding eccentricity should be taken with several caveats. In our present preliminary analyses, the "pre-task" passenger eye data included reading the instructions. Furthermore, if a driver more fully concentrated on a secondary task to the point of becoming a primary task, it might be expected that eccentricity would decrease rather than increase. Without world knowledge, the eccentricity measure is agnostic as to what specifically is being concentrated on, but instead reflects more only the presence/absence of concentration/control. Lastly, the differentiating impact of blink duration changed across our analysis period segments whereas eccentricity did not. Future studies should examine the attententional impact of contextual aspects of the driving scene for example such as velocity, road curvature, and other traffic.

Acknowledgments. The authors are supported by and performed the current work under the Marie Curie ITN: HF Auto, Human Factors of Automated Driving (PITN-GA-2013605817), www.hf-auto.eu. We would like to thank Dr. Magnus Hjalmdahl for inspiration towards simultaneous eye tracking of driver vs. passenger while driving across Sweden between VTI office while playing with one of the eye tracking glasses used in the present study. Additionally, we are also indebted to Peter van Leeuwen for his insights into the potential value of driver visual eccentricity behavior, specifically regarding both time and distances away from center taken jointly together.

References

1. Kyriakidis, M., de Winter, J.C.F., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., et al.: A human factors perspective on automated driving. Theor. Issues Ergon. Sci. (2017)
2. Kyriakidis, M., Happee, R., de Winter, J.C.F.: Public opinion on automated driving: results of an international questionnaire among 5000 respondents. Transp. Res. F **32**, 127–140 (2015)
3. SAE: taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. On-road automated vehicle standards committee, SAE International (2014). http://standards.sae.org/j3016_201401
4. Smith, B.W.: Human error as cause of vehicle crashes. Center for Internet and Society at Stanford Law School (2013). http://cyberlaw.stanford.edu/blog/2013/12/human-error-cause-vehicle-crashes
5. Baxter, G., Rooksby, J., Wang, Y., Khajeh-Hosseini, A.: The ironies of automation … still going strong at 30? In: Proceedings of the ECCE Conference, Edinburgh (2012)
6. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. Trans. Syst. Man Cybern. Syst. Hum. **30**, 286–297 (2012)

7. Hergeth, S., Lorenz, L., Vilmek, R., Krems, J.: Keep your scanners peeled: gaze behavior as a measure of automation trust during highly automated driving. Hum. Factors **58**(3), 509–519 (2016)
8. Mackworth, N.: The breakdown of vigilance during prolonged visual search. Q. J. Exp. Psychol. **1**(1), 6–21 (1948)
9. Cabrall, C.D.D., Happee, R., de Winter, J.C.F.: From Mackworth's clock to the open road: a literature review on driver vigilance task operationalization. Transp. Res. F **40**, 169–189 (2016)
10. Caird, J.K., Horrey, W.: Twelve practical and useful questions about driving simulation. In: Fisher, D.L., et al. (eds.) Handbook of Driving Simulation for Engineering, Medicine, and Psychology, pp. 5.1–5.16 (2011). Chap. 5
11. Chapman, P., Underwood, G.: Visual search of dynamic scenes: event types and the role of experience in viewing driving situations. In: Underwood, G. (ed.) Eye Guidances in Reading and Scene Perception, pp. 369–393. Elsevier, Oxford (1998)
12. Sheridan, T.B.: Recollections on *Presence* beginnings, some challenges for augmented and virtual reality. Presence **25**(1), 75–77 (2016)
13. Regan, M.A, Williamson, A., Grzebieta, R., Tao, L.: Naturalistic driving studies: literature review and planning for the australian naturalistic driving study. Australian College of Road Safety (ACRS) (2012). http://acrs.org.au/files/papers/57_Regan-PR.pdf
14. Baltodano, S., Sibi, S., Martelaro, N., Gowda, N., Ju, W.: The RRADS platform: a real road autonomous driving simulator. In: Proceedings of the 7th International Conference on Automotive User Interfaces and Vehicular Applications, pp. 281–288 (2015)
15. Takeda, Y., Sato, T., Kimura, K., Komine, H., Akamatsu, M., Sato, J.: Electrophysiological evaluation of attention in drivers and passengers: toward an understating of drivers' attentional state in autonomous vehicles. Transp. Res. F **42**, 140–150 (2016)
16. Wang, Y., Reimer, B., Dobres, J., Mehler, B.: The sensitivity of different methodologies for characterizing driver's gaze concentration under increase cognitive demand. Transp. Res. F **26**, 227–237 (2014)

Taxonomy of Traffic Situations for the Interaction between Automated Vehicles and Human Road Users

Tanja Fuest[1(✉)], Lenja Sorokin[2(✉)], Hanna Bellem[2], and Klaus Bengler[1]

[1] Department of Mechanical Engineering, Chair of Ergonomics, Technical University of Munich, Boltzmannstraße 15, 85747 Garching b. München, Germany
{Tanja.Fuest,Bengler}@TUM.de
[2] BMW Group Research, Innovations, New Technologies, Parkring 19, 85748 Garching b. München, Germany
{Lenja.Sorokin,Hanna.Bellem}@bmw.de

Abstract. The aim of this research is to develop and present a taxonomy of traffic situations relevant for the assessment of communication processes between automated vehicles and other road users. The taxonomy is intended to serve as a basis for concept development and further research on the effect of automated vehicles on human road users like pedestrians, cyclists and manual drivers. The taxonomy can be used to choose attributes and value facets which are relevant for a specific research question. This facilitates a systematic definition of traffic situations, which ensures that all relevant aspects are addressed when eliciting requirements or validating concepts. The overview of attributes also facilitates the identification of potential confounders. Another benefit is the enhanced comparability between different studies. The proposed taxonomy differs from previous taxonomies of traffic situations by focusing not only on the driver's but on the interaction partner's perspective, too.

Keywords: Taxonomy of traffic situations · Automated vehicle · Interaction with traffic participants · Explicit communication · Implicit communication

1 Introduction

An increasing number of automobile manufacturers is planning to integrate automated test cars into manual traffic within the next years [1]. Although it is not yet foreseeable when highly or fully automated vehicles will be sold, measures must be taken at an early stage to ensure a smooth mixed traffic. As interaction between traffic participants is one underlying necessary condition of safe and efficient traffic, it is important not just to solve technical issues but also to ensure the social acceptance of automated vehicles (AV) [2]. This implies that AVs should be able to interact and communicate in order to drive more safely and comprehensibly for other road users. This task is particularly challenging for mixed traffic consisting of self-driving and human-driven cars.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_68

It is therefore important to research the main challenges regarding communication and interaction with AVs:

1. An AV does not necessarily have a passenger. Therefore, other road users do not have a (human) communication partner, meaning that eye contact, mimic and gestures are not possible anymore. It remains to be investigated in which way this sort of communication needs to be replaced.
2. As driving behavior, itself is a mean of communication, it seems necessary for AVs to show an unambiguous driving behavior. This way a clear message, for example "the AV will stop and let you cross the road", could be delivered and ease traffic flow. If the influence of driving behavior is not regarded, misleading communication situations can occur ("you cannot not communicate" [3]).
3. It remains unclear how the labeling of self-driving cars can affect other road users' behavior.

Communication in traffic is intentional, it may differ from situation to situation and always depends on the circumstances. Therefore, every traffic situation needs to be assessed as a situation in which social interaction takes place [4]. Interaction refers to reciprocal actions of two or more entities [2]. For example, drivers make use of informal communication in car parks where only few legal regulations exist [5] whilst informal communication is less relevant for high-speed traffic situations. When developing future communication systems, it is therefore necessary to define relevant situation characteristics [6].

Therefore, it is crucial to have a systematic overview over the different attributes that influence the interaction between involved road participants before conducting a study. This can be done by means of a taxonomy. It defines all relevant aspects of social communication and interaction between AVs and human road users (HRU), which can occur in traffic. This allows for testing potential communication designs of AVs and to control disturbances. Thus, the taxonomy contributes to the comparability of different studies and helps to structure and conduct further studies as well as to design and verify the interaction between AVs and HRUs. Relevant characteristics for different research questions can be filtered and varied in the simulator or on the test track.

2 Communication

The term communication is based on the Latin word "communicatio" and means message or statement [7]. Communication between different road users in traffic is mainly used to express intentions of initiating interactions.

Most traffic situations, such as the simultaneous approach of a car and a pedestrian to a crossover, are regulated by the German traffic regulations (StVo, §26). Apart from this legal framework, there are many informal rules that intend to ensure or optimize traffic flow. Here, nonverbal communication plays an important role. Usually, a short eye contact or a gesture is enough to make the respective intentions clear. For an AV, these means of communication are not available and might be compensated [8].

Various forms of transmitting information exist [7]. Communication can be directed to either one individual or to a bigger group of individuals. For example, the driver of a car can convey to a pedestrian that he noticed him through a short eye contact or he can indicate in order to pass his intention to turn to the following car and the oncoming traffic. All these messages can be transmitted implicitly or explicitly [9].

2.1 Explicit Communication

If a message is explicit, the transmitter conveys his intention in direct signs to the receiver. For example, the transmitter can tell the receiver, that he is from Germany. [9]

Explicit communication can also be found in traffic. The driver can transmit his intentions by using light and sound signals. For example, if he is indicating left, other road users know that the driver is going to turn left or is overtaking another car. Light signals in traffic are, among others, the turn indicator, hazard warning lights, warning lights or flashing blue police lights. Sound signals include the horn, the siren, or, when referring to electric cars, the sound transmission of a beep to the outside to warn pedestrians. Additionally, messages may be transmitted through projections or monitors on the street in the near future.

Labeling a car as an AV is also a form of explicit communication.

2.2 Implicit Communication

A message is transmitted implicitly, if the content is not spoken out loud, but indirectly included in the message. For instance, a message that someone is coming from Germany can also be recognized in the form of an accent [9].

Usually, nonverbal methods of communication are used to transmit implicit messages [10]. The pronunciation and stress of words as well as gesture and mimic can always add additional information to the content of the message [9].

However, in traffic it is common to use nonverbal methods of communication not for implicit, but for explicit messages [11]. If a driver is waving to a child to indicate that it can cross the street, it means that the driver already checked that the street is free on both sides and that the child can cross the street safely. Therefore, implicit communication on the street means the transmission of intentions via the usage of specifically designed trajectories. For example, a car driver can also encourage a pedestrian to cross the street by reducing speed. Lehsing et al. [12] have shown that it is possible to quantify the interaction based upon this communication.

Both implicit and explicit forms of communication occur in traffic. As a driver does not communicate in an AV, the systems communication must be analyzed and designed. Because communication is dependent on the situation, it is important to have a structured approach when designing use cases or studies to keep the number of potential cases for development and assessment manageable but still ensure validity.

3 Approaches to Cluster Traffic Situations

Several approaches to cluster traffic situations from different perspectives exist. For example, error classifications and accident types have a strong situational dependency, meaning that certain types of errors can only be caused in specific situations [13]. The causes for accidents are numerous (for example wrong estimation of the distance to the vehicle ahead, inattention, impairment from alcohol or drugs, etc.) [14–17].

Within the UR:BAN project, scenarios for assistance systems were defined based on their characteristics (reliability, field of application, e.g.), the driving task level (navigation, maneuvering or stabilization) and the driver (time requirements and operating load). As a result, they were able to coordinate a vast number of different tests [18].

There are, however, other taxonomies of traffic situations that are not based on accident types. Benda [19] has developed a taxonomy which focuses on the driver's perspective. The taxonomy combines 14 categories with up to 10 specifications each and is therefore very complex and difficult to use. Fastenmeier [20] adapted the categories to make the taxonomy more manageable by eliminating redundant categories and some specifications.

Bengler et al. [2] have developed a taxonomy of the most important aspects of human-machine cooperation. They differentiate between five layers: intention, modes of cooperation, allocation, interfaces and contact.

Other authors define use cases for automated driving by distinguishing, for example, the maximum speed limit, dynamic elements (AVs or manual drivers) and the scenario (off-road, dirt road, parking area, highway, etc.) [21].

So far, no taxonomy defining situations, in which an interaction between two road users takes place, yet exists.

4 Taxonomy of Traffic Situations for the Interaction between Automated Vehicles and Human Road Users

A taxonomy of traffic situations, which classifies the interaction between an AV and a HRU according to relevant influence factors, was developed. It differs from previous approaches to cluster traffic situations, which focus on the driver's perspective or define specific use cases [19–21]. The proposed taxonomy addresses interactions between the AV and other traffic participants such as pedestrians, cyclists and other car drivers. It addresses implicit as well as explicit communication and provides an overview over the most relevant attributes that influence these forms of communication. Attributes are situation specific characteristics, such as driving speed, the AV's intentions and the right of way. All attributes are divided into different value facets. For the selection of the attributes it was relevant whether the attributes could affect the comprehension of the interaction, the HRU's acceptance of the AV, the target achievement of the AV and the HRU, and the objective und subjective safety.

Table 1. Overview of all attributes and value facets.

Attribute	Value facets
Right of way	AV
	HRU
	Undefined
AV's intention regarding right of way	Let the HRU go first
	Go first
HRU's intention regarding right of way	Let the AV go first
	Go first
HRU Character	Vehicle drivers
	Cyclists
	Pedestrians
Longitudinal distance (Headway)	<3 m
	3–10 m
	>10 m
Lateral distance	0 m
	≤3 m
	>3 m
Attention HRU	Yes
	No
Impairment of the HRU's perception	View
	Acoustic
	Both (view and acoustic)
	No impairment
Speed AV	0 km/h
	30 km/h
	50 km/h
	130 km/h
Speed HRU	0 km/h
	4.4 km/h
	17.5 km/h
	30 km/h
	50 km/h
	130 km/h
Driving direction AV	Driving forward
	Reverse
Perspective (from the perspective of the AV)	Ahead
	Sideways / diagonal
	Backward

The taxonomy can be used by choosing attributes and value facets which are relevant for a specific research question. This facilitates a systematic definition of traffic situations. This, in turn ensures that all relevant aspects are addressed when eliciting

requirements for the communication design of AVs or validating concepts. The overview of attributes also facilitates the identification of potential confounders. Another benefit is the enhanced comparability between different studies. An explanation of the different attributes will be given below (see Table 1).

Right of Way. The taxonomies of Benda [19] and Fastenmeier [20] already distinguish different forms of right of way. The attribute has been added to the taxonomy because the right of way in combination with the AV´s intention regarding the right of way defines the tone of the AV (ask for the right of way vs. enforce the right of way) and, thus, the HRU's acceptance of the AV could be influenced. Possible value facets could be:

- *AV*: The AV has the right of way.
- *HRU*: The HRU has the right of way.
- *Undefined*: The right of way is undefined and the road users have to communicate with each other in order to resolve the situation.

AV's Intention regarding Right of Way. The intention of AVs with regard to the right of way determines the message which should be communicated by the AV (go first vs. go on) and is thus relevant for the driving behavior. This will affect the AV's capabilities of achieving its objective. In addition, the intention in combination with the right of way could influence the HRU's understanding of the interaction. There are traffic situations where a discrepancy between the right of way and the AV´s intention regarding the right of way is important. For example, if a red traffic light makes it impossible for the AV to drive a few meters forward, it should allow a left turn of another car on the opposite lane, to drive into a private driveway (see Fig. 1). The AV leaves the HRU the right of way. This discrepancy could have an impact on the understanding of the interaction, on the acceptance of the AV by the HRU, and in more critical situations also on safety. The following value facets are distinguished:

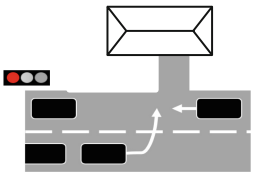


Fig. 1. Example of a specific situation

- *Let the HRU go first*: AV yields the right of way to the HRU.
- *Go first*: AV claims the right of way.

HRU's Intention regarding Right of Way. Not only the AV's intention regarding the right of way, but also the HRU's intention can change the interaction and therefore the goal achievement of both AV and HRU. In combination with the AV's behavior the HRU's acceptance of the AV could change. Possible value facets could be:

- *Let the AV go first*: HRU yields the right of way to the AV.
- *Go first*: HRU claims the right of way.

HRU Character. It cannot be assumed that the communication methods and strategies of the AV will be equal amongst different interaction partners. For example, the scope of information may have to be adapted according to the interaction partner. Also, the need for safety will differ depending on the HRU type. A pedestrian may have a higher need for safety and information than a vehicle driver. Therefore, the AV's influence on objective and subjective safety plays an important role.

The value facets are limited to interacting partners who must be assumed to have no vehicle-to-vehicle (V2V) communication facilities and will not be equipped with the latter in the next years:

- *Vehicle drivers*: Communication is restricted by the bodywork.
- *Cyclists:* Need for safety may be higher than for cars, because cyclists are not protected by a bodywork.
- *Pedestrians*: They may have a high need for safety, because they are not protected by a bodywork. Additionally, their next action is often not predictable and the possible speed of pedestrians is very low.

Speed AV. Other studies have already used the top speed to characterize situations [15, 21]. In addition, excessive speed is a common cause of traffic accidents [22] and has thus an impact on the objective safety.

Especially at high speeds, communication has to be executed very fast, but it can be communicated in more detail at lower speeds. Thus, the communication strategy of the AV must be adapted to the driving speed. At a maximum speed of 130 km/h other ways of communication by using the trajectory have to be used than at a lower speed. When designing the communication of an AV three time states need to be defined (before, during and after the interaction), because the speed adapts dynamically. The speed of the AV was split into the following value facets:

- 0 km/h: The AV is standing still.
- 30 km/h: Low speed range.
- 50 km/h: Urban speed range.
- 130 km/h: Maximum permissible speed in most countries throughout Europe.

Speed HRU. The inclusion of the attribute "speed HRU" is based on the same reasons as stated above. However, the value facets are extended because they need to account for the average speeds of pedestrians and cyclists:

- 0 km/h: The HRU is standing still.
- 4.4 km/h: Average speed of pedestrians [23].
- 17.5 km/h: Average speed of cyclists [24].

- 30 km/h: Low speed range.
- 50 km/h: Urban speed range.
- 130 km/h: Maximum permissible speed in most countries throughout Europe.

Longitudinal Distance (Headway). The longitudinal distance was included in the taxonomy, because the distance between the AV and the HRU may influence the objective safety and be related to accidents [14, 22]. Related to that, it may also affect the remaining communication time: the less distance to the HRU remains, the more urgent it is to communicate quickly. This will also influence the remaining time for the perception and interpretation of explicit and implicit communication signals. Three time states are required, because the longitudinal distance could change before (e.g. announcement of the interaction), during (e.g. adaptation of the interaction) and after the interaction (e.g. thank). The following three value facets are distinguished:

- <3 m: At a small distance a very fast external communication is necessary to communicate with faster HRUs, like vehicle drivers. Transmitting a massage can hardly be made in time through implicit communication.
- 3–10 m: At an average distance, it can be assumed that the external communication can be similar to that at low distances. On the other hand, the trajectory may no longer be recognized from a certain distance from slow HRUs (like pedestrians).
- >10 m: At a long distance, the external communication must be clearly visible. By using implicit communication to interact with slower HRUs, there is the danger that the pedestrians no longer refer to the trajectory because the distance is too large.

Lateral Distance. The lateral distance may have an effect on the subjective feeling of safety. As a result, the need for information could increase with a smaller lateral distance, for instance if a cyclist and a car drive on the same lane. In order to be able to specify all possible situations, three states are required: the lateral distance must be specified before, during and after the interaction. For example, the lateral distance changes if an AV drives from a driveway onto a road on which a manual vehicle is driving. Before the interaction, the AV is at a higher lateral distance from the manual vehicle. This distance decreases during the interaction. After the AV is on the road, the lateral distance between the AV and the HRU is negligible, because the AV is in front of or behind the HRU. The following three value facets are distinguished:

- 0 m: There is no lateral distance between AV and HRU, for example, when a manual driver drives behind or in front of an AV on the same lane. Therefore, no communication in lateral direction is necessary.
- ≤3 m: The lateral distance between AV and HRU is small. Therefore, the need for safety of the HRU is likely to be high and communication in the lateral direction is useful, for instance if a cyclist is on a cycling lane next to the road, the AV is driving on.
- >3 m: Due to the large lateral distance between AV and HRU it can be assumed that the two are on separate lanes. As a result, the communication should be very clear or is not necessary, depending on the direction of movement and the planned route of the AV and HRU.

Attention HRU. An interaction can only occur between AV and HRU if the attention of the HRU is directed at the AV. In case of missing attention, the AV's communication has to escalate in critical situations to draw the HRU's attention to itself and to the traffic situation. In particular, because HRU distraction contributes to a high accident risk [14–16]. The following value facets are possible:

- *Yes*: The attention of the HRU is given to the AV.
- *No*: The attention is not being given to the AV.

Impairment of the HRU's Perception. In the taxonomies which focus on the driver's perspective, a dimension of visual impairment exists [19], and perceptual disorder is mentioned by Böhm and Schneider [14] as a cause of accidents. The impairment category was extended for this taxonomy, because there can be not only a visual impairment but also an acoustic one. Depending on the situation, the AV needs to take further actions if the HRU has an impairment of perception. In the case of visual impairment, the AV has to communicate with sounds (e.g. if an HRU is driving backwards or if there is a barrier between the AV and HRU), or, in the other case with light signals or monitors (e.g. if the HRU is listening to music). As a result, there are four value facets:

- *View*: There is a visual impairment of the HRU
- *Acoustic*: There is an acoustic impairment of the HRU.
- *Both (View and Acoustic)*: There is visual, as well as acoustic impairment.
- *No Impairment*: There is no impairment of the HRU's perception.

Driving Direction AV. When driving backwards, drivers usually need to deal with a limited view through the rear window. Since other drivers know about the restricted view, they may expect a reduced speed of the backwards driving car. An AV has a good view in every direction, because of its sensors and can drive backwards as well as forward. Especially if the AV is not labelled as self-driving, the HRU's expectations regarding the speed of the AV whilst driving backwards and the real speed could differ. As a result, the subjective safety could decrease. Two value facets will be distinguished:

- *Driving forward:* The AV drives forward.
- *Reverse*: The AV moves backwards.

Perspective (from the Perspective of the AV). The AV has to communicate in several directions depending from which direction the HRU is approaching. It is important to check whether the HRU perceives and evaluates the explicit and implicit communication differently from various points of view. As in the case of the longitudinal distance, three time states are required, as the perspectives change before (e.g. communication to the front: announcement of the interaction), while (e.g. communication to the side: adaptation of the interaction) and after the interaction (e.g. communication to rear: ending the interaction). Following perspectives can be distinguished

- *Ahead:* The communication of the AV is directed forward, for example when letting a pedestrian cross the road.
- *Sideways/Diagonal*: The communication of the AV is directed sideways, for example when turning to the right, if a cyclist is located on the cycling track next to the AV.

- *Backward*: The communication of the AV is directed backwards, for example, when reversing or in the case of traffic jam with following traffic.

5 Application

The taxonomy is intended to support and structure research activities on the design of communication and interaction between AVs and HRUs. In this application example the taxonomy is used to the following question: *How quickly can the AV's intention regarding the right of way be expressed by using a targeted design of the trajectory?* Therefore, it is appropriate to vary the attributes "right of way" and the "AV's intention regarding right of way" (see Table 2).

Table 2. Possible combinations of the attributes "right of way" and "AV's intention regarding right of way" (it can be assumed that the AV drives rule based).

Right of way	AV's intention
AV	Let the HRU go first or Go first
HRU	Let the HRU go first
Undefined	Let the HRU go first or Go first

Furthermore, the attributes "HRU character" could influence the research question. A vehicle driver may understand the trajectory of an AV later or earlier than a pedestrian or a cyclist. This should be evaluated in different studies using the same trajectory of the AV in comparable situations (see Fig. 2). Therefore, it is important to know all attributes that could influence the interaction to keep them constant across all studies.

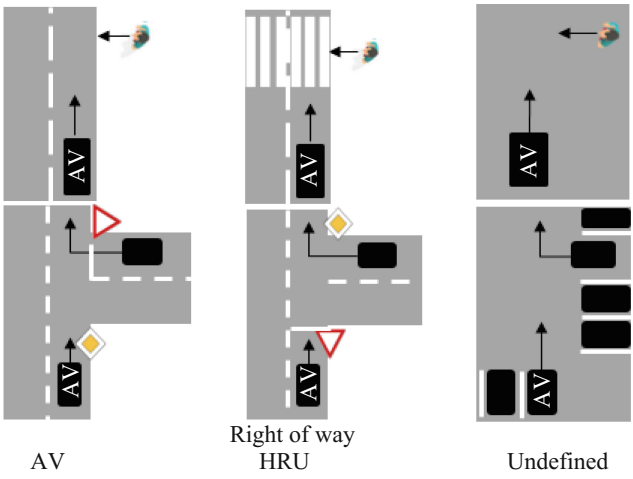


Fig. 2. Example of specific situations with pedestrians and vehicle drivers.

Another attribute that should be varied is the "longitudinal distance". Perhaps the HRU understands the planned trajectory earlier, if the AV begins to communicate with the HRU at a greater distance. Changing the "speed of the AV" could also influence the HRU's understanding of the trajectory and should therefore be varied. At a last step, the HRU's "attention" and the "impairment of the perception" may have an influence on the research results. Without paying attention or being restricted by an impairment, the HRU may need a longer time to understand the intention of the AV. The mentioned attributes need to be taken into account when answering the question. All other attributes should be kept constant across all studies.

Other research questions could evaluate the detectability of explicit communication, for which especially the perspective and the speed could be relevant. Furthermore, the time until collision could be evaluated, for which the speed and the distance should be manipulated systematically.

6 Limitations

The proposed taxonomy of traffic situations for the interaction between one AV and one HRU is based on discussions and reasonable assumptions, which were derived from real traffic situations. However, the taxonomy does not aim to be exhaustive. It is meant to give a first overview of relevant attributes and value facets, which should be modified by other researchers introducing further attributes and value facets. For example, the taxonomy could be extended to multiple road user situations.

7 Conclusion

A taxonomy of traffic situations, which classifies the interaction between an AV and a HRU was developed. It provides an overview over the most relevant attributes and related value facets that may influence the communication between the AV and the HRU. The overview of attributes facilitates the identification of potential confounders and increases the comparability between different studies.

References

1. SZ.de: BMW will selbstfahrende Autos in München testen. http://sz.de/1.3280168
2. Bengler, K., Zimmermann, M., Bortot, D., Kienle, M., Damböck, D.: Interaction principles for cooperative human-machine systems. Inf. Technol. **54**, 157–164 (2012)
3. Watzlawick, P.: Man kann nicht nicht kommunizieren. Das Lesebuch. Hogrefe, Bern (2016)
4. Bliersbach, G., Culp, W., Geiler, M., Heß, M., Schlag, B., Schuh, K.: Motive. In: Gefühlswelten im Straßenverkehr. Bonn (2002)
5. Witzlack, C., Beggiato, M., Krems, J.: Interaktionssequenzen zwischen Fahrzeugen und Fußgängern im Parkplatzszenario als Grundlage für kooperativ interagierende Automatisierung. In: VDI Wissensforum GmbH (ed.) Fahrerassistenz und automatisiertes Fahren. VDI-Berichte 2288, pp. 323–336 (2016)

6. Geyer, S., Baltzer, M., Franz, B., Hakuli, S., Kauer, M., Kienle, M., Meier, S., Weißgerber, T., Bengler, K., Bruder, R., et al.: Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance. IET Intell. Transp. Syst. **8**, 183–189 (2014)
7. Röhner, J., Schütz, A.: Psychologie der Kommunikation. Springer, Wiesbaden (2016)
8. Färber, B.: Kommunikationsprobleme zwischen autonomen Fahrzeugen und menschlichen Fahrern. In: Maurer, M., Gerdes, J.C., Lenz, B., Winner, H. (eds.) Autonomes Fahren, pp. 127–146. Springer, Berlin (2015)
9. Schulz von Thun, F.: Miteinander reden 1. Störungen und Klärungen. Allgemeine Psychologie der Kommunikation. Rowohlt E-Book, Reinbek (2013)
10. Zimmermann, M., Bauer, S., Lutteken, N., Rothkirch, I.M., Bengler, K.J.: Acting together by mutual control: evaluation of a multimodal interaction concept for cooperative driving. In: International Conference on Collaboration Technologies and Systems (CTS), pp. 227–235 (2014)
11. Jiang, X., Wang, W., Mao, Y., Bengler, K., Bubb, H.: Situational factors of influencing drivers to give precedence to jaywalking pedestrians at signalized crosswalk. Int. J. Comput. Intell. Syst. **4**, 1407–1414 (2011)
12. Lehsing, C., Kracke, A., Bengler, K.: Urban perception-a cross-correlation approach to quantify the social interaction in a multiple simulator setting. In: 2015 IEEE 18th International Conference on Intelligent Transportation Systems-(ITSC 2015), pp. 1014–1021
13. Weller, G., Schlag, B.: Kriterien zur Beurteilung von Fahrerassistenzsystemen. In: BDP-Kongress für Verkehrspsychologie (ed.) (2002)
14. Böhm, H., Schneider, W.: Verkehrsteilnehmergruppen und Verkehrserziehungsmittel. In: Ministerium Wirtschaft, Mittelstand Verkehr Land NRW, Forschungsgemeinschaft (ed.) Der Mensch im Verkehr. Köln (1965)
15. Shinar, D., McDonald, S.T., Treat, J.R.: The interaction between driver mental and physical conditions and errors causing traffic accidents: an analytical approach. J. Saf. Res. **1**(10), 16–23 (1978)
16. Otte, D., Kühnel, A., Suren, E.G., Weber, H., Gotzen, L., Schockenhoff, G., Vu Han, V.: Erhebungen am Unfallort. In: Bundesanstalt für Straßenwesen (ed.) Unfall- und Verkehrssicherheitsforschung Straßenverkehr, 37. Köln (1982)
17. Hacker, W.: Allgemeine Arbeitspsychologie. Huber, Bern (1998)
18. Bengler, K., Pütz, A., Purucker, C., Götze, M., Färber, B.: Das UR:BAN Assistenzszenario als Konstrukt für die Entwicklung und Bewertung integrierter Fahrerassistenzsysteme. Z. Arb. wiss. **69**, 65–70 (2015)
19. Benda, H. V.: Die Skalierung der Gefährlichkeit von Straßenverkehrssituationen, I Teil: Ein Klassifikationssystem für Verkehrssituationen aus Fahrersicht. FP 7320, im Auftrag der Bundesanstalt für Straßenwesen, TU München, Lehrstuhl für Psychologie (1977)
20. Fastenmeier, W.: Die Verkehrssituation als Analyseeinheit im Verkehrssystem. In: Fastenmeier, W. (ed.) Autofahrer und Verkehrssituationen - Neue Wege zur Bewertung von Sicherheit und Zuverlässigkeit moderner Straßenverkehrssysteme, 33. Verlag TÜV Rheinland, Köln (1995)
21. Wachenfeld, W., Winner, H.: Use Cases des Autonomen Fahrens. Gemeinsames Ergebnis des Kernteams Villa Ladenburg Kollegs „Autonomes Fahren“ (2014)
22. Statistisches Bundesamt: Verkehr. Verkehrsunfälle. Wiesbaden (2016)
23. Federal Highway Administration: Manual on uniform traffic control devices (2003)
24. Panis, L.I., de Geus, B., Vandenbulcke, G., Willems, H., Degraeuwe, B., Bleux, N., Mishra, V., Thomas, I., Meeusen, R.: Exposure to particulate matter in traffic. A comparison of cyclists and car passengers. Atmos. Environ. **44**, 2263–2270 (2010)

Effect of Warning Levels on Drivers' Decision-Making with the Self-driving Vehicle System

Wenjie Cui[1], Ronggang Zhou[1(✉)], Yunhao Yan[1], Linghua Ran[2], and Xin Zhang[2]

[1] School of Economics and Management, Beihang University, Beijing, China
zhrg@buaa.edu.cn
[2] Ergonomics Laboratory, China National Institute of Standardization, Beijing, China

Abstract. This paper introduces a systematic study on the influence of interface warning levels on users' decision-making in self-driving vehicles. Three scenes are designed in the driving simulator with the warning level interfaces displayed in each scene in a sequence. The lateral and longitudinal control of the driving simulator can be managed either by an automated controller in the self-driving mode or by the driver in the manual mode. A simulated driving experiment is performed with 23 participants. Their driving behavior data under different warning level interfaces are collected for empirical research. Results show that in the self-driving mode, the first level warning interface produces significant effect on users' driving behavior. Therefore, they usually take over control of the vehicle more quickly, the average speed becomes lower and the larger deviation of the vehicle from road center may appear after the take-over.

Keywords: Human factors · Self-driving vehicle · Decision-making · Warning level

1 Introduction

In recent years, self-driving vehicles have received increasing attention with the development of the Internet and intelligent technologies. Although the self-driving technology has some potential disadvantages, it will bring tremendous changes to our lives. First of all, self-driving vehicles are safer than manual ones. According to the statistics released by WHO, traffic accidents bereave 1.25 million people of their lives per year in the world [1], and this figure is expected to increase to 2.2 million by 2030. In the United States alone, about 33 thousand deaths are caused each year by traffic accidents. The popularity of self-driving vehicles will substantially reduce traffic accidents. According to a study by the Eno Centre for Transportation, if 90% of the vehicles on American roads are self-driving vehicles, the annual number of traffic accidents is expected to drop from 6 million to 1.3 million, along with a reduction of casualties from 33 thousand to 11.3 thousand [2].

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_69

Secondly, self-driving vehicles may contribute to environmental protection. With optimized acceleration, braking and energy recycling systems, self-driving vehicles can help reduce traffic congestions as well as energy consumption and emissions [3–5]. In addition, self-driving vehicles can save time for drivers. According to a report from McKinsey, 80% Americans averagely spend 50 min commuting each day. With self-driving vehicles, drivers can make use of this time to relax or work. McKinsey predicted that self-driving vehicles could save 1 billion hours a day for drivers in total [6]. In addition, self-driving vehicles can relieve traffic congestions and bring more convenience to people who are physically impaired.

In recent years, self-driving vehicles have attracted global attention, thanks to the research and development by Google, Tesla, Baidu and other companies (see USA Today [7–9], BBC News [10], Daily Mail [11]). Indeed, as suitably highlighted by McKnight and McKnight [12], maintaining a vehicle in the road center and ensuring a steady speed is perhaps not the most difficult part of the driving task, and would indeed be relatively easy to achieve on an empty road, even for novice drivers. However, problems arise when the automated vehicle is required to interact with more complex road environments, and the technology currently reaches only L2 or L3 level automation (SAE) [13]. Therefore, human decisions are quite important under special circumstances.

Until recently, studies have emerged on driver behavior in self-driving vehicles of L2 and L3 levels. For example, studies with ACC (Adaptive Cruise Control) suggest changes in workload, when ACC is compared to manual driving. Conclusions drawn by these studies are conflicting, with both reduced [14, 15] and increased [16] workloads are reported. Studies with ACC have also shown reductions in situation awareness (SA), and stress [17, 18]. In addition, drivers' ability to regain control of the vehicle is better if they are expecting automation to be switched off [19]. Drivers' visual attention to the road center decreases [20], driving performance degrades [21], and other activities are more engaged in as the level of automation increases [22].

However, research into the effect of warning levels in self-driving vehicles of the L3 and L4 is still inadequate. Quantitative research in this regard is particularly meaningful. In this study, the effect of warning levels on the decision-making of drivers is examined. Indicators such as the non-takeover rate (the rate of failing to resume manual control), average decision time (the average time for the drivers to resume manual control when they receive the warning), average vehicle speed and vehicle lateral offset (distance to road border) are considered to research this area.

2 Method

2.1 Participants

In order to investigate the effect of warning levels on decision-making of drivers, an experiment is designed with 23 participants (12 males and 11 females) aged between 18 and 55, and their average annual mileage is 2,546 km. All the participants are drivers who have been licensed for more than 3 years in the People's Republic of China.

2.2 Experiment Design

In this section, the experiment design is explained. A driving simulator is used to simulate the driving experience. The interface of the simulator is shown in Fig. 1. Since the objective of this experiment is to study the effect of warning levels on the decision-making of drivers, three warning levels are designed with different remainders, colors and countdowns, as can be seen in Table 1.

Fig. 1. Interface of the driving simulator.

Table 1. Remainders, colors and countdowns of each warning level.

Warning levels	Remainder	Color	Countdown (s)
Level 1	Please take over immediately	Red	10
Level 2	Please get ready to take over	Green	30
Level 3	Please consider taking over	Blue	None

Three scenes are designed with the warning levels displayed in each scene in a sequence generated based on the Latin square theory. The sequence of warning levels displayed in each scene is as follows: Scene 1, 123123; Scene 2, 312312; Scene 3, 231231. These warning levels serve as the independent variables. In the experiment, participants are required to drive 7,800 m in a simulated driving environment with the speed limit of 50 km/h. The simulated driving environment is a two-lane two-way road with six roadblocks, including accident scenes, traffic cones and speed bumps. The distance between every two blocks is 1,300 m. When vehicles approach these roadblocks, corresponding warning levels will be displayed to warn the driver, in which Warning Level 1 corresponds to the accident scene, Warning Level 2 to the traffic cone and Warning Level 3 to the speed bump. According to the displayed warning levels, participants can decide whether or when to take over control of the vehicle. When the

vehicle passes all the roadblocks, the on-board system will switch to the self-driving mode. To enhance the experience of participants in the self-driving mode, they are demanded to read a certain article on the cellphone, thus distracting their attention from the vehicle. Each participant takes part in two scenes in the experiment, and rests for five minutes in between. The dependent variables in the experiment are indicators designed to evaluate the driving behavior, including the non-takeover rate, average decision time, average vehicle speed and vehicle offset. The non-takeover rate records the probability of situations when drivers fail to resume the control of the vehicle when warning is displayed. The average decision time is adopted to compare the time for the driver to respond to the warning in different warning levels. The vehicle offset is to measure the distance between the vehicle and road center after the warning is displayed. The route and traffic are kept the same for all participants, and each participant conducts the experiment independently, thus unaffected by other participants.

2.3 Experiment Equipment

The equipment engaged in this experiment includes a three-screen computer, an automatically controlled steering wheel and a control board that can overtake the throttle and braking system of the vehicle. The UC-win/Road Professional 10.2 program is run on the vehicle to record the speed, offset and takeover status of the vehicle. An iPhone 6s cellphone is used as the mobile device provided for participants to read articles while driving.

2.4 Experiment Procedures

The procedures of the experiment are explained as follows. First, participants are introduced with the general information of the experiment and are asked to sign an informed consent form if they approve the conditions as listed in the form concerning information collection, confidentiality and revocation rights. They are also asked to fill in a questionnaire about their personal information and driving experience.

Thereafter, participants are seated on the driving simulator. The height and position of the seat and steering wheel are adjusted according to the driving habit of each participant. Participants are required to obey the traffic rules and limit the vehicle speed to 50 km/h. A ten-minute rehearsal is performed before the experiment to familiarize participants with the equipment, vehicle condition and road condition. Afterwards, the experiment starts and participants conduct the simulated driving in two scenes with a 5-minute break in between.

3 Results and Discussion

The experimental data of 23 participants are collected, and a statistical analysis of the data is conducted using SPSS 22.0. In order to analyze the effect of warning levels on decision-making of drivers, the four aforementioned indicators are observed, namely the non-takeover rate, average decision time, average vehicle speed and vehicle offset.

A 1-way Analysis of Variance (ANOVA) is adopted to analyze the experimental data in which F and P indices are used to identify the differences between data.

3.1 The Non-takeover Rate

This study obtains 276 decision-making records from the experimental data. Each of the 23 participants is supposed to make six decisions in each scene (23 * 2 * 6 = 276). These data can be further classified into three levels according to the warnings displayed to drivers; therefore, in each level there are 92 records. In general, there are 266 records of overtake actions undertaken by participants when they receive the warnings, accounting for 96% of all warnings. The non-takeover rate in each warning level is calculated and shown in Fig. 2. It can be concluded that when warnings are displayed, almost all participants will take over control of the vehicle, and the non-take over rate increases from Level 1 to Level 3.

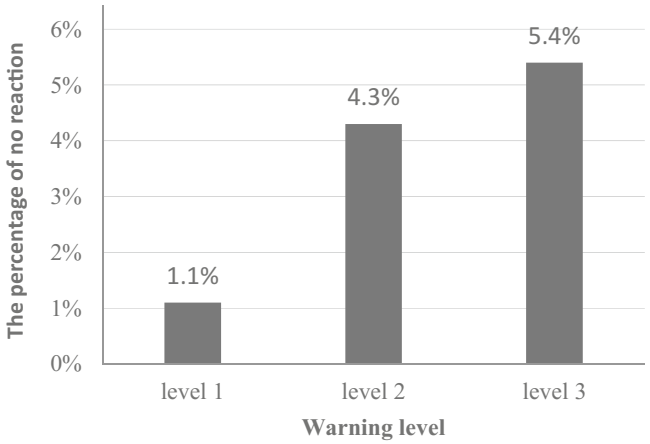


Fig. 2. Non-takeover rate in each warning level.

3.2 The Average Decision Time

Decision time refers the time between the appearance of warning and the drivers' takeover action in the self-driving mode. It reflects drivers' responses to different warning levels. The average decision time in different warning levels is illustrated in Fig. 3.

According to the decision time, the warning level has a significant main effect ($F = 10.36$, $P < 0.001$). Based on univariate analysis of the general linear model, the average decision time of Warning Level 1 and Level 2 shows significant difference ($P < 0.001$). The difference between average decision time of Level 2 and Level 3 is not significant ($P = 0.904$). The decision time when drivers face different warning levels shows significant difference ($P < 0.05$).

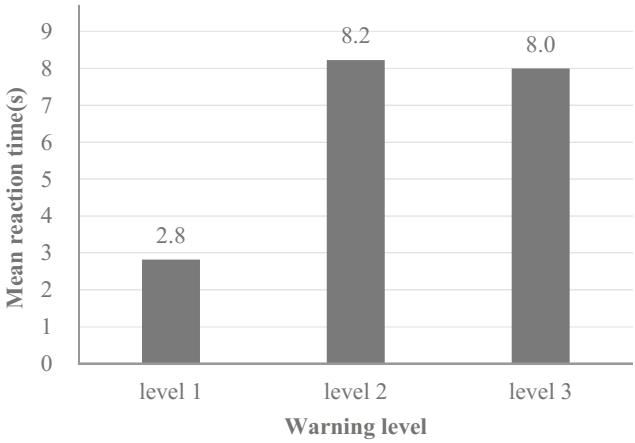


Fig. 3. Average decision time in different warning levels.

Next, the result of descriptive statistical analysis indicates that in Warning Level 1, it takes drivers only 2.8 s on average to respond to the warnings, while in Level 2 and Level 3, the average decision time is about eight seconds.

To sum up, the Warning Level 1 has the greatest influence on drivers' behavior. The decision time of drivers when facing Warning Level 1 is much shorter, which means they tend to take over and decelerate more quickly. In contrast, Level 2 and Level 3 have less influence on drivers' behavior. No significant difference can be identified between Level 2 and Level 3. This indicates that drivers have a quicker decision-making behavior (take over or not take over) when they face high-level warnings.

3.3 The Average Speed

The average speed of the vehicle is a fundamental index to reflect the running state of the vehicle, and it is the most effective indicator of the driving situation.

In order to compare the influence of each warning level on driving behavior, the vehicle speed is observed in four stages according to the distance to the roadblock, as shown in Fig. 4. The average vehicle speed at each stage is presented in Table 2.

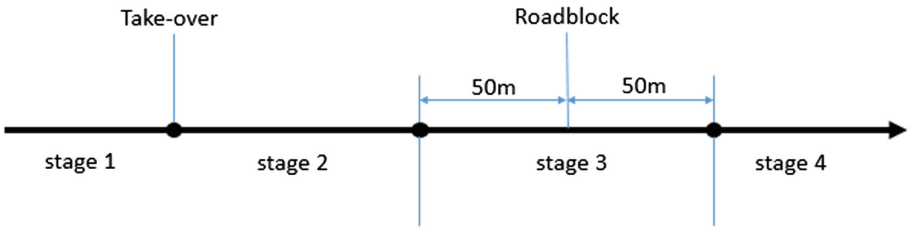


Fig. 4. The four stages near the roadblock

Table 2. The average speed (km/h) of the three warning levels during the four stages

Stage	Warning level		
	Level 1	Level 2	Level 3
Stage 1	50.00	50.00	50.00
Stage 2	41.42	49.07	49.73
Stage 3	36.01	44.91	45.07
Stage 4	51.75	55.22	53.67

The changes of the average speed at different stages in each warning level show a V-pattern trend, indicating that drivers decelerate first and then accelerate. The average speed at Stage 3 is much smaller than that at other stages. It suggests that most drivers exhibit sudden behavior of deceleration when they pass the roadblocks. The average speed of different levels is significantly different ($F = 8547.568$ (Level 1), $F = 2937.114$ (Level 2), $F = 2465.114$ (Level 3), $P < 0.001$). The average speed of stages in a given level also shows a significant difference ($P < 0.001$), except for Stage 1 and Stage 2 ($P = 0.681$) in Warning Level 3.

Then, we compare the average speed of different warning levels at each stage. At each stage, the average speed in Warning Level 1 is smallest compared to that of other warning levels. The average speed in different warning levels at each stage is significantly different ($F = 2499.189$ (Stage 2), $F = 3328.967$ (Stage 3), $F = 3328.967$ (Stage 4); $P < 0.001$).

In summary, a more urgent warning has a greater impact on drivers' behavior, resulting in a larger magnitude of vehicle deceleration. The average speed of Stage 3 in each warning level is smaller than other stages; hence, we need to pay more attention to Stage 3 when designing the warning interface.

3.4 The Lateral Offset

This experiment requires drivers to drive in the middle of the lane. However, influenced by various factors, the vehicle always have some lateral offset. Therefore, we recorded the left and right lateral offsets as indicators to reflect the control ability of drivers.

Left lateral offset refer to the distance between the vehicle and the left side of the road, and right lateral offset could be calculated from left offset, vehicle width and lane width. Based on the data of left and the right lateral offsets, the lateral offsets (distance between vehicle and center of the road) can be derived. The experimental data of lateral offsets at different stages of each level are shown as Table 3, in which AVG refers to

Table 3. Lateral offsets (m) of three stages in each warning level

Stage	Warning level					
	Level 1		Level 2		Level 3	
	AVG	STD	AVG	STD	AVG	STD
Stage 2	2.37	2.64	1.75	1.79	1.96	1.98
Stage 3	3.75	2.16	3.62	1.63	3.12	2.00
Stage 4	1.34	1.46	1.59	1.49	1.52	1.37

the average lateral offset, and STD the standard deviation. At Stage 1, since the vehicle is always kept at road center by the self-driving mode, there is no lateral vehicle offset. Therefore, the data of Stage 1 are omitted in Table 3. If the lateral offset becomes larger after the take-over action, the warning has a greater influence on drivers' behavior, or vice versa.

The lateral offsets of different stages in each level are significantly different ($F = 21859.474$ (Level 1), $F = 18366.423$ (Level 2), $F = 10045.799$ (Level 3); $P < 0.001$). The average lateral offset at Stage 3 is much larger than that at the other two stages. The average lateral offset of stages in a given level also shows much difference ($P < 0.001$).

Then, we compare the lateral offset of different warning levels at each stage. At Stages 2 and 3, the average lateral offsets of Warning Level 1 is the largest. Nevertheless, the average lateral offset in Level 1 is the smallest at Stage 4. The reason for this situation is that drivers still maintain their vigilance. In other words, they are still concentrated on the road after they pass the roadblocks of Warning Level 1. The average lateral offsets in different warning levels at each stage are significantly different ($F = 1927.606$ (Stage 2), $F = 509.962$ (Stage 3), $F = 515.320$ (Stage 4); $P < 0.001$). The average lateral offsets between each two warning levels also shows a significant difference ($P < 0.001$).

In summary, a more urgent warning level can lead to a greater impact on drivers' behavior, resulting in a bigger magnitude of steering. This conclusion can also be suggested by the standard deviation of lateral offsets. Standard deviation of lateral offsets in Warning Level 1 is biggest compared with other levels at Stage 2 and Stage 3. In addition, we need to pay more attention to Stage 3 when designing the warning interface because the average lateral offset at Stage 3 is biggest compared with the average lateral offset at other stages in each warning level.

4 Summary and Outlook

The purpose of this paper is to explore the influence of warning levels on the driver's decision-making behavior in the self-driving situation.

Three warning levels are designed to remind drivers of different levels of emergency, and four indicators are used to quantify the influence. As the urgency of warning increases, the average decision time, non-takeover rate and average vehicle speed reduce while the lateral vehicle offset increases, indicating that a more urgent warning leads to a greater influence on driver's decision-making behavior.

This study provides significant guidance for the design of self-driving vehicle interface. In spite of its fruitful findings, further research is still necessary. The effect of more factors, such as warning color and countdown time, can be considered. The eye focus of drivers can be included as an indicator to reflect their decision-making behavior. Drivers' emotional and psychological reaction can also be considered using the multichannel physiologic device. Moreover, the stages of this experiment can be further divided to include more details. For example, Stage 3 can be divided into two sub-stages, covering the 50 m before and after the roadblock.

Acknowledgments. This study was work supported by the National Natural Science Foundation of China (71640034 and 31271100) and the National Key Technology R&D Program (2014BAK01B01).

References

1. World Health Organization. http://www.who.int/mediacentre/factsheets/fs358/en/
2. Eno Centre for Transportation. https://www.enotrans.org/wp-content/uploads/2015/09/AV-paper.pdf
3. Barth, M., Boriboonsomsin, K., Wu, G.: Vehicle automation and its potential impacts on energy and emissions, p. 103 (2014)
4. Morrow III, W.R., Greenblatt, J.B., Sturges, A., Saxena, S., Gopal, A., Millstein, D., Shah, N., Gilmore, E.A.: Key factors influencing autonomous vehicles' energy and environmental outcome, p. 127 (2014)
5. Brown, A., Gonder, J., Repac, B.: An analysis of possible energy impacts of automated vehicle, p. 137 (2014)
6. Mckinsey & Company. http://www.mckinsey.com/industries/automotive-and-assembly/our-insights/ten-ways-autonomous-driving-could-redefine-the-automotive-world
7. USA Today. http://www.usatoday.com/story/tech/talkingtech/2017/01/19/waymo-blog-stop-shaming-humans/96797988/
8. USA Today. http://www.usatoday.com/story/money/cars/2017/01/12/uber-open-self-driving-car-lab-near-detroit/96501386/
9. USA Today. http://www.usatoday.com/story/tech/news/2017/02/10/ford-bets-1b-self-driving-car-startup/97745230/
10. BBC News. http://www.bbc.com/news/technology-38839071
11. Daily Mail. http://www.dailymail.co.uk/money/investing/article-4269054/Cash-driverless-car-revolution.html
12. McKnight, A.J., McKnight, A.S.: Motor vehicle technology: automation of driving tasks (2003)
13. Society of Automotive Engineers. http://www.sae.org/servlets/pressRoom?OBJECT_TYPE=PressReleases&PAGE=showRelease&RELEASE_ID=3544
14. Ma, R., Kaber, D.B.: Situation awareness and workload in driving while using adaptive cruise control and a cell phone. Int. J. Ind. Ergon. **35**, 939 (2005)
15. Stanton, N.A., Young, M.S.: Driver behaviour with adaptive cruise control. Ergonomics **48**, 1294 (2005)
16. Rudin-Brown, C.M., Parker, H.A.: Behavioural adaptation to adaptive cruise control (ACC): implications for preventive strategies. Transp. Res. F Traffic Psychol. Behav. **7**, 59 (2004)
17. Merat, N., Jamson, A.H.: Is drivers' situation awareness influenced by a highly automated driving scenario?
18. Parasuraman, R., Molloy, R., Singh, I.L.: Performance consequences of automation-induced 'complacency'. Int. J. Aviat. Psychol. **3**, 1 (1993)
19. Merat, N., Jamson, A.H., Lai, F.C., Daly, M., Carsten, O.M.: Transition to manual: driver behaviour when resuming control from a highly automated vehicle. Transp. Res. F Traffic Psychol. Behav. **27**, 274 (2014)
20. Carsten, O., Lai, F.C., Barnard, Y., Jamson, A.H., Merat, N.: Control task substitution in semiautomated driving: does it matter what aspects are automated? Hum. Factors **54**, 747 (2012)

21. Strand, N., Nilsson, J., Karlsson, I.M., Nilsson, L.: Semi-automated versus highly automated driving in critical situations caused by automation failures. Transp. Res. F Traffic Psychol. Behav. **27**, 218 (2014)
22. Merat, N., Jamson, H.A., Lai, F., Carsten, O.: Human factors of highly automated driving: results from the EASY and CityMobil projects. In: Meyer, G., Beiker, S. (eds.) Road Vehicle Automation, p. 113. Springer, Berlin (2014)

Does Shifting Between Conditionally and Partially Automated Driving Lead to a Loss of Mode Awareness?

Anna Feldhütter(✉), Christoph Segler, and Klaus Bengler

Chair of Ergonomics, Technical University of Munich, Boltzmannstr. 15, 85748 Garching, Germany
{anna.feldhuetter, christoph.segler, bengler}@tum.de

Abstract. In near future, several complex automation modes – like SAE-level 2 and 3 – may be employed in one vehicle. In order to investigate the relevance of mode awareness and mode errors in the context of vehicle automation, a study with 49 participants was conducted. In the experiment, the participants experienced two stages: one stage with alternating partially and conditionally automated driving and another stage with only partially automated driving. Mode awareness and the occurrence of mode errors were compared in the two stages in order to examine the effect of shifting between the two modes. Additionally, the influence of a cognitive-auditive and a visual-motoric non-driving-related task as well as an adapted HMI was examined. Results showed that depending on the type of the non-driving-related task shifting between partially and conditionally automated driving leads to a loss of mode awareness and results in more mode errors compared to having only one automation mode. An enhancement of mode awareness by the suggested adapted HMI could not be found.

Keywords: Conditional automation · Partial automation · Mode awareness · Mode errors · Human-machine interaction · Human-machine interaction concept

1 Introduction

The technical progress in vehicle automation promises many advantages, such as an enhancement of road safety and efficiency as well as an increase in drivers' comfort by reducing their workload. However, the larger number of assistance systems and automation modes and the subsequent rise in the complexity of the technical system also entails risks. It is well known that the availability of multiple assistance and automation modes in one technical system may lead to a lack of mode awareness [1]. Especially if the driver is not aware of the current automation mode and its functional structure or assumes a different mode is currently activated, incorrect driver behavior can occur, which is referred to as mode error. This can in turn lead to dangerous situations or even accidents. In the aviation sector, a lot of research has been carried out, as several incidents, among them severe accidents, occurred with an increase in the number of modes. They were caused by a lack of mode awareness or by mode errors [2] (cited after Dornheim [3]). In this context, Sarter and Woods [2, 4] have identified

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_70

the following important factors that encourage insufficient mode awareness/mode confusion: insufficient/incorrect monitoring of the automation due to absent/poor feedback, lack of understanding of the functional structure of the automation, high level of automation and shifting between automation modes as well as a high degree of coupling between several automation modes. Especially in time-critical situations, insufficient mode awareness may result in mode errors, as the pilots recognize the error belatedly or not at all due to a high workload [5]. Several approaches were developed to either enhance mode awareness (e.g. adopting the interface concept dependent on the automation mode or implementing acoustic and haptic feedback for shifting between different modes [6, 7]) or to prevent the occurrence of mode errors [1]. In contrast to the aviation sector, almost no research has been carried out that investigates mode awareness and error in the automotive automation context. In particular, partially automated driving (PAD, level 2, according to SAE International [8]) and conditionally automated driving (CAD, level 3) are similar for the user regarding the obvious functionality, however, the driver's role differs significantly: In partially automated vehicles the driver still needs to monitor the system, whereas in conditionally automated vehicles the drivers can engage themselves in non-driving-related tasks [8] (NDRT). It is, therefore, important in terms of controllability to investigate whether shifting between these two modes leads to a lack of mode awareness and to the occurrence of mode errors. For this purpose, a driving simulator experiment in a fix-based driving simulator was conducted. Further, an adapted human-machine interface (HMI) was evaluated, which aimed at enhancing mode awareness and thus preventing mode errors.

2 Methods and Apparatus

2.1 Simulation and Vehicle Automation

The experiment was conducted in a fix-based driving simulator. The simulator was equipped with an active steering wheel, pedalry, a projector for displaying the driving environment and a touchscreen for performing the NDRT. The driving simulation and the different automation modes were programmed using the Unity 3D 5.4.0 game engine (Unity Technologies) and the Visual Studio 2015 programming environment (Microsoft Corporation).

In order to investigate mode awareness in two different automation modes, the Automated Lane Change Test (ALCT) of Spießl [9] was adapted. The ALCT was derived from the standardized Lane Change Test (LCT), which was developed for measuring drivers' distraction by a NDRT in the context of manual driving [10]. The LCT consists of a simulated drive on a three-lane highway. At regular instances, road signs appear which specify a certain lane. The driver is instructed to steer the simulated vehicle onto this required lane (see Fig. 1). In the ALCT, the automated system takes over longitudinal and lateral control, scans the signs and changes lanes if necessary. In the experiment, the automation maintained a constant velocity of 60 km/h. The distances between the signs ranged between 100 m and 140 m. Even though the signs were visible very early on, the texture did not become clear before the

defined distance of 60 m was reached in order to ensure that the participants had the same reaction time for each sign. At the same time when the texture became clear, the automation initiated the according maneuver. In the CAD mode, the system carried out each maneuver as indicated by the signs. In the PAD mode, the system committed errors at a defined number of signs, meaning that the automation selected the wrong lane. If the participants recognized the automation error, they could switch off the automation by oversteering the system and correct the error. After reaching the correct lane, the participants were instructed to switch on the automation again. As a system reliability of less than 90% leads to a decrease in trust [11] (cited after [9]), the error rate in the partial automation mode was set to 10%. The position of the automation errors was randomized across all participants. The number of signs and the track length were the same for all participants.

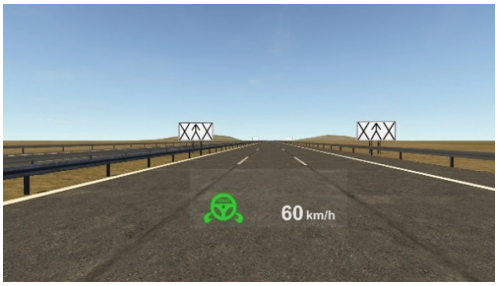


Fig. 1. Automated Lane Change Test and head-up display in the PAD mode.

2.2 Non-driving-Related Tasks

It was assumed that an adaption of the drivers' reaction to the two automation modes depends on the type of NDRT. Therefore, two different types of NDRT were tested in the two automation modes: a visual-motoric and an auditive-cognitive task. Both tasks were designed to be interruptible in order to give the driver the possibility to monitor the system if necessary. Further, a baseline for testing the two HMI concepts with no additional task was conducted.

The visual-motoric task consisted of a destination-entering task (DET) on a simulated navigation system screen (see Fig. 2). The destination included a city and a street, which needed to be entered consecutively. Each destination had to be entered only once. The task was performed on a touch screen, which was located in the middle console. The application of this visual-motoric task aimed at examining the influence of this task type on mode awareness. In addition, the effect of a supporting adaptive human-machine interface implemented on the same screen as the task was investigated.

The auditive-cognitive task consisted of the Twenty Questions Task (TQT), inspired by the study of Merat, Jameson, Lai and Carsten [12]. The participants were encouraged to guess an item from a predefined category by asking a maximum of twenty questions, the response to each question always being 'no' or 'yes'. The question-answer style of this task is similar, for instance, to a complex telephone

Fig. 2. The destination-entering task in German language. The orange-labeled text displays the required city (*Ort: Muenchen*) and street (*Strasse: Knorrstrasse*)

conversation, as it utilizes cognitive resources but can be properly replicated. In addition to the effect of this type of NDRT on mode awareness, the influence of an adaptive HMI on mode awareness was also investigated.

2.3 Procedure

The procedure of this study is shown in Fig. 3. Before the experiment, the participants were informed about the ALCT and the capacity of both automation modes, PAD and CAD, as well as the corresponding responsibility of the driver according to SAE International [8]. Subsequently, the participants completed a questionnaire, which asked for demographics and experience with automated driving and assistance systems. To examine the effect of different HMIs (initial and adaptive), the participants were subdivided into two equally large groups (between-subject design). Before beginning the experimental drive, the participants familiarized themselves with the driving simulator, including the two automation modes (PAD and CAD), the ALCT and the two NDRTs.

After completion of the familiarization drive, the experiment started. The experimental drive was subdivided into two stages: the two-mode stage and the one-mode stage (see Figs. 3 and 4). Each participant completed both stages, counterbalancing the order of the stages between the participants in order to avoid sequence effects. In the two-mode stage, the automation mode changed between CAD and PAD several times (see Fig. 4). After driving manually onto the three-lane highway and activating the conditional automation, the participants started with the first NDRT. After driving 1.5 km, the system initiated a mode change to PAD, which needed to be confirmed by the participants. Afterwards, the participants drove 2.0 km in partially automated mode, followed by a further mode change to CAD. After completing both mode changes, the next NDRT started. This procedure was repeated twice for each NDRT and once without any NDRT (baseline), permuting the order of both NRDTs and the baseline, which led to a total stage duration of 21 min. During the 15-minute one-mode stage no mode changes occurred. As in the two-mode stage, the participants completed each NDRT twice and the baseline once. In order to examine the effect of mode transitions,

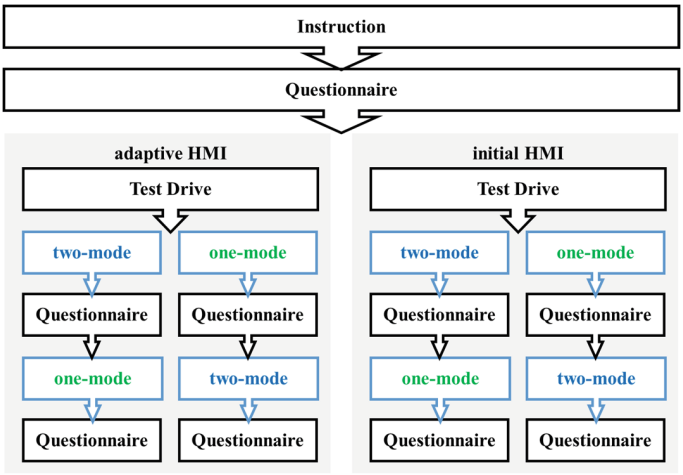


Fig. 3. Study procedure

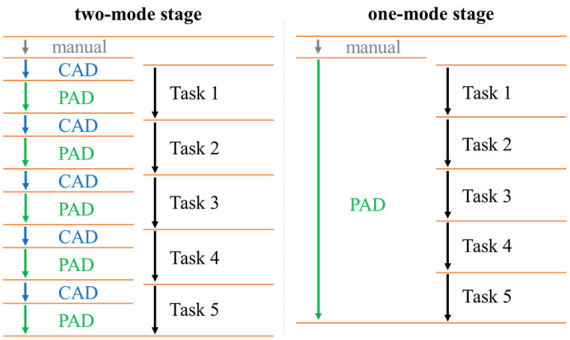


Fig. 4. Procedure of the two-mode and one-mode stage. The two different non-driving-related tasks (destination-entering task and Twenty Question Task) were performed twice and the baseline without any task once per stage. The order of the tasks and the baseline was permuted and the same task was not performed consecutively.

mode awareness was compared in the one-mode stage and the PAD periods of the two-mode stage (within-subject design).

After each stage, the participants rated their subjective mode awareness, the distraction by the NDRT and the performance of either monitoring the automation (in the PAD mode) or of the NDRT. The analysis of this rating is not part of this paper.

2.4 Human-Machine Interface

The HMI consisted of two parts: the head-up display (see Fig. 1), which was available to all participants during the entire experiment, and the adaptive HMI, which was only

available to the adaptive HMI group. Figure 5 shows the symbols used for indicating the automation modes in the present study. In the one-mode stage, only two symbols were necessary, as only partial automation and manual driving (MD) was available. An activated partial automation mode was indicated by the green PAD symbol. In the case of a change to manual driving, the white MD symbol was displayed. In the two-mode stage, the blue symbol for CAD and the symbols for mode transition were added. If CAD was available, in addition to the green PAD symbol, the text Autopilot available was displayed and a one-time acoustic signal sounded. After confirmation with a specific button on the steering wheel, the CAD mode was activated (blue CAD symbol was displayed). In the case of a transition from CAD to PAD, the CAD symbol flashed red, and the acoustic signal gonged continuously until the participants confirmed the transition with the button on the steering wheel. All transitions were triggered by the system.

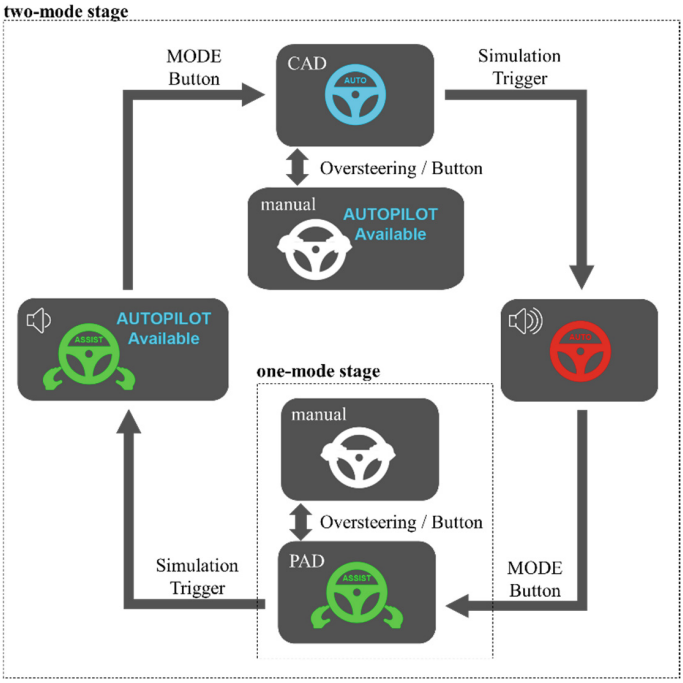


Fig. 5. Symbols used for indicating the automation modes

While the initial HMI only consisted of acoustic feedback and the above-mentioned information in the head-up display, the adaptive HMI additionally featured the same information on the same screen where the DET was presented (see Fig. 6). During PAD and MF, the afore-mentioned symbols were displayed in the left upper corner of the screen. Only in CAD, the blue CAD symbol was supplemented with a blue declining bar, which showed the remaining time in this mode. This implementation was

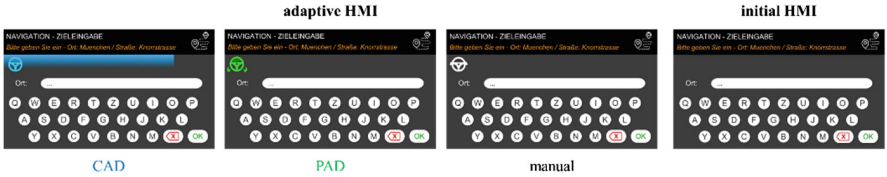


Fig. 6. Adaptive human-machine interface (left) and initial human-machine interface (right)

based on the approach of Monk [6] and Sellen et al. [7], as different symbols were used depending on the automation mode.

2.5 Dependent and Independent Variables

In order to assess mode awareness, the monitoring performance during PAD was measured by the dependent variables (DV) reaction time to an automation error (time between the appearance of the sign texture and the intervention by the participant), the proportion of undetected automation errors and the proportion of false interventions. These DVs can be analyzed for all three conditions of NDRT (TQT, DET and no additional task). Further, mode awareness was evaluated by means of the task performance in entering the destination, which was assessed by the mean hitting rate for entering the city or street names (hits per second while entering one word), the mean pause between entering a city and a street and the mean pause between entering two destinations. These metrics indicate the attention directed toward the driving environment during PAD, as the NDRT needed to be interrupted for this purpose. In addition to measuring the task performance of the DET, eye-tracking was used for this NDRT to evaluate the distraction from the driving environment and, hence, to assess mode awareness. Therefore, the gaze-shifting proportion to either the driving surroundings or the NDRT, the mean gaze duration of one gaze towards either the driving environment or the NDRT and the mean gaze-shifting rate between the driving environment and the NDRT were measured.

The effect of two independent variables (IV) on mode awareness was tested: the HMI concept and mode transition. In order to examine the effect of different HMI concepts, the above-mentioned DVs were compared for the two groups, where one group tested the initial HMI concept and the other group the adaptive one. For assessing the effect of mode transition on mode awareness, the DVs were compared for the one-mode stage (no mode transitions) and the PAD period of the two-mode stage (with mode transitions).

3 Results

A total of 50 volunteers participated in the experiment. One participant had to be excluded from data analysis due to technical issues during the experiment. Of the remaining 49 participants, 15 were female (31%), and the mean age of all participants was 34.65 years (SD = 9.92). Eleven participants (23%) had never experienced

automated driving or driver assistance systems, whereas 22 participants (45%) use these systems occasionally and 16 (33%) regularly.

To evaluate the effect of mode transition and HMI on mode awareness, a 2 (HMI, between-subject design) × 2 (stage, within-subject design) analysis of variance (ANOVA) was conducted for each NDRT condition (TQT, DET, none). In the case of a significantly differing variance between two groups, Welch's T-Test was carried out. Throughout the entire analysis, a significance level of .05 was assumed. Due to different technical issues, further participants had to be excluded from data analysis. The number of participants varied for each NDRT condition.

Results of the ANOVA showed that there was no significant effect of the HMI concept on mode awareness considering all NDRT conditions.

Figure 7 shows the reaction times to an automation error in the two stages (one-mode and two-mode), depending on the NDRT. Analyzing the no-task (N = 42) and the DET (N = 47) condition, the ANOVA revealed no significant effect of mode transitions, considering the reaction time to an automation error and the proportion of undetected automation errors. In the no-task condition there was no false intervention, so this metric could not be analyzed. In the DET condition, two participants intervened erroneously while the PAD mode was activated (one in the one-mode stage and one in the two-mode stage). Thus, this metric could also not be analyzed for the DET condition.

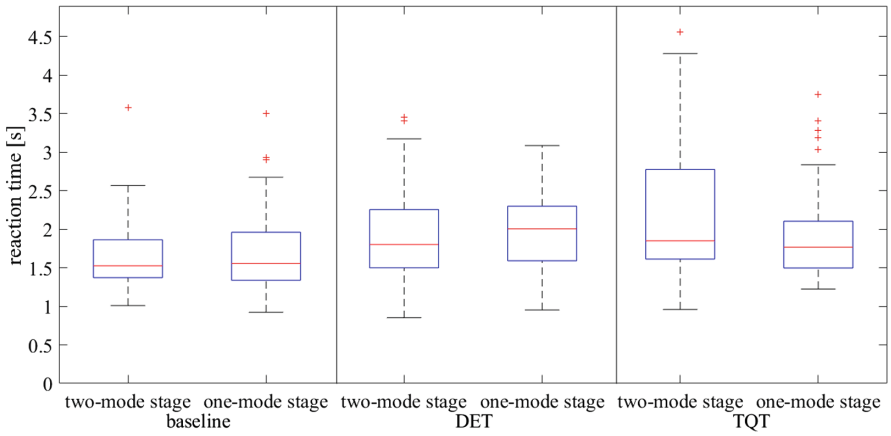


Fig. 7. Average reaction time to an automation error depending on the non-driving-related task and the stage

Considering the DET performance (hitting rate, pause between entering a city and a street and pause between entering two destinations), there was also no significant effect of mode transitions. Processing the eye-tracking data resulted in a remaining sample size of 29 participants, where both the number of participants in the two HMI groups and the order of the two stages were still balanced. The ANOVA showed that regarding the gaze-shifting proportion to either the driving environment or the DET, no

significant difference between the stages emerged. Welch's Test (ANOVA: $F(1, 27) = 34.83$, $p < .001$, $r = .60$), Levene Test: $p = .001$) revealed a significant effect of the mode transition on the mean gaze duration towards the DET ($t(39.22) = 5.62$, $p < .001$, $r = .67$). Hence, the participants on average looked 3.89 s longer at the screen of the DET in the two-mode stage than in the one-mode stage (Fig. 8). Consistently, the ANOVA also showed that the mean gaze-shifting rate between the driving environment and the DET was significantly less in the two-mode stage ($F(1, 27) = 9.39$, $p = .005$; $r = .17$).

Evaluating mode awareness for the TQT ($N = 48$) results of the ANOVA revealed that the reaction time to an automation error was significantly higher in the stage with two automation modes ($\Delta t = 0.31$ s) than in the stage with only one mode ($F(1, 46) = 7.07$, $p = .011$, $r = .18$, see Fig. 7). The analysis of undetected automation errors did not show any differences between the stages. The number of false interventions could not be analyzed, as only one participant corrected the automation erroneously (in the one-mode stage).

Throughout the entire analysis, no significant interaction effects were found.

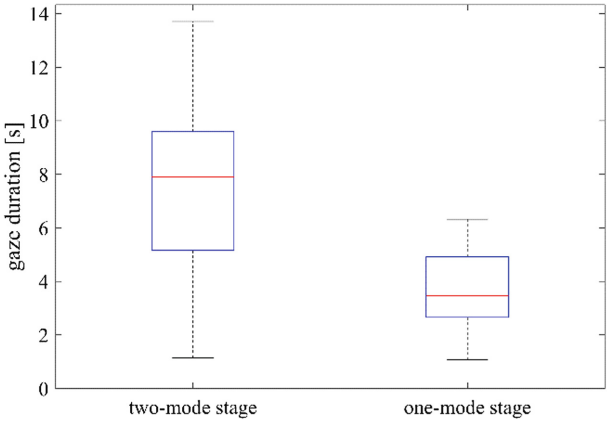


Fig. 8. Mean gaze duration in partially automated driving towards the destination-entering task depending on the two different stages.

4 Discussion

4.1 Effect of Mode Transition on Mode Awareness

With regard to the effect of mode transition on mode awareness, the results differed according to the type of NDRT. With no additional task, based on the variables investigated, there were no differences in mode awareness between the stages, with only one or two alternating automation modes (CAD and PAD).

When the DET (visual-motoric task) was performed, a significant effect of mode transition on the visual monitoring performance was found: Changing the modes from CAD to PAD and vice versa led to significantly stronger distraction from the driving

environment, as the participants averted their eyes from the road for a significantly longer period of time, compared to when only PAD was activated. Consequently, this effect could also be found in the gaze shifting rate, as in the two-mode stage the participants shifted their gaze back to the road significantly less often. It is assumed that the participants simply did not change their gaze behavior when shifting from CAD to PAD due to a lack of mode awareness. However, the different gaze behavior in the two stages did not lead to differences in reaction times in the case of an automation error. This is plausible, though, as looking away from the driving environment would probably be more likely to result in missing signs and therefore in an increase of undetected automation errors. This assumption cannot be definitely confirmed, however. A possible reason may be that the number of automation errors per participant was relatively low, with an error rate of 10%. Additionally, even if the participants missed some signs, the probability that an automation error occurs at one of the missed signs was relatively low. Hence, the sample size available for the analysis was relatively small. Even though no change in reactions to automation errors could be found, the decrease in visual monitoring of the partly automated system and in checking the driving environment due to a lack of mode awareness represents a safety-relevant issue, as the driver would detect suddenly emerging automation errors late or not at all.

When the TQT (auditive-cognitive task) was performed, a significant increase in reaction times to automation errors was detected during the stage with the two alternating automation modes. The number of undetected automation errors could not confirm this decrease in performance. However, here, too, the sample size was relatively small. There are already several studies that have found a deterioration of driving performance caused by auditive-cognitive secondary tasks such as talking on the phone [13, 14] or using the onboard voice assistant [15] due to limited mental resources. As the TQT always started during CAD, the participants were able to use the majority of their attentional resources for performing this task. With transition to PAD, the main task shifted to monitoring the automation. Obviously, in the two-mode stage the participants did not adapt their behavior by diverting enough attentional resources back to the monitoring task, as in this stage longer reaction times to automation errors occurred. It is assumed that the participants were not aware of the current automation mode PAD and did therefore not adequately adapt their behavior. However, it is also conceivable that the participants did not remember their responsibility to monitor the system in PAD. Further, it cannot be ruled out that, due to the safe driving simulator environment, the participants were more willing to take risks and, hence, concentrated more strongly on the NDRT.

4.2 Effect of Different HMI Concepts on Mode Awareness

The adaptive HMI did not lead to an improvement of mode awareness in comparison to the initial HMI. An interview after the experimental drive revealed that only nine participants recognized the adaptive HMI, which might be one reason why no differences were found. A further reason for the lack of improvement may be that the modification of the HMI was not conspicuous enough. An extended application of the automation symbol colors (green, blue and red) in the vehicle interior, using for

instance LED concepts, would be conceivable. Apart from improving mode awareness, a reasonable further approach could be, according to Sarter and Woods [1], to prevent the occurrence of mode errors. Results of this study showed that it seems to be crucial for mode awareness which type of NDRT is performed by the driver. It might be expedient, as a further alternative approach, to interrupt these specific tasks, depending on the duration of engagement in the tasks or on the driver state regarding mode awareness. However, this could lead to low acceptance among users.

5 Summary

Results of this driving simulator study showed that shifting between the PAD and CAD automation modes affected mode awareness for two different types of NDRT. While performing a cognitive-auditive NDRT (Twenty Question Task), the alternating mode transitions led to longer reaction times when an automation error occurred. While performing a visual-motoric NDRT (destination-entering task), the mode alternating transitions resulted in less monitoring and less checking of the partially automated system, as the time spent looking at the driving environment and the automation decreased significantly. Without an additional task, no effect on mode awareness was found when shifting between the modes.

It can therefore be stated that depending on the task type, shifting between the PAD and CAD automation modes leads to a loss of mode awareness and results in more mode errors compared to having only one automation mode.

The suggested adaptive HMI concept, which was based on the proposals of Monk [6] and Sellen et al. [7] and which was compared to the initial concept, did not improve mode awareness, considering both the two task types and the control group without an additional NDRT.

This study identified that the increase in the number of automation modes and in system complexity can lead to a loss of mode awareness. In particular, when the PAD and CAD modes are employed in one vehicle, the occurrence of mode errors is a safety-critical issue, as the responsibilities of the driver are significantly different. It is therefore important to develop a suitable HMI which supports the driver optimally in being aware of the current automation mode and to further examine under which circumstances mode errors may occur. Further studies should have a closer look on the reported effects.

References

1. Sarter, N.B., Woods, D.D.: How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Hum. Factors **37**, 5–19 (1995)
2. Sarter, N.B.: Investigating mode errors on automated flight decks. Illustrating the problem-driven, cumulative, and interdisciplinary nature of human factors research. Hum. Factors **50**, 506–510 (2008)
3. Dornheim, M.A.: Dramatic incidents highlight mode problems in cockpits. Aviat. Week Space Technol. **142**(5), 57–59 (1995)

4. Sarter, N.B., Woods, D.D.: Pilot interaction with cockpit automation II. An experimental study of pilots' model and awareness of the flight management system. Int. J. Aviat. Psychol. **4**, 1–28 (1994)
5. Sarter, N.B., Woods, D.D.: Mode error in supervisory control of automated systems. Hum. Factors **36**, 26–29 (1992)
6. Monk, A.: Mode errors. A user-centred analysis and some preventative measures using keying-contingent sound. Int. J. Man Mach. Stud. **24**, 313–327 (1986)
7. Sellen, A., Kurtenbach, G., Buxton, W.: The prevention of mode errors through sensory feedback. Hum. Comput. Interact. **7**, 141–164 (1992)
8. SAE International: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. SAE International, Warrendale (2014)
9. Spießl, W.: Assessment and Support of Error Recognition in Automated Driving. Universitätsbibliothek der Ludwig-Maximilians-Universität, München (2011)
10. ISO 26022: Road vehicles—Ergonomic aspects of transport information and control systems—Simulated lane change test to assess in-vehicle secondary task demand (2010)
11. Buld, S., Krüger, H.-P.: Wirkungen von Assistenz und Automation auf Fahrerzustand und Fahrsicherheit. Projekt: EMPHASIS, Effort-Management und Performance-Handling in sicherheitsrelevanten Situationen. Abschlussbericht, Würzburg (2002)
12. Merat, N., Jamson, A.H., Lai, F.C.H., Carsten, O.: Highly automated driving, secondary task performance, and driver state. Hum. Factors **54**, 762–771 (2012)
13. Caird, J.K., Willness, C.R., Steel, P., Scialfa, C.: A meta-analysis of the effects of cell phones on driver performance. Accid. Anal. Prev. **40**, 1282–1293 (2008)
14. Strayer, D.L., Johnston, W.A.: Driven to distraction: dual-task studies of simulated driving and conversing on a cellular telephone. Psychol. Sci. **12**, 462–466 (2001)
15. Strayer, D.L., Cooper, J.M., Turrill, J., Coleman, J.R., Medeiros-Ward, N., Biondi, F.: Measuring cognitive distraction in the automobile. In: AAA Foundation for Traffic Safety, 53ff (2013)

Information Expectations in Highly and Fully Automated Vehicles

Cyriel Diels[1(✉)] and Simon Thompson[2]

[1] Centre for Mobility and Transport, Coventry University, Coventry CV1 5FB, UK
Cyriel.Diels@coventry.ac.uk
[2] Jaguar Land Rover, Whitley CV3 4LF, UK
sthom261@jaguarlandrover.com

Abstract. Vehicle automation fundamentally changes the relationship with our vehicles whereby system trust and occupant comfort will become key user requirements. We here explored what information users expect to receive across a wide range of traffic scenarios in urban and peri-urban environments for both highly and fully automated vehicles. Results indicated that users expected the HMI to provide information to promote system transparency, comprehensibility, and predictability. Visual presentation using Head Up Display technologies was the preferred method but older users in particular expressed a preference for auditory interfaces. In hazardous or uncertain situations, participants expressed a desire to be able to override or confirm vehicle actions. Results indicated that information requirements might however change with experience signifying the potential role for adaptive and configurable HMI solutions for future vehicles.

Keywords: Vehicle automation · Human machine interaction · User requirements

1 Introduction

Vehicle automation fundamentally changes the relationship with our vehicles. System trust and occupant comfort are arguably the most significant human factors challenges. In the context of future HMI concepts that facilitate trust and comfort, we explored what information people expect to see, where and when, and how this information should be presented.

In this study, we considered both highly and fully automated vehicles and included a range of traffic scenarios in both urban and peri-urban environments. *Highly automated vehicles*: In this study, the definition of highly automated vehicles largely corresponded to BASt level 3 [1]. Participants were informed that the driver would not be required to permanently supervise the system whilst in this mode. Whereas the driver can do other tasks such as reading emails, the driver must be awake/alert at all times. The system will take responsibility for the operation of the vehicle whilst in this mode, but may request the driver to resume control within a short period of time, e.g. less than 10 s. In case the driver fails to regain control, the vehicle would bring itself to a safe stop. Finally, participants were informed that the vehicle provides automation on discrete parts of a

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_71

journey that requires no driver supervision, but may not be able to complete the journey 'mission' without driver intervention. *Fully automated*: Fully automated driving followed BASt level 4 [1] definition. Participants were informed that the vehicle would be in control under all conditions and that the vehicle would not have any primary or secondary vehicle controls with reference being made to autonomous pods.

Since automated vehicles are not yet available in the public realm, users have no experience with such vehicles and may find it difficult to imagine possible interactions. This also limits the validity of studies looking into the acceptance and expected interactions with such vehicles using survey based approaches. Typically, new technologies with which people have not had any experience with are under- or overvalued [2, 3]. Furthermore, surveys tend to refer to "vehicle automation" but may not differentiate different automation concepts such as highly or fully automated vehicles. Of course, different automation levels can be expected to also refer to different user expectations in particular when also taking into account different traffic scenarios.

In an attempt to circumvent some of these limitations, we created an immersive environment consisting of a vehicle buck and large field of view projection screen to display the different traffic scenarios. To raise awareness of the absence of any primary controls in fully automated vehicles, participants were seated in the passenger side of the vehicle buck. In the highly automated vehicle condition in which vehicle control is still required when automation is not available, participants were seated in the driver seat. Finally, participants received clear explanations and illustrations of the different automation levels and were shown specific traffic scenarios.

2 Methods

2.1 Participants

A total of 38 participants took part in the study consisting of young (18–25 yrs; n = 18), middle (30–50 yrs; n = 14) and older (60+; n = 6) participants with an equal gender split in each age category. All participants held a full UK driving license for at least 2 yrs.

2.2 Study Design

A within subjects design was employed with each participants asked to respond to all traffic scenarios in both the highly and fully automated vehicle conditions.

2.3 Facilities and Traffic Scenarios

Participants were seated in a vehicle buck consisting of the front half only (see Fig. 1). Traffic scenarios were projected onto a screen 2 m in front of the vehicle projecting an image with a horizontal and vertical angle of 60 and 40° of visual arc, respectively. Associated in-vehicle sound was played via speakers located behind the participants. They were shown a total of 22 traffic scenarios which each consisted of 20 s real- world

footage taken on urban and peri-urban roads. In the fully automated vehicle condition 12 scenarios were presented, 10 scenarios were presented in the highly automated vehicle condition (see Table 1). The presentation order of the different scenarios were randomized to control for any order effects.

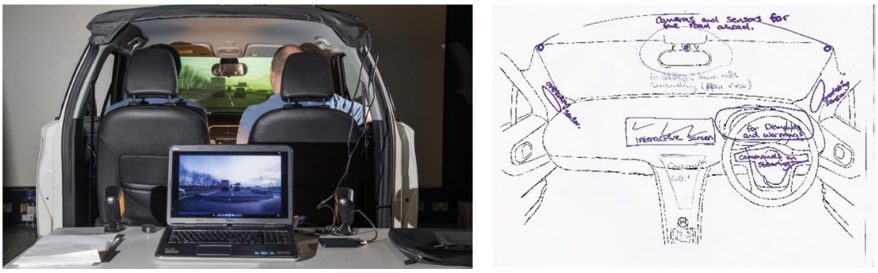


Fig. 1. Left: Experimental setup showing the vehicle cab and projected traffic scenarios. For the highly and fully automated scenarios, participants sat in the driver and passenger seat, respectively. Right: Example of participant drawing illustrating what information they expected to see where.

Table 1. Overview of traffic scenarios for the fully (left column) and highly automated (right column) conditions.

Fully automated	Highly automated
01 – Vehicle changes lane in order to overtake a slower vehicle	13 – Vehicle enters the motorway and can be switched into autonomous mode
02 – Vehicle in slowing traffic congestion	14 – Vehicle receives an unexpected exit request from the user who instructs the vehicle to exit the motorway
03 – Vehicle sets off and enters the lane from the rest area	15 – Vehicle HMI of route/expected exit/notification
04 – Vehicle takes a motorway exit	16 – Vehicle exits the motorway and requests to handover to manual control to the user
05 – Vehicle approaches a speed restriction area, due to roadworks	17 – Vehicle approaches an unexpected scenario (e.g., no road markings or V2X) (Bin Lorry)
06 – Vehicle is overtaken by another car	18 – Vehicle approaches crosswalk with pedestrians crossing
07 – Vehicle approaches an unexpected scenario (e.g., no road markings or V2X) and the AV comes to a safe stop or expects user take over control	19 – Vehicle approaches pedestrians not crossing on the crosswalk
08 – Vehicle approaches crosswalk with pedestrians crossing	20 – Vehicle approaches pedestrians walking towards the autonomous vehicle
09 – Vehicle approaches pedestrians crossing the road	21 – Vehicle approaches cyclist travelling in same direction
10 – Vehicle approaches pedestrians walking towards the vehicle	22 – Vehicle approaches traffic lights
11 – Vehicle approaches cyclist travelling in same direction	
12 – Vehicle approaches traffic lights	

2.4 Measures

Following each scenario, participants were asked to indicate what information they expected to see, where and when, and how this information should be presented (e.g. type of display, modality, location, frequency, duration). In addition, participants were also encouraged to sketch their ideas on a generic sketch of a vehicle interior (for an example see Fig. 1 (right)).

2.5 Procedure and Participant Instructions

Participants were welcomed and introduced to the concepts of highly and fully automated vehicles. Participants were asked to sit in the driver or passenger seat depending on whether they started with the highly or fully automated vehicle condition, respectively. They were then shown an example traffic scenario and were informed as to the questions they would be asked in response and were shown the generic sketch of a vehicle interior which they could use to illustrate their answers. Participants were encouraged not to feel restricted by their knowledge of current HMI and to consider that information could be communicated using any modality (vision, haptic, auditory, olfactory, proprioception) or combination thereof at any location in the vehicle. They were encouraged to think beyond the traditional instrument cluster and centre display to present information.

3 Results

For brevity, we here present a selection of the results which illustrate some common findings across scenarios. Figure 2 (left) shows screen shots of the traffic scenario 9 in which a fully automated vehicle approaches pedestrians crossing the road. Figure 2 (right) shows the types of information participants expected to receive in scenario 9. The vast majority (34 out of 38) of participants expected to receive some sort of acknowledgement that the vehicle had seen the pedestrians. Fourteen participants expected to receive information regarding the proximity of people to the car. This type of

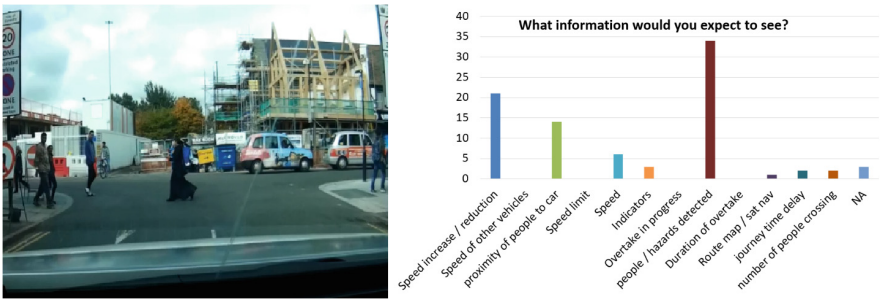


Fig. 2. Screen shot (left) and frequency and type of information expected to see in scenario 9 "Fully automated vehicle approaching pedestrians crossing the road"

information can be categorized as conveying *situation awareness*, i.e. what the vehicle "sees". 21 participants expected the vehicle to communicate information on upcoming maneuvers, i.e. information conveying *behavioural awareness*, i.e. informing the occupant what the vehicle is going to do.

When asked *where* people would expect to receive this information, the following locations were mentioned in decreasing order of importance: centre display, windscreen Head Up Display (HUD), user display, i.e. display mounted in front of occupant, audio, user HUD, i.e. HUD in from of occupant, visor HUD, i.e. HUD in visor location of windshield, and A-pillar. Of interest is to note that older participant category expressed a preference for the use of audio. The majority of participants expected to receive this information visually and/or auditory, with only few participants referring to haptic feedback. When asked *when* this information should be displayed, all participants indicated to expect the information immediately on detection of hazards/people. Finally, participants were asked what interactions they would expect. A considerable proportion expected to interact with an interface (touchscreen, voice control) to instruct the vehicle to proceed with caution. About a third of participants expected an emergency stop pedal or button to be available, while about an equal number of only eleven participants did not expect to see any interaction method at all.

Table 2. Overview of main findings related to information expectations in highly and fully automated vehicles

Information expectations highly and fully automated vehicles
• Vehicle intentions related to hazardous situations
• Upcoming vehicle maneuvers
• Speed limit in speed restricted area
• Safe to proceed
• Hazard detection
• Ability to override or confirm vehicle actions in hazardous/uncertain situations

The above exemplifies many of the responses observed across the different traffic scenarios as well as the two automation levels. In summary, information expectations were related to system transparency, comprehensibility, and predictability. Table 2 summarizes the main findings. In hazardous situations such as approaching an intersection involving other road users, participants expected the vehicle to inform them of its intentions. Participants also expected the vehicle to provide information with regards to upcoming maneuvers such as changing lanes or taking a motorway exit. Participants expressed an interest in being shown the speed limit in speed restricted areas even though the system was in charge of setting and driving at the appropriate speed. The detection, identification of hazards and its visualization was a consistent requirement among scenarios, automation levels and participants. A large number of participants expressed the desire to be able to override or confirm vehicle actions in hazardous/uncertain situations. Finally, participants indicated that information requirements might change with experience.

4 Discussion

In this study we explored information requirements for a range of traffic scenarios in both urban and peri-urban environments. We considered both highly and fully automated vehicles and used a participant sample of young, middle-aged, and older participants. The overall results indicated that users expected the HMI to provide information to promote system transparency, comprehensibility, and predictability.

One of the most consistent requirements found was the detection, identification of hazards and visualization of these and this was consistent across scenarios, automation levels, and participants. Among the more surprising finding was that participants expressed an interest in being shown the speed limit in speed restricted areas even though the system was in charge of setting and driving at the appropriate speed. A large number of participants also expressed the desire to be able to override or confirm vehicle actions in hazardous/uncertain situations. Taken together, these results would suggest a lack of trust in such automated vehicles, which can be understood given that the participants had no experience with the performance of automated vehicles. It suggests that in particular during the introductory phase, user acceptance may benefit from providing quite extensive amount of information [4, 5]. As also indicated by participants, with experience and growing trust in automated vehicles, the need for such information may diminish and points towards adaptive or personalized HMI solutions. This would also help to avoid user frustration and annoyance.

Results related to presentation mode indicated that the vast majority of participants expected information to be presented visually either in a center display, windshield Head Up Display, or instrument cluster. For vehicle control transition scenarios, there was a preference to use the instrument cluster which can be understood given the fact that current driving task related information is provided here and suggests to co-locate such information under these conditions.

While acknowledging the relatively small sample size (n = 6), it is noteworthy that older participants expressed a preference for audio feedback as opposed to visual information. Although this study did not allow to delve into further details as to the reasoning, it is worth considering that the largely visual HMI approach currently being proposed may not be the most suitable or desirable option for all users. Furthermore, most participants would have no previous experience with HMI implementations that make use of tactile and proprioceptive information, which may able to provide alternative solutions.

In the context of information location, it should be mentioned that in the current study there was no reference to the idea of being able to engage in leisurely or economically productive non-driving related tasks (NDRT) such as reading, watching movies, or emailing. There may be a potential trade-off between system trust and the ability to comfortably engage in NDRT [6]. That is, providing information via the HMI to develop, calibrate, and maintain trust may interfere with the ability to comfortably engage in NDRT leading to annoyance, frustration, and an overall compromised user experience. Therefore, understanding the relative importance of the different trust factors over comfortably engaging in NDRT and how to incorporate these into the user interface will become an important line of research.

References

1. Gasser, T.M., Westhoff, D.: BASt-study: definitions of automation and legal issues in Germany. In: Presentation at the TRB Workshop on Road Vehicle Automation, Irvine, CA (2012)
2. Skippon, S., Garwood, M.: Responses to battery electric vehicles: UK 9 consumer attitudes and attributions of symbolic meaning following direct 10 experience to reduce psychological distance. Transp. Res. D **16**(11), 525–531 (2011)
3. Vilimek, R., Keinath, A.: User-centred design and evaluation as a prerequisite 5 for the success of disruptive innovations: an electric vehicle case study. In: 6 Driver Acceptance of New Technology: Theory, Measurement and Optimization, vol. 7, pp. 169–186 (2014)
4. Payre, W., Cestac, J., Delhomme, P.: Fully automated driving impact of trust and practice on manual control recovery. Hum. Factors **58**(2), 229–241 (2016). doi:10.1177/0018720815612319
5. Payre, W., Cestac, J., Delhomme, P.: Impact of training and in-vehicle task performance on manual control recovery in an automated car. Transp. Res. F. (2017). doi:10.1016/j.trf.2017.02.001
6. Miglani, A., Diels, C., Terken, J.A.M.: Compatibility between trust and non driving related tasks in UI design for highly and fully automated driving. In: Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (2016)

Obtaining Design Requirements from the Public Understanding of Driverless Technology

Patrick Langdon[1(✉)], Ioannis Politis[1], Mike Bradley[1], Lee Skrypchuk[2], Alex Mouzakitis[2], and John Clarkson[1]

[1] Department of Engineering, Engineering Design Centre, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK
{pml24,ip332,mdb54,pjc10}@cam.ac.uk
[2] Research & Technology, Jaguar Land Rover, Coventry, UK
{LSKRYPCH,AMOUZAK1}@jaguarlandrover.com

Abstract. Despite a large literature on autonomous vehicles and a large number of relevant surveys examining public perception of self-driving vehicles, recent studies suggest that reported public perception is rapidly changing, volatile, and highly dependent on context and potential biases in reporting. In addition, some conclusions from very recent studies are supported: namely, that very little of the survey literature addresses key user centred design consideration that arise from inclusion of the wider population in the use of autonomous vehicles. We report initial studies that are intended to form part of a method currently under development that aims to bring together automotive design and Inclusive design; an analytical functional approach to design for user capability in the wider population. This paper describes the development stages of a user centred design capable methodology for sampling public opinion from a range of different qualitative studies. The finding are made available at two levels. Firstly, a general report on public perceptions', examines key elements of sensitivity in responses to autonomous vehicle issues. Secondly, the methodology shows how to address in an unbiased way the specification and design a number of potential HMI concepts for managing takeover from car to driver in an autonomously capable vehicle in urban and highway usage cases.

Keywords: Human factors · Inclusive design · Surveys · Autonomous cars · Public opinion

1 Background and Motivation

Although there is a burgeoning and long standing literature on Self Driving Cars (SDC) and their effect on society [1] a good deal of this literature has been carried out within the domain of the research lab, or specialists in socio economics, transport researchers, or Business information [2] reporting.

Older survey studies have reported predictions regarding the prevalence of specific technologies that were found to be highly conservative. Recent studies [1, 3–7] suggest that reported public perception is divided, rapidly changing, and highly dependent on context and potential biases in reporting [2]. Hence, older survey and laboratory studies

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_72

are likely to be less informative about current public opinion due to the effect of public exposure to publicity. This has come from various sources, including OEMs, advertising, the provision and advertising of automated and autonomous feature in new market offerings and a number of new initiatives aimed at putting autonomous vehicles onto UK roads (e.g. Venturer; UK Autodrive) [8, 9] and their surrounding media comment.

There are, however, a number of very recent reviews of public opinion and social implications of attitudes towards autonomous vehicles. Kyriakidis et al. [10] presented the most recent academic review of public opinion and report on a 5000 person study carried out in 109 countries in 2015. Noting the diversity of responses, they found that manual driving was preferred and that 33% thought full automation likely to be enjoyable. However, most concern was about hacking, legal and safety issues, at different levels of sophistication depending on demography. Respondents saw high automation as premium product worth in excess $30,000, which appears to reflect a specific view of technology as the domain of the high earning bracket.

A very recent study carried out in the UCL Transport Institute report for the Department of Transport [2], reports that urban men more likely to be interested in AV than women and that older people were less likely to be interested in AV, in general. They found that interest varied with Geographical area and that those enjoying driving less were more likely to be interested in the possibilities of AVs. The study identified three key factors in the public perception of AVs:

1. Safety and cybersecurity
2. Perceived usefulness and benefits
3. Cost (should be less than conventional cars)

These themes summarise the main considerations reported by many studies and were well matched to the present thematic analysis, taking safety and security to be effectively "trust". These themes can also be seen repeatedly in a number of other recent studies.

The THINKGOODMOBILITY 2016 [11], survey of Autonomous Vehicles sampled 12000 respondents and 48 people in focus groups in four countries. They found that 44% were uncomfortable with AVs, or in sharing the road with AV (41%). When asked about activities people expected to do while autonomous, 42% stated text, 37% internet, 28% reading, 27% e-mail, 19% sleep, and 18% video.

Of the whole sample 43% felt that AV were safer than driving and 19% did not, with the remainder (62%) neutral. This large neutral figure may be a characteristic of recent sampling in surveys as around 60% stated that AVs did not know enough to judge. They also felt that AVs did not have enough "common sense" for a human driver mix. The majority, 73%, thought AVs could malfunction, and 70–80% thought that the driver should be in control and should have access to a steering wheel. A large response, 82% *preferred to be aware of the road at all times*.

The very recent Venturer Project carried out a survey based on understanding the socioeconomic adoption scenarios for autonomous vehicles based on key concepts from Socio-technical regimes [12] and Technology Acceptance models [13]. Aimed at level of awareness and general attitudes to AV, they made a number of findings summarised here. They found a clear awareness of AV and generally positive expectations of their benefits. The driving task was still enjoyed and manual control was still seen positively,

while AVs were associated with a loss of control. There were doubts over safety but agreement that benefits may be in boring, undemanding driving, not complex urban manoeuvring.

Currently therefore, despite many surveys with comparable findings, recent studies suggest that it is not clear what the public understanding may be of driverless technology and how it is perceived as part of their lives in the near and middle future time scales. Very little work has focused on the totality of the issue from the users' perspective, which concerns the specific technology, and extends into both the likely users in society and, importantly, what journeys they make. This will be related to peoples' usage of vehicles, the journeys they make, and the relationship between these and the lifestyles they are able to lead in their home locations. Their perceptions may well be quite distinct from that of technology developers, governance and service planners. Very little rigorous research has been devoted to the task of identifying potential users' understanding, opinions, needs, and preferences for the *specifics* of new generation of autonomous vehicle technology. In particular, there is a need to understand how effective human machine interfaces (HMI) can be designed to best provide the driver safe control when either handing control to the car or taking control back from the car in urban and highway driving situations [14]. This may be the key to establishing technological acceptance in both driving and non-driving elements of the population.

2 Methodology

The collected data was essentially qualitative and hence methodology was driven by the principles of qualitative research. For this to be effective, a range of sources was necessary using differing methods such as interviews, focus groups, surveys and field studies. Each of these sources were fully documented; fully referenced and described, and human data taken with Video and Audio recordings and transcription of interactions. The outcomes in Initial interviews, focus groups and research were followed by an interpretation and consolidation stage where a triangulated interpretation was formed taking into account all the data sources [15, 16]. All sources of data and the extent of their agreement and disagreement were decided, moderated by the methodological strength of their origins. For example, survey data was compared for agreement with thematic analysis but themes common only to one were included or abandoned depending on the strength of the data collection method. It was necessary to use a compressed format for the outcomes of the triangulated interpretation stage. A convenient representation of this sort was a concept or requirements table. In addition, once the Thematic Analysis for both Interviews and Focus Groups had been completed, interpretation of the resulting themes provides a set of key insights, and these are reported.

The project addressed the problem through a human centered design approach using qualitative research techniques [15–17]. The analysis of transcribed data was made using standard social science techniques to analyse the data using qualitative research approaches; such as that of Grounded Theory [16]. A general qualitative design tools NVIVO [18], yielding a thematic analysis used during different stages of interpretation. The method followed the plan shown in Fig. 1.

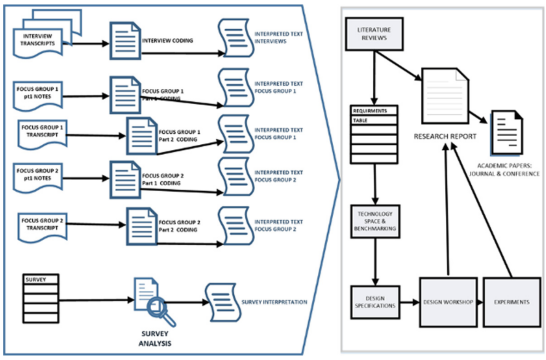


Fig. 1. Overview of methodology plan. Other aspects; technology review, focus groups and the design workshop stage are reported elsewhere [Ibid 14]

Interview participants were sampled to cover an appropriate range of the population that it is intended to generalise the results to. Some criteria for choice were established based on age, gender, demographic group and geographical region and these were matched to the traveller needs survey; described below, on a quota basis. No respondents with strong impairments of vision, physical or hearing capability were used and the sampling was chosen to include, young, late middle-aged males and females from metropolitan and rural areas. The sampling aim was also to achieve inclusion in functional capability, age, gender, social economic status and geographical location. This was achieved by using Inclusive design criteria such as those illustrated in Fig. 2. The field of inclusive design relates the capabilities of the population to the design of products by better characterising the user–product relationship. As such, its advantage is that of inclusion of functionality for varying capability into mainstream society. Inclusion refers to the quantitative relationship between the demand made by design features and the capability ranges of users who may be excluded from use of the product because of those features. Capability ranges were characterised by Functional analytical scales that can be numerically compared to a product function feature analysis [19]. The primary aim of inclusive design is to make mainstream products and services accessible to, and usable by, as many people as reasonably possible [19]. Designing inclusively ensures that the needs of user groups that are often under-represented in the design process are accounted for. This process not only can facilitate a quantification of the baseline inclusion, but also informs the design process and provides a metric to evaluate new design concepts [13]. The benefit of using an inclusive user group lies in the fact that a wide spectrum of capabilities is taken into consideration in the design phase. These capabilities can be auditory, visual, cognitive, or related to dexterity and mobility. Considering an inclusive sample of users with a variety of such capabilities results in designs that can provide a satisfactory user experience not only for the population with capability limitations, but also for a larger proportion of the population [13].

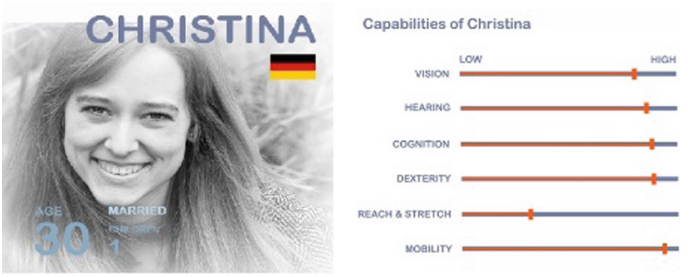


Fig. 2. Inclusive design persona profile showing functional scales [19]

As basis for sampling that took into account the journeys and personas of users, a source of information was the accurate National Survey data (Fig. 3). The *Intelligent Mobility, Traveller Needs and UK Capability Study (2015)*, an recent survey of UK journeys and transportation usage with 10,000 respondents in the UK, and enabled the formation of sampling categories of "persona-journeys". Hence potential users were classified by the journey types they undertake [20]. For example:

Traveller[1] Type	Description
Progressive Metropolites	Living in the heart of the city, typified by the technology-savvy young professional, with significant amounts of personal and business travel. Want to reduce their transport footprint.
Default Motorists	High mileage drivers, with a mix of those who enjoy driving and many for whom it is a functional choice.
Dependent Passengers	Dependent on others for their mobility needs, representing a mix of students, elderly, and those with impairments.
Urban Riders	City dwellers, who travel less frequently than the Progressive Metropolites, making use of public transport available to them.
Local Drivers	Mainly retirees or stay at home parents, making low mileage local journeys.

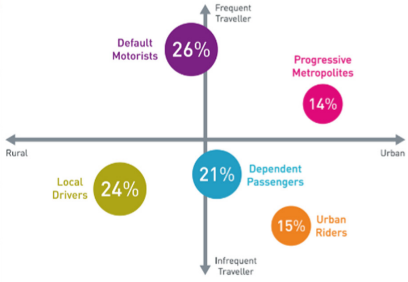


Fig. 3. User journey personas (top-level) identified by the IM UK traveller needs survey

'Dependent Passengers' covers 21% of UK population and 18% of journeys. It is a segment that is dependent on others to meet its mobility needs. This traveller type consists of a number of groups, such as young people (who typically get driven by their parents), elderly people, and travellers with impairments. They take a majority of their journeys as car passengers and the remainder is typically covered by either bus or by walking, with journeys split representatively between work and leisure.

Other aspects of the method; technology benchmarking, the Focus groups and the design workshop stage, are reported elsewhere. The relevant literature sampled issues concerning autonomous vehicle takeover including: Automotive research papers, trade journals and press releases.

2.1 Semi-structured Interviews

The outcomes of several stages of the qualitative research are reported: The **initial public interviews** with representative of groups identified by the Traveler needs survey. A total of eight semi-structured interviews were carried out aiming to obtain a cross-section of opinions and preferences from interviews based on the questioning regarding the appearance, function and use of future autonomous vehicles without any introduced biases of either explanation or suggestion regarding the HMI technology involved. Exposure to autonomous technology was based on two examples of current state of the art offerings represented by commercial Public Relations videos:

(1) Video 1 OEM concept autonomous interior video:
(2) Video 2 OEM video of fully autonomous car:

The transcribed interviews were analysed using thematic analysis software (Nvivo v.11). Two coders' themes were separately developed then compared and merged on the basis of content and statements by the respondents, the number and strength of the mentions, and the commonality with the other coder. The merging was made with the reduction of themes where repetitive or the reclassification of individual personal themes by merging them with a larger theme. New themes were added to re-classify idiosyncratic statements and some small themes were deleted if they were no longer significant in the new scheme.

2.2 Two Focus Groups

These were carried out addressing the issues of the understanding of autonomous vehicle takeover process and the HMI technology that may be involved; (reported in this volume *Ibid*). An assessment tool (*e.g.* a survey) identifying the participants' demographics and initial attitudes towards the design topic in order to recruiting representative users. This identified detailed user requirements for concept designs. The participants to the focus group were: technically minded individuals, in one case, and: inclusive group with a range of ages and gender to represent the desired user group for the design. A series of focus groups with different participant sets can provide a wider set of outputs, which were then triangulated into the interpretation process. The focus groups were video recorded and transcribed, and analysis of the script and video recordings is given in (*Ibid*) [21]. Facilitation had a pre-agreed structure, which enabled dialogue between participants while audio-visual tools display the focus points of discussion, in this case road layout diagrams and toy cars. The facilitator's role was to trigger continued discussion without predisposing the participants to any particular philosophy and which were kept technology and issue neutral to HMI in relation to the primary questions of takeover.

2.3 The Interpreted Outcomes Analysis

The interpreted outcomes analysis took as input the interpreted findings of these previous stages of the research. Hence, and following Goodman-Deane et al. [16] the analysis was made using modified qualitative research approaches, such as that of Grounded

Theory. This yielded a thematic analysis with an accompanying interpretation by the researchers. Importantly, the analysis requires a continued process of coding during the theme development. Themes were recoded and monitored iteratively. Multiple coders, without interaction, compared their coding's and converged onto a final set. Themes from initial sources were merged using a triangulated approach to generate a definitive interpretation and representative examples of the theme, and a rationale for it. A descriptive analysis was undertaken with the main themes described as a hierarchical list (see, for example in Fig. 4). A number of top level item themes subsumed sub-themes, as shown. For example, "HMI Issues" subsumed a number of themes, including "Algorithm expectations" and "Distractions". A detailed interpretation was then made in text format, breaking down each theme following an overview of the entire theme.

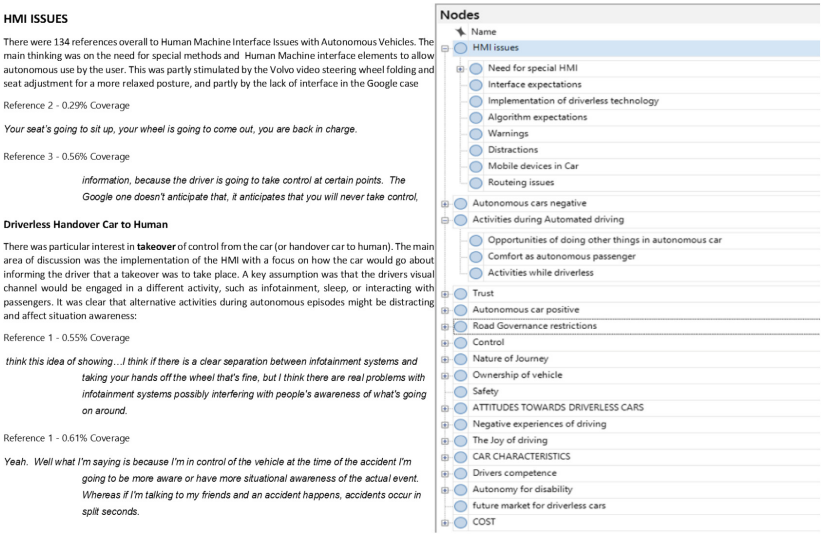


Fig. 4. (Left) Example of the interpreted analysis the top-level thematic nodes organized by number of references in the 8 sources. On the right is the themes resulting from the various stages of recoding, including hierarchies where they were apparent.

2.4 The Concepts and Requirements Insights

This table summarised the interpreted findings. The interpretation of the final themes forms a large corpus of text even with a second stage of recoding. One purpose of the interpreted analysis was to form identified micro-concepts or insights directly related to the issues identified and used as input to the themes. These were, by their nature repetitive and convergent. However, an alternative representation was used that allowed the comparison of the numerous occurrences in the interpretation and that was also useful as input to the Design process. The representation used was the engineering design requirements table, modified for social science requirements and used as a representation that captured the richness of the interpreted findings. It had the advantage of being a compressed format for the outcomes of the triangulated interpretation stage, especially

but allows human centred design requirements to be generated. At a minimum this gave each captured entity a unique identifier for later referencing; a name, a description, and a rationale. In principal, this allowed inclusive requirements to be generated for input into the design stages and final concepts. A convenient representation of this sort can be a requirements table that enumerates each idea-concept from the interpreted thematic analysis, gives it a description, a rationale for function and indication of its application in a design, including whether it is mandatory or optional. Such a table was constructed for the analysis from the two levels of interpreted findings of the interviews, focus groups and literature reviews. An initial table with all insights included a large number of insights, including repetitions and hierarchically related constructs, which were removed to reduce the total number.

Output insights from the previous process were used both as input in to the HMI concept development stages and also formed a stand-alone set for wider interpretation. For this purpose, insights were recoded into graphic card format and a complete set recorded for further analysis. A sub-set of these were generated, ignoring repetition or duplication of theme insight and context and this smaller set was used in the design workshops. This set is exemplified here, as a short example list. The full list contains 96 items and is the basis for further work (Table 1).

Table 1. Examples of insights, the result of triangulated interpreted qualitative data from interviews, surveys and focus groups.

Identifier	Decription
INT1: multimodality	In addition to a visual, there should be a series of increasingly strident signals as the takeover point approaches
INT14: response to inaction from driver	Car adopts safest strategy
INT8: system capability	The system should simple and easy to understand
FC50: countdown to takeover	Seen as similar to Satnav countdown instructions to a junction
INT12: warning by changing ambient interior configuration	Advantages of interior ambience were clear
FC6: integration of autonomy with traffic	Uncertainty how autonomous cars can perform alongside conventionally driven vehicles
ID29: time recovery	Facilities for activities during autonomous driving
ID33: impairment compensation	Provide inclusive design
INT17: regain control	Driver should be able to take control of car

3 Discussion and Conclusions

A number of insights were distilled from the qualitative data collected from a series of interviews and focus groups, These were triangulated and interpreted in the light of a number of sources of data, including technology benchmarking and recent survey

findings on attitudes to a social and economic issues associated with the introduction of autonomous vehicles technology at Level 3 and also Level 4.

Some manufacturers have autonomous cars on the roads effectively already functioning at SAE Automation level 3 [4] where the vehicle takes control of both steering and acceleration in specific circumstances but the driver must monitor the vehicles progress, ready to intervene to takeover control. Considerable media attention has been directed at accidents and incidents involving these Level 3 vehicles as a response based on public safety. Level 5 vehicles also represented in test form on UK roads and these require on intervention by the driver as they can handle the entire dynamic driving task [4].

In their thematic analysis of previous work and their own survey data, Cavoli et al. [2] draw wide ranging conclusions. However, they also identify a clear pathway for concern based on their extensive thematic review on Social and behavioural questions associated with automated vehicles that concerns the design of autonomous takeover technology and its HMI (Fig. 5). In particular, it would appear that the general perception is that engagement with non-driving tasks at Level 3 will introduce potentially unsafe situations leading to traffic incidents. These then may affect the uptake and acceptance of the technology as public perception is affected.

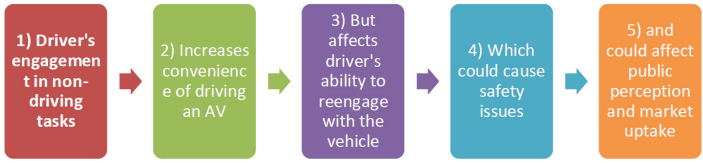


Fig. 5. The UCL transport institute report for the department of transport: potential safety risks in cases of partially or highly automated AVs

Despite the general finding that around 33% of sampled populations are generally positive towards AV technology, a considerable number of studies report 30%–50% that are either neutral or have no view regarding the technology. These participants are presumably related to the Mainstream and Price-drop" groupings of the Technology Acceptance Models [22] and their neutrality may arise from lack of knowledge of technology, or reluctance to commit unless further evidence from media and personal experience is forthcoming. This result was common to all the recent studies reviewed. On the other hand, a solid proportion of around 25% are reluctant to hand over control, would retain awareness of the road and see the AV driving environment as potentially dangerous and stressful. Again, it is likely that these may represent the "Last to Adopt" and "Technology Avoiders" groups of such models, who resist technology offerings. Interestingly, the focus group findings reported here (e.g. FC6: Integration of Autonomy with Traffic; and INT17: Regain Control) suggest that these negative attitudes can be associated with technology sophisticated urban and suburban dwellers who appreciate the potential failings of engineering [7, 23], rather than those who rely on blind trust in technology.

In addition to this, a number of conclusions made in the Cavoli study indicate that this field is far from understood. For example, they state that it will be necessary to further understand older people's perception of AVs and their potential reluctance to use the technology and also, importantly, that further research is needed to assess non-drivers' perception and expectations of AVs, and why a large percentage of females seem less interested in AVs compared to men. Nevertheless, they also found that 42% would currently force their car into a traffic stream, as a measure of driving sociability. Higher technology respondents showed more openness to technology and were more open to AVs. The least open to AV's were found to be more sociable drivers with low technology understanding but high levels of optimism, indication a form of "blind-faith" acceptance.

During the THINKMOBILITY survey there was a reported ~ 6% improvement of attitude as a result of social exposure to AV material suggesting that attitudes are changing rapidly in the UK and internationally. The Venturer Project survey [8] found that, in reviewing considerable body of previous literature, that variability in responses may have been due to response biases in sampling, particularly for surveys. They concluded that if we are to understand the nature of public opinion in this area surveys were not sufficient. They propose that further collection of specific qualitative data will be needed based upon identified populations responding to more detailed usage and technology cases. The present authors would concur, with the further specification of the importance on inclusive populations in sampled groups; catering for the wider non-mainstream population, including the older drivers and those whose capabilities may not be well matched to driving tasks. In addition, with reference to widening the usage of autonomous vehicles, the greater proportions of these groups are likely to be found among current non-drivers.

Acknowledgements. This work was funded by EPSRC and Jaguar-Land Rover, as part of the project Human Interaction: Designing Autonomy in Vehicles (HI:DAVe), Project Grant Number: EP/N011899/1.

References

1. Clark, B., Parkhurst, G., Ricci, M.: Understanding the Socioeconomic Adoption Scenarios for Autonomous Vehicles: A Literature Review. UWE, Bristol (2016)
2. Cavoli, C., et al.: Social and Behavioural Questions Associated with Automated Vehicles: A Literature Review. Department for Transport, London (2017)
3. Begg, D.: A 2050 vision for London: What are the implications of driverless transport? http://www.transporttimes.co.uk/Admin/uploads/64165-Transport-TimesA-2050-Vision-for-London_AWWEB-READY.pdf (2014)
4. SAE: http://cyberlaw.stanford.edu/files/blogimages/LevelsofDrivingAutomation.pdf. Accessed 27 Feb 2017
5. Schoettle, B., Sivak, M.: A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia (2014)
6. Schoettle, B., Sivak, M.: Motorists' Preferences for Different Levels of Vehicle Automation (2016)
7. Sommer, K.: Continental Mobility Study 2013 (2013)

8. Venturer: www.venturer-cars.com/
9. UK Autodrive: www.ukautodrive.com/
10. Kyriakidis, M., Happee, R., De Winter, J.C.F.: Public opinion on automated driving: results of an international questionnaire among 5000 respondents. Transp. Res. Part F Traffic Psychol. Behav. **32**, 127–140 (2015)
11. Tennant, C., Howard, S., Franks, B., Stare, S., Bauer, M.W.: Autonomous Vehicles - Negotiating a Place on the Road: A Study on How Drivers Feel about Interacting with Autonomous Vehicles on the Road. London School of Economics and Political Science, Goodyear Europe, Middle East and Africa (EMEA) (2017)
12. Geels, F.: The dynamics of transitions in socio-technical systems: a multi-level analysis of the transition pathway from horse-drawn carriages to automobiles (1860–1930). Technol. Anal. Strat. Manag. **17**(4), 445–476 (2005)
13. Langdon, P., Thimbleby, H.: Inclusion and interaction: designing interaction for inclusive populations. In: Editorial for Special Edition of Interacting with Computers, vol. 22, pp. 439–448, Elsevier (2010)
14. Politis, I., Brewster, S., Pollick, F.: Language-based multimodal displays for the handover of control in autonomous cars. In: Automotive UI 2015, pp. 3–10. ACM Press (2015)
15. Flick, U.: An Introduction to Qualitative Research - Fourth Edition. SAGE (2009)
16. Goodman-Deane, J., Langdon, P., Clarkson, J.: Key influences on the user-centred design process. J. Eng. Des. **21**(2–3), 345–373 (2010)
17. Breakwell, G.M., Hammond, S., Fife-Schaw, C.: Research Methods in Psychology, 2nd edn. SAGE, London (2000)
18. Nvivo: http://www.qsrinternational.com/product. QSR International (2017)
19. Keates, S., Clarkson, P.J.: Countering Design Exclusion – An Introduction to Inclusive Design. Springer, London (2004)
20. IMTNS: IM Travellers needs Survey (2015)
21. Politis, I., Langdon, P., Bradley, M., Skrypchuk, L., Mouzakitis, A., Clarkson. J.: Designing autonomy in cars: a survey and two focus groups on driving habits of an inclusive user group, and group attitudes towards autonomous cars (ibid.) (2017)
22. Legris, P., Ingham, J., Collerette, P.: Why do people use information technology? A critical review of the technology acceptance model. Inf. Manag. **40**(3), 191–204 (2003)
23. Kunur, M., Langdon, P., Bradley, M., Bichard, J.-A., Glazer, E., Doran, F., Clarkson, J., Loeillet, J.: Creating inclusive HMI concepts for future cars using visual scenario storyboards through design ethnography. In: International Conference on Universal Access in Human-Computer Interaction, pp. 139–149. Springer (2015)
24. https://ts.catapult.org.uk/current-projects/traveller-needs-uk-capability-study/
25. Missel, J.: Ipsos MORI Loyalty Automotive Survey (2014). http://www.ipsosmori.com/researchpublications/researcharchive/3427/Only-18-per-cent-of-Britons-believe-driverless-cars-to-bean-important-development-for-the-car-industry-to-focuson.aspx?utm_campaign=cmp_325684&utm_source=getanewsletter

Manual Takeover and Handover of a Simulated Fully Autonomous Vehicle Within Urban and Extra-Urban Settings

Phillip L. Morgan[1,2(✉)], Chris Alford[1,2], Craig Williams[1,3], Graham Parkhurst[1,4], and Tony Pipe[1,5]

[1] University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, United Kingdom
{phil.morgan,chis.alford,craig6.williams, graham.parkhurst,tony.pipe}@uwe.ac.uk

[2] Psychological Sciences Research Group, University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, United Kingdom

[3] Centre for Health and Clinical Research, University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, United Kingdom

[4] Centre for Transport and Society, University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, United Kingdom

[5] Bristol Robotics Laboratory, University of the West of England – Bristol, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, United Kingdom

Abstract. Relatively little is known about human behavior and performance when retaking control of highly autonomous vehicles (AVs) at different speeds and under varied driving conditions. Past research has tended to focus on long periods of high-speed extra-urban autonomous driving before participants switch to manual mode. The current study investigated time to 'takeover' (reengage with vehicle controls) and 'handover' (regain baseline level of driving) when switching frequently between automated and manual modes within extra-urban and slower urban settings. Thirty-one drivers completed scenarios at different speeds (20–30-mph/urban; 40–50-mph/extra urban). Dependent measures included speed and lateral lane position. Takeover time was consistent in 30–50-mph conditions (~2-seconds) but significantly slower at 20-mph. Some other measures took 15–20-seconds to match baseline in 20–30-mph conditions and participants tended to drive slower than recommended speed limits after takeover. Baseline performance was not achieved at 50-mph. Implications for handover system design and AV insurance frameworks are discussed.

Keywords: Automated vehicle · Takeover · Handover · Urban settings

1 Introduction

Autonomous vehicle (AV) technology is developing at an accelerating pace, with highly and fully autonomous vehicles likely to be a conventional form of motor transport in the near future [1, 2], perhaps overtaking non-autonomous vehicles as the leading

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_73

automotive engineering solution. There have been many studies on human behavior and performance evaluating one or more automated vehicle functions (such as adaptive cruise control: e.g., [3–5]). However, less is known about how humans cope with taking over control of highly and fully AVs; especially under different speeds and driving conditions (see [6] for a review). In line with recent research [7], the current driving simulator study set out to investigate 'takeover' (time taken to reengage with vehicle controls) and 'handover' (time taken to regain a baseline/normal level of driving) behaviour and performance when switching frequently between automated and manual driving modes and not in response to critical events. The study employed different speed conditions ranging from 20 to 50-mph within urban and extra-urban driving scenarios. The study forms part of a major trial conducted for an Innovate UK funded project – Venturer - Driverless Cars on UK Roads (2015–2018) – with a parallel study completed (not reported here) involving a highly autonomous road vehicle.

The potential problems associated with safe and efficient 'takeover' and 'handover' arise as a feature of developing AVs which are highly but not fully autonomous [1]. These are referred to as level 3 and 4 in frameworks of vehicle autonomy [8]. The current study focusses on level 3 autonomy defined by SAE International [9] as *'conditional autonomy: the driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene'*. With such systems, the AV controls key vehicle functions such as steering, acceleration, and deceleration as well as monitoring the driving environment via sensors and/or a radar(s). The human is an agent within the system that may need to takeover the dynamic driving task from the automated system in the event of a handover request (e.g., due to loss of signal). Level 3 vehicle autonomy involves hands and feet of controls and eyes of the road but should not promote 'mind off ' the road (level 5 autonomy), to minimize possible compromises of safe and efficient takeover and handover behavior and performance even with handover support systems.

Research involving highly automated driving (HAD) has begun to soar over the past 7–8-years [10–13]. HAD involves multiple automated systems whereas fully automated driving (FAD) involves automation of all vehicle systems that would otherwise be controlled by a human in a manual vehicle [12]. Much of the research on HAD and FAD has focused on the handover phase between automated and manual driving with studies testing conditions that may impact performance, such as: critical incidents [14], traffic density [13], distractions [10] and fatigue [15].

Many studies have adopted paradigms involving long driving scenarios of mainly highway driving [10, 14, 16], with fewer adopting shorter scenarios [7, 15]. Many have tested quite experienced high mileage drivers [16]; sometimes with experience of driving simulation [17, 18]. This can be beneficial for reducing possible acclimatization effects of engaging with simulator controls and scenarios although there can be a risk of expertise effects affecting some findings. The current study employs short driving scenarios involving urban and extra-urban non-highway driving and tests participants with varying degrees of driving experience and little or no driving simulator experience.

An early study by Merat and Jamson [19] examined the effects of lateral and longitudinal lane position automation on handover and takeover behaviour and performance

during 40-mph driving scenarios. Participants with driving simulation experience undertook an automated condition involving having to take back control of the simulator once an auditory signal occurred followed by a takeover lag in response to critical events. During handover, driving measures such as braking and anticipation time were slower in the automated condition. It was concluded that participants had experienced a loss of situation awareness and suggested that frequent updates during automated driving might help to combat these issues. However, the automated system dealt with some but not all critical incidents and this may have increased the expectation that manual intervention was not always needed. This could have affected decision and reaction times to brake and the efficiency of the response thereafter.

To further explore the loss of situation awareness hypothesis, Merat et al. [14] examined handover occurring at regular versus irregular intervals and initiated in response to time spent looking away from the road. Experienced high annual mileage drivers completed extra-urban scenarios at speeds of 70-mph. Gaze patterns towards the road were higher and more stable in a regular handover condition whereas drivers in an irregular condition were more susceptible to fixating on environmental distractors. Following takeover, it took drivers a considerable 35–40-seconds to stabilize lateral lane control irrespective of being in the regular or irregular condition. The authors stressed that most 'real life' handover situations are likely to be unpredictable and therefore future research is needed to determine how situation awareness can be regained and manual driving takeover achieved more efficiently.

Recent studies on autonomous driving have focused on factors that can impact safe and efficient handover behavior and performance. These include frequency of handover requests and performing secondary tasks [7, 20], traffic density and complex driving conditions [21], and factors such as age [22]. Gold et al. [21] tested 72 participants having to make an evasive maneuver soon after takeover in scenarios involving different degrees of traffic density. Higher traffic density led to longer takeover times and more near collisions than in lower traffic density conditions, with performance of a secondary task prior to the takeover request only affecting takeover quality and not time. Eriksson and Stanton [7] noted that many past studies have involved infrequent handover requests and many have operationalized these by having to respond to critical events soon after the handover request. They tested 26 participants on two three-lane highway scenarios driving at 70-mph. Participants performing a secondary reading task or monitored vehicles system during the automated period. They found evidence of longer individual takeover of control resumption times (e.g., up to 25.75-seconds) and greater individual variance than generally previously reported and stressed that we need to proceed with caution when trying to generalize some findings.

More research is needed on behavior and performance when retaking control and trying to resume safe and efficient manual driving of highly autonomous vehicles under different driving conditions (e.g., various speeds, different environments). The current study investigated takeover and handover driving behavior when switching frequently between automated and manual modes within low speed urban inner city settings (20- and 30-mph) and relatively higher speed extra-urban outer city settings (40- and 50-mph). It was predicted that takeover time and time taken to regain a baseline manual level on key driving measures such as lateral lane position would increase in line with autonomous

mode driving speed. It was also expected that takeover and handover times would be lower than in some studies [7, 14] employing higher speeds (e.g., 70-mph) and highway driving scenarios.

2 Method

Participants. Thirty-one people took part comprising mostly of University of the West of England (UWE) – Bristol and University of Bristol staff and students. This was adequate to detect a medium-large effect size (Cohen's $f = .25–.4$) with power of .8 [23]. Ages ranged from 18 to 69 years ($M = 41.0$, $SD = 13.9$, $3 > 60$-years), with 16 males and 15 females. Inclusion criteria included: full UK driving license (or non-UK driving license with permission to drive in the UK), normal-corrected vision and hearing, and fluency in English comprehension. Individuals who had experienced a major road traffic accident within 5-years were not permitted to take part.

Design and Materials. A factorial repeated measures design was employed. The main independent variable was recommended speed limit with four levels: 20-, 30-, 40-, and 50-mph. The 20- and 30-mph urban conditions involved inner-city built-up scenarios with moderate traffic density. The 40–50-ph extra-urban conditions involved outer-city less built-up scenarios; with lower traffic density. Each speed condition constituted a five-minutes 'segment' of the main handover trials (see below) and was repeated twice such that there were two handover trials at each speed limit. Handover point within a trial also varied, such that there were two handovers per segment and 16 handovers in total. This meant that handover requests were not too frequent to risk high predictability and elevated arousal levels, determined through pilot studies. Speed and handover point treatment combinations were counterbalanced using a Latin Square.

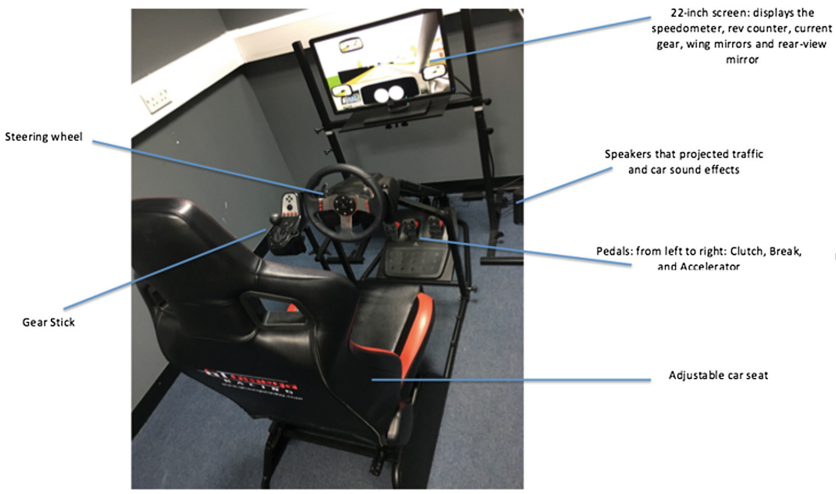


Fig. 1. Simulator and controls. *Note.* Gear shift not used and monitor was upgraded to 50″

A STISIM DriveTM Model M100 driving simulator (Fig. 1) was used to develop and run driving scenarios in fully automated and manual modes. Hardware consisted of a Logitech G27 steering wheel with horn and indicator buttons, GamePOD GT2 Stem, an upgraded 50" widescreen computer monitor, and pedals (clutch, accelerator, brake). The simulator was always in automatic transmission mode during manual driving.

Scenarios were programmed using STISIM Drive 3.1 scenario builder and Open Module Programming software. Handover (automated to manual driving) and handback (manual to automated) were signaled by an auditory alert. Phases were as follows:

- *Orientation (5-minutes)*: mixed urban (inner-city: built-up, traffic lights, junctions, moderate volume of pedestrians, moderate traffic density) and extra-urban (outer-city: less built-up, no traffic lights, no junctions, low volume of pedestrians, low traffic density) environments and recommended speed limits of 20-mph (urban) and 50-mph (extra-urban);
- *Initial Handover Practice (5-minutes):* mixed urban and extra urban environment and a speed limit of 30-mph with two handover-handback requests;
- *Manual Baseline (10-minutes):* 2.5-minutes for each speed limit;
- *Handover trials (40-minutes):* 2 × 5-minutes driving for each speed limit. Each trial including 3-minutes in fully automated driving mode and 2-minutes in manual driving mode (i.e., two handovers per trial).

Each handover (automated to manual) phase involved 60-seconds of manual driving. Many past autonomous driving simulator studies have adopted long driving scenarios upwards of 60-minutes of mainly automated highway driving [10, 14, 16] with times to regain baseline manual driving performance reaching 40–50-seconds. Thus, 60-seconds handover periods were deemed sufficient for the much shorter 5-minutes driving scenarios used within the current experiment.

Driving-related dependent measures included: longitudinal acceleration; lateral acceleration; standard deviation of lateral lane position (SDLP); average speed; and steering, gas, and brake pedal input count. These measures were compared between the baseline and experimental handover phases.

When the STISIM was in fully automated mode, participants proofread a health and safety document. This was positioned on a document stand approximately 45° to the left of STISIM screen.

Procedure. Participants were tested within the Automated Driving Zone of Bristol Robotics Laboratory on the UWE-Bristol Frenchay campus. Once adjustments were made for their comfort, the experimenter read aloud a set of instructions about the experiment and tasks. Participants completed an 'orientation' phase to familiarise themselves with the controls and driving scenarios, as well as to determine whether they were feeling the effects of simulator sickness. They were instructed to drive in the lane closet to the center line and to stay as close to the center line as possible during all driving scenarios. If happy to continue with the remaining trial phases, participants signed a consent form. They then completed 'initial handover practice', followed by 'manual baseline' and then by 'handover trials'. Participants received a full written and verbal debrief. The experiment took 120–125 min to complete, with breaks between each phase.

3 Results

All statistical tests reported are two-tailed with alpha levels of .05. Effect sizes were determined using Cohen's *f* [24] with ≤.1, ≥.25 and ≥.4 indicating small, medium and large effect sizes. Error bars in figures represent ±standard deviation.

First, it is important to establish the time to takeover simulator controls (steering and/or gas pedal or brake pedal input) following the onset of the auditory handover request. Driving controls were handed back to participants immediately at the onset of the auditory alert. Means displayed in Fig. 2 reveal that takeover time was close to 2-seconds in the 30–50-mph conditions and almost 2.5-seconds in the 20-mph condition.

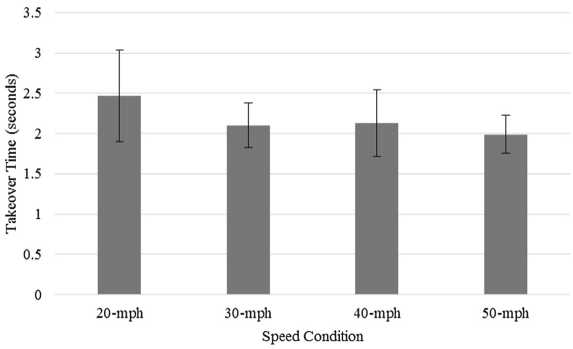


Fig. 2. Mean time to takeover driving controls following a handover request.

A repeated measures ANOVA revealed a significant main effect of speed condition on takeover time, $F(3, 90) = 13.47, p < .001, f = .67$. Bonferroni post-hoc tests revealed that takeover time was significantly higher in the 20-mph condition than the 30-, 40- and 50-mph conditions (ps $< .05$). Takeover time did not differ between the 30-, 40-, and 50-mph conditions (ps $> .05$).

In order to examine whether there were generic differences in handover manual driving performance compared with baseline manual driving, we next compared the entire period of manual driving following handover across all speed conditions. This involved 55- of the 60-seconds period of manual driving; excluding the last 5-seconds due to noise in the data. A 4 × 2 repeated measures MANOVA with speed condition as one independent variable (20-, 30-, 40-, and 50-mph) and trial phase (baseline, handover trials) as the other independent variable revealed a significant main effect of speed condition, $F(24, 255) = 125.70, p < .001, f = 3.44$. There was also a significant main effect of trial phase, $F(8, 23) = 55559.01, p < .001, f = 3.50$, and a significant interaction between both variables, $F(24, 155) = 24.56, p < .001, f = 1.52$. The interaction was explored using Bonferroni post-hoc tests and revealed the following:

– *Average Speed:* participants drove significantly slower during handover than in baseline in the 20-, 40-, and 50-mph conditions ($p < .001$) but there was no difference in the 30-mph condition ($p = .09$);

– *Longitudinal Acceleration:* participants did not accelerate as much during handover compared to baseline in the 30–50-mph conditions (*p*s < .001) but did accelerate more in the 20-mph condition during handover (*p* < .001);
– *SDLP:* technical issue in the 20-mph condition; significantly lower SDLP mean during handover than in baseline within the 30- and 40-mph conditions (*p*s < .01); and significantly higher during handover than in baseline within the 50-mph condition (*p* < .001);
– *Brake Pedal Input:* no difference between handover and baseline in the 20-mph condition; significantly greater during handover than in baseline within the 30–50-mph conditions (*p*s < .01, .05, and .001 respectively).

Next, we compared performance during handover versus baseline with more fine analysis; specifically at 5-seconds cumulative increments post-handover request.

Within the 20-mph condition, gas pedal input was similar in handover to baseline, although only after 10-seconds (all *p*s > .05 after this point). Brake pedal input was also similar in handover to baseline, although only after 15-seconds (all *p*s > .05 after this point). Whilst these findings indicate that baseline driving performance (arguably normal manual driving) was achieved on some measures within 5–15-seconds after a handover request, this was not the case for other measures, including: longitudinal acceleration (higher during handover), lateral acceleration (lower during handover), and steering input (lower during handover). Within the 20-mph condition, participants drove consistently and significantly slower during handover (Fig. 3). This represented a 1.68-mph reduction (first 0–5-seconds of handover, including takeover period) falling to a 0.97-mph difference between baseline and handover over the entire 55-seconds handover period. This may represent more cautious driving during handover.

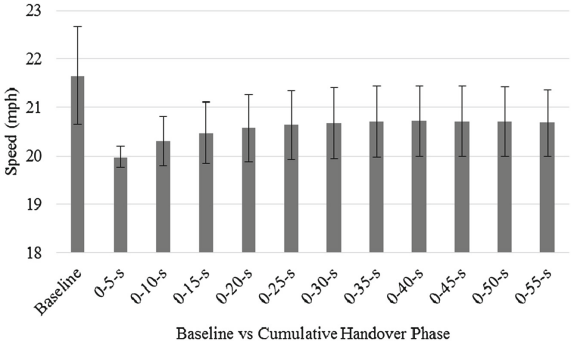


Fig. 3. Mean speed for baseline versus handover segments in the 20-mph condition.

Within the 30-mph condition, SDLP was similar between handover and baseline immediately after the handover request and for the 55-seconds handover period (*p*s > .05). Longitudinal acceleration was similar between handover and baseline after 35-seconds (*p*s > .05). Steering input was similar between handover and baseline immediately post-handover (*p*s > .05 until 25-seconds and *p*s > .05 after 45-seconds). These findings are partially comparable to the 20-mph condition with the addition of SDLP that did not differ

between handover and baseline. Other driving measures (e.g., lateral acceleration) did not match baseline for cumulative periods of more than 5–10-seconds. Participants drove consistently and significantly slower during handover than in baseline, representing a 1.64-mph difference (first 0–5-seconds of handover) falling to a 0.30-mph difference at 55-seconds. This represents more cautious driving during handover.

Within the 40-mph condition, SDLP was similar between handover and baseline immediately after the handover request and for the 55-seconds handover period ($ps > .05$). Brake pedal input was also similar in handover vs baseline immediately post-handover ($ps > .05$ after this point until 20-seconds with some significant differences after). In general, these findings represent a lower degree of similarity between handover and baseline findings compared with the 20–30-mph conditions. Other driving measures did not match baseline levels at handover for cumulative periods of more than 5–10-seconds. As in the 20–30-mph conditions, participants drove consistently and significantly slower during handover than in baseline, representing a marked 2.76-mph reduction (first 0–5-seconds of handover) falling to a 1.48-mph difference at 55-seconds. These findings again seem to represent more cautious driving during handover.

Within the 50-mph condition, none of the driving measures matched baseline at handover for cumulative periods of more than 5–10-seconds apart from steering input after 35-seconds. SDLP was significantly higher during handover. Longitudinal acceleration, gas and brake pedal inputs were significantly lower during handover; fitting of more cautious driving.

As in 20–40-mph conditions, participants drove consistently and significantly slower during handover than in baseline (Fig 4), representing a 2.30-mph reduction (first 0–5-seconds of handover period) and reducing further to up to 4.08-mph slower. Again, these findings seem to represent more cautious driving during handover, although variance was much higher after 0–10-seconds than in the other speed conditions.

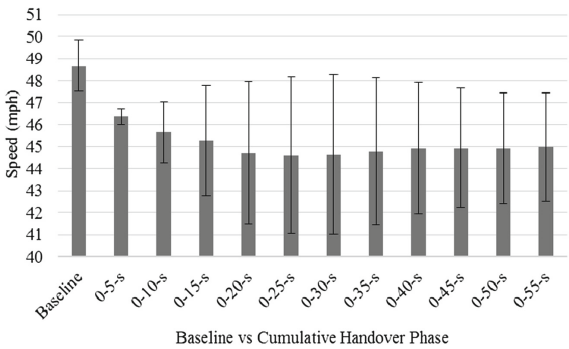


Fig. 4. Mean speed for baseline versus handover segments in the 50-mph condition

4 Discussion

The current study investigated performance when taking over control of a simulated highly autonomous vehicle following handover requests to non-critical events under

different driving speeds (20–50-mph) and within different environments (urban inner city and extra urban outer city). Some of the findings have been validated by another study forming part of the same Venturer project involving a road-based highly autonomous vehicle; to be reported soon.

Within the current simulation study, it was predicted that takeover and time taken to regain baseline performance on key driving measures during handover (e.g., SDLP) would be lowest in the 20- and 30-mph urban conditions and would increase in higher speed extra urban conditions. It was also expected that these times would be lower than reported in some past studies that have tended to employ higher speed conditions (e.g., 70-mph or higher) and mainly multi-lane highway driving scenarios [7, 14].

Takeover time ranged from just below 2-seconds to almost 2.5-seconds and was lower in extra urban outer city scenarios than urban inner city scenarios. However, it was only significantly higher in the 50-mph extra urban speed condition whereas the 20–40-mph speed conditions did not differ. These means are lower than median values recently reported by Eriksson and Stanton: 4.57-seconds with no-secondary task and 6.06-seconds with a secondary task [7]. The takeover time averages within the current study are more in line with some other previous studies [13, 25, 26] that have tended to involve much higher driving speeds.

The longest takeover mean time for an individual participant within the current study was 4.77-seconds ($SD = 1.74$-seconds) in the 20-mph condition and the fastest takeover time was 1.32-seconds ($SD = 0.22$-seconds) in the 40-mph condition. Overall, takeover times in the current study were lower than in the Eriksson and Stanton [7] study and this is most likely due to the urban and low speed (40–50-mph) extra-urban speed conditions as well as possibly the frequency of handover requests during the relatively short driving scenarios employed. Handover requests occurred twice during each 5-minutes driving scenario with total automated driving time set at 3-minutes per driving scenario. Thus, participants within the current study might have been even more poised – despite completing the proofreading secondary task – to retake vehicle controls. In addition, whilst the current study tested takeover and handover in response to non-critical events as in the Eriksson and Stanton [7] study, participants did not need to make any maneuvers (e.g., leaving the operational domain) and instead only had to try and maintain the recommended speed limit and stay in the same lane during the handover period. This may also have accounted for the relatively shorter takeover times.

Some post-takeover driving measures within the current study resembled baseline manual driving within the 20–40-mph speed conditions. In the 30- and 40-mph speed conditions, SDLP was the same as baseline immediately after the handover request and was stable for the entire 55-seconds handover manual driving period. However, SDLP was higher (i.e., greater deviation) than in baseline within the 50-mph extra urban condition. (A technical problem meant that an adequate set of SDLP data was not available in the 20-mph condition although trends in available data suggest similar findings to the 30- and 40-mph conditions.) These findings are important as they suggest that maintaining a stable lane position is not compromised following takeover of vehicle controls after a period of automated driving within simulated 20–40-mph speed conditions. The SDLP findings from the 50-mph extra-urban condition might suggest less safe driving, although future research is needed with simulated road and traffic

conditions with potential hazards (e.g., vehicles overtaking close by) to further investigate whether this raises a cause for concern. This would be needed before roadworthy AVs are tested in outer-city settings at speeds higher than 40-mph.

In the urban driving scenarios, some other driving measures stabilized to baseline levels in less than 20-seconds. For example, in the 20-mph condition, gas pedal input was similar to baseline 10-seconds after the handover request and brake pedal input was similar 15-seconds post-handover request. Whilst gas and brake pedal inputs did not match baseline during the post-handover request period in the 30-mph condition (participants tended to accelerate more and brake less), steering input did match baseline immediately after the handover request for up to 25-seconds post-handover. There was little resemblance between baseline & handover manual driving within the highest 50-mph speed condition on all other driving measures reported here as well as SDLP (e.g., lateral/sideway acceleration was also higher during handover than in baseline). Again, this raises a cause for concern for AVs with handover capabilities travelling at speeds higher than 40-mph within or on the outskirts of urban settings.

Another interesting and consistent finding was that average speed following a handover request was lower than in comparative manual baseline conditions and mostly lower than the recommended speed limit across conditions. In fact, average speed was up to 1.68-mph slower than baseline following a handover from autonomous to manual driving modes in the 20-mph condition and over 4-seconds slower in the 50-mph condition. This suggests more cautious manual driving following a handover request than during baseline manual driving that persisted for the entire 55-seconds post-handover request period assessed in all speed conditions. However, given the variability in the time taken to regain manual baseline control levels across a range of parameters, a phased handover of controls is recommended.

There are limitations to the current study. First, all participants performed a straightforward proofreading secondary task when the simulator was in autonomous mode to ensure eyes were off the road prior to handover. The scenarios we used were relatively short in comparison to most other AV studies, and it would be interesting to examine whether participants would be poised to takeover with eyes on the road ahead. Second, participants may have learned that handover requests did not occur in response to critical events which might have affected speed to takeover controls. This also needs testing within future studies. Third, any consequences of e.g., slow takeover and/or poor driving performance after taking over controls within a driving simulator are minimal. There is no threat to the safety of the participant, other people, or risk of damaging a vehicle(s). Thus, parallel testing using a road-based autonomous vehicle is crucial if findings are to inform e.g., AV handover system design.

5 Implications

The findings of the current study and those from another of our studies using a road-based highly autonomous vehicle suggest that designers of autonomous vehicle technology with handback functionality need to proceed with caution and consider human performance under multiple driving conditions and scenarios in order to plot accurate

takeover and handover time safety curves. This recommendation is in line with suggestions by other researchers investigating takeover and handover during automated driving [7, 14]. Our scenarios were designed to mimic urban automated driving conditions and frequency of handover requests representative of early generation level 3 autonomous road vehicles for use in urban settings whereas the higher speed highway scenarios used in studies such as [7] are more representative of level 3 autonomous road vehicles for use in extra urban settings. Taken together, we posit that our findings are complimentary and important in terms of informing the design and deployment of AVs with handover capabilities deployed in urban and extra-urban settings, and that environment specific handover control specifications are required.

Through our current and planned driving simulator and road-based autonomous experiments, we are deriving and validating design principles for safe handover support systems within highly autonomous vehicles. Our findings are also informing the development of autonomous vehicle insurance standards, principles, and policies with collaborators. This includes the development of parameters for liability frameworks within level 3 autonomous vehicles designed for use within and around urban environments.

Acknowledgements. The reported research forms part of an Innovate UK research project - VENTURER: Introducing driverless cars to UK roads (2015–2018). See http://www.venturer-cars.com/. We thank Carolyn Mitchell, Rebecca Tommey, and Peter Blackley (all at Atkins UK) for reading an earlier draft of this paper, and Dr. Alex Lenz and Jason Welsby (both at Bristol Robotics Laboratory) for assisting with programming the autonomous elements of the driving scenarios.

References

1. Department for Transport: The pathway to driverless cars: a detailed review of regulations for automated vehicle technologies (2015). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/401562/pathway-driverless-cars-summary.pdf
2. Walker, G.H., Stanton, N.A., Young, P.M.: Human Factors in Automotive Engineering and Technology. Ashgate, Surrey (2015)
3. Ma, R.: Kaber.: Situation awareness and workload in driving while using adaptive cruise control and a cell phone. J. Ind. Ergon. **35**(10), 939–953 (2005)
4. Stanton, N.A., Young, M.S.: Developing a psychological model of the driver. In: Eleventh Seminar. Behavioural Research in Road Safety (2002)
5. Stanton, N.S., Young, M.S.: Driver behavior with adaptive cruise control. Ergonomics **48**, 1294–1313 (2005)
6. Morgan, P.L., Alford, C., Parkhurst, G.: Handover issues in autonomous driving: a literature review. Innovate UK Venturer Project Report D9, University West of the England, Bristol, UK (2016)
7. Eriksson, A., Stanton, N.S.: Take-over time in highly automated vehicles: Non-critical transitions to and from control. Hum. Factors **59**(4), 689–705 (2017)
8. SAE International: U.S. Department of transportation's new policy on automated vehicles adopts SAE International's levels of automation for defining driving automation in on-road motor vehicles (2016). https://www.sae.org/news/3544
9. SAE International: Automated driving: levels of driving automation defined in new SAE international standard J3016 (2014). https://www.sae.org/misc/pdfs/automated_driving.pdf

10. Jamson, A.H., Merat, N., Carsten, O.M.J., Lai, F.C.H.: Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. Transp. Res C **30**, 116–125 (2013)
11. Merat, N., Jamson, H.: How do drivers behave in a highly automated car? In: Proceedings of the Fifth International Driving Symposium on Human Factor in Driver Assessment, Training and Vehicle Design, pp. 514–521 (2009)
12. Merat, N., Waard, D.: Human factors implications of vehicle automation: current understanding and future directions. Transp. Res. F **27**(Part B), 192–195 (2014)
13. Radlmayr, J., Gold, C., Lorenz, L., Farid, M., Bengler, K.: How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. In: Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting, pp. 2063–2067 (2014)
14. Merat, N., Jamson, A.H., Lai, F.C., Daly, M.R., Carsten, O.M.J.: Transition to manual: driver behaviour when resuming control from a highly automated vehicle. Transp. Res. F. **27**(Part B), 274–282 (2014)
15. Neubauer, C., Matthews, G., Langheim, L., Saxby, D.: Fatigue and voluntary utilization of automation in simulated driving. Hum. Factors **54**(5), 734–746 (2012)
16. Merat, N., Jamson, H., Lai, F., Carsten, O.: Highly automated driving, secondary task performance and driver state. Hum. Factors **54**, 762–771 (2012)
17. Gold, C., Damböck, D., Lorenz, L., Bengler, K.: "Take over!" How long does it take to get the driver back into the loop? In: Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting (2013)
18. Gold, C., Damböck, D., Bengler, K., Lorenz, L.: Partially automated driving as a fallback level of high automation. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren. 28(29.11) (2013)
19. Merat, N., Jamson, A.H.: Is drivers' situation awareness influenced by a fully automated driving scenario? In: de Waard, D., Godthelp, J., Kooi, F.L., Brookhuis, K.A. (eds.) Human Factors, Security and Safety, pp. 1–11. Shaker, Maastricht (2009)
20. Naujoks, F., Purucker, C., Neukum, A.: Secondary task engagement and vehicle automation—comparing the effects of different automation levels in an on-road experiment. Transp. Res. F **38**, 67–82 (2016)
21. Gold, C., Körber, M., Lechner, D., Bengler, K.: Taking over control from highly automated vehicles in complex traffic situations the role of traffic density. J. Hum. Factors Ergon. Soc. **58**(4), 642–652 (2016)
22. Körber, M., Gold, C., Lechner, D., Bengler, K.: The influence of age on the take-over of vehicle control in highly automated driving. Transp. Res. F **39**, 19–32 (2016)
23. Faul, F., Erdfelder, E., Lang, A.G., Buchner, A.: G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods **39**(2), 175–191 (2007)
24. Cohen, J.: Statistical Power Analysis for the Behavioral Sciences. Lawrence Earlbaum Associates, Hillsdale (1998)
25. Louw, T., Merat, N., Jamson, H.: Engaging with highly automated driving: to be or not to be in the loop. In: 8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design (2015)
26. Kerschbaum, P., Lorenz, L., Bengler, K.: A transforming steering wheel for highly automated cars. In: Intelligent Vehicles Symposium (IV), pp. 1287–1292. IEEE (2015)

Road and Rail: Infrastructure and Road Safety

Driving Performance, Adaptation, and Cognitive Workload Costs of Logo Panel Detection as Mediated by Driver Age

Mei Ying Lau and David Kaber(✉)

Department of Industrial Engineering, North Carolina State University,
400 Daniels Hall, 111 Lampe Drive, Raleigh, NC 27695, USA
{mylau,dbkaber}@ncsu.edu

Abstract. This study investigated the effect of on-road visual sign usage on young (≤22 years), middle-aged (23–64 years), and elderly (65+ years) driver performance, adaptive vehicle control behavior and mental workload. Eight simulated highway driving scenarios (the primary task) were presented along with an embedded secondary task of familiar business logo sign identification. Driver workload was manipulated by varying logo panel counts on service signs, including 6- and 9-panels. Results revealed comparable driving performance among the 9- and 6-logo conditions but significant differences in sign identification among age groups. Elderly drivers produced significantly lower sign identification accuracy than other age groups (at 54–57%). Both elderly and young drivers demonstrated significantly higher cognitive workload (reduced blink duration) in response to secondary task demands as compared to middle-aged drivers. The workload during sign identification led to significantly greater elderly adaptation behaviors (higher speed reductions) and performance degradations (greater lane deviations).

Keywords: Driving performance · Driver age · Adaptation · Cognitive workload · Logo signage · Blink duration · Driving simulator

1 Introduction

1.1 Aging and Driving

Driver age has been previously associated with decrements in driving performance [1]. While age itself does not lead to accidents, age-related declines in attentional, perceptual, cognitive and psychomotor abilities have been found to significantly contribute to unsafe driving incidents [2]. An analysis of crash data from 1995–2010 revealed crash involvement rates for young and older drivers to be consistently above average [3]. High rates of roadway regulatory infractions among young drivers, including collisions and speeding, can be largely accounted for by inexperience [4]. Such issues can, however, be addressed through training, safety education, and increased driving experience. As young drivers gain driving experience over time, crash rates sharply decrease [3]. On the other hand, elderly drivers are particularly at risk due to age-related declines in

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_74

information processing and motor capabilities, which cannot simply be resolved through training and education [2]. Unfortunately, many contemporary highway designs have been based on physiological and cognitive abilities of younger and middle-aged drivers and such designs may not be suitable for elderly drivers.

Due to potential age-related declines in various abilities, elderly drivers may not only be more prone to crash involvement but also be more vulnerable to accidents compared to younger driver. Crash fatality data for 2013 revealed that while crash fatalities commonly decreased among all age groups under 55 years with additional driving experience, fatalities actually increased in 55+ year old community [5]. In addition, odds of motor vehicle crash-related injuries, such as serious head and extremity injuries, has also been found to increase with age [6]. Related to these statistics, elderly drivers are a continuously increasing percentage of the US driving population.

With respect to driver use of on-road signage, two previous studies [7, 8] found elderly drivers to require significantly more time, closer proximity and more eye movements (fixations) to acquire targets than younger counterparts. Even with extra response time, older adults are less accurate in target identification. Ho et al. [8] reported that in driving situations involving search and identification of targets in visually complex scenes within a fleeting time window, older adults are more likely to misidentify or miss a sign. Such conditions can also aggravate elderly driver's driving behavior as the driving task poses greater mental demand for elderly drivers than younger counterparts [9]. On this basis, there is a need to consider elderly driver capabilities in roadway/highway design in order to support effective target detection along with vehicle control to prevent crashes and injuries.

1.2 Motivation

The present study focused on "costs" of driver external distraction in terms of performance decrements in encounters with visually-complex traffic signs requiring target search during normal freeway scenarios. Specific service ("blue") signs include multiple business logo panels (often referred to as "logo signs"). Prior research on driver use of service signs [7, 10–12] has focused on performance variations under different roadway conditions but none examined cognitive workload and performance decrements in driver successful identification of logo targets. Based on this review, the main research question for our study was whether workload or performance costs of sign use would vary among elderly, middle-aged and younger drivers. Driver ability to maintain vehicle control while correctly identifying targets is essential to successfully reaching a destination and ensuring safety. Drastic changes in speed or lane position are especially problematic in high-paced freeway environments with respect to crash potential. The "cost" of successful freeway sign target detection in terms of simultaneous driver vehicle control and cognitive workload may be greater for elderly due to effects of aging. In terms of highway design, it is critical that drivers of all ages are able to successfully identify targets without excessive cognitive demands and loss of vehicle control.

1.3 Hypotheses

We hypothesized that greater density of external visual information in a driving environment, specifically the number of logo panels on freeway service signs, would lead to an increase in driver mental workload (Hypothesis (H)1), adaptive vehicle control behavior (H2) as well as performance degradations (H3) across all age groups. However, based on the previous crash rate statistics, it was also expected that these effects would be most pronounced for elderly and possibly young age drivers in using specific service signs. Therefore, we also hypothesized that in successful sign target detection, elderly drivers would experience higher mental workload (H4) and exhibit greater adaptation behavior (H5) as well as driving performance degradations (H6) as compared to middle-aged drivers but comparable responses to young drivers.

2 Methodology

2.1 Prior Study

A prior freeway driving simulation study [13] was conducted with a focus on driver distraction and performance effects of roadway sign familiarity, logo format and panel count (6 vs. 9). The present study used data collected during that investigation. In Zahabi et al. [13] study, participants experienced eight test trials with half familiar and half unfamiliar target detection conditions (in random order). Logos were classified as familiar if they represented a common US chain business and unfamiliar if they presented an independent/local business. The familiar food target ("Dunkin' Donuts") remained consistent across trials; however, unfamiliar logos varied and were presented only once during the experiment. Driver responses to roadway signage were characterized and analyzed in terms of signal detection theory (SDT) outcomes. When a target was present, a 'hit' was recorded when a driver successfully identified the target; a 'miss' was recorded when the driver failed to detect the stimulus. When a target was absent from a sign and a driver identified the absence, a 'correct rejection' was recorded. A 'false alarm' was recorded when a driver detected a stimulus that was absent.

Contingency analyses on driver detection of food business targets on specific service signs revealed a highly significant effect of target familiarity, with higher accuracy for familiar targets than unfamiliar (94.13% vs. 85.80%). Signal detection accuracy marginally decreased from 6 to 9 logo panels per sign when all targets were familiar (94.57% vs. 93.68%). In general, overall accuracy rates within age group for familiar targets in both 6 and 9 logo panel signs were high (see Table 1).

For purpose of the present study, only trials presenting familiar pictorial food logo targets were analyzed in order to eliminate any potential confounds in assessment of the workload and performance costs of successful sign target identification (i.e., familiar vs. unfamiliar targets). Consequently, only four of the eight trials completed by each participant in Zahabi et al. [13] study were considered in this investigation. Related to this, trials yielding incorrect target identifications were not analyzed as the outcome could have occurred due to participant failure to attend to roadway signs as instructed.

This approach allowed us to focus on driver workload and performance costs associated with successful target identifications ('hits' and 'correct rejections').

Table 1. Familiar target counts by logo count within age group (reprinted from [13])

Age group	Logo count	Hit	Correct rejection	Miss	False alarm	Overall accuracy	Accuracy: target present	Accuracy: target absent
Young	6	50	102	1	0	99.35%	98.04%	100.00%
	9	36	72	0	0	100.00%	100.00%	100.00%
Middle	6	34	74	3	0	97.30%	91.89%	100.00%
	9	35	76	3	0	97.37%	92.11%	100.00%
Elderly	6	19	69	16	0	84.62%	54.29%	100.00%
	9	24	83	18	1	84.92%	57.14%	98.81%

2.2 Apparatus

A driving simulator and an eye tracking system were used in conjunction to collect driver response measures and eye behavior. A STISIM Drive Model 400 driving simulator (System Technology, Inc., Hawthorne, CA) was integrated with three 37″ HDTV screens, audio speakers and a realistic vehicle cab in order to present simulated driving scenarios to participants. A FaceLAB 5.1 eye tracking system (Seeing Machines, Canberra, AUS) integrated two cameras an infrared lighting emitter to record eye movements at 60 Hz with an accuracy of 0.5° to 1° of rotational error. Eye Works Record software (Eye Tracking, Inc., Solana Beach, CA) was used to integrate the eye movement data from FaceLAB system with visual imagery shown on the simulator screens.

The simulator driving scenarios were designed to represent a realistic Interstate environment. All aspects of the roadway environment design, including the signage, followed guidelines of the North Carolina Department of Transportation (NCDOT) and Manual of Uniform Traffic Control Devices (MUTCD) [14]. Based on the resolution of simulation imagery and viewing distances of drivers, all roadway graphics were scaled (relative to real dimensions) so signs resembled actual freeway appearances.

2.3 Participants

In order to provide a basis for age group selection, a statistical analysis was conducted on crash data from 1995–2010 [3]. Results revealed three distinct groups. Drivers under 23 years of age or above 64 years old exhibited significantly higher crash rates than those who were in-between. With this in mind, three groups were defined for the study, including Young drivers (≤22 years of age), Middle-aged drivers (between 23 and 64 years of age), and Elderly drivers (≥65 years of age). Inclusion criteria for participant recruiting included a valid North Carolina driver's license and use of a motor-vehicle for a minimum of 5 h per week. Sixty participants completed the experiment. Each of the three age groups included 20 participants with balanced gender within group.

2.4 Experimental Design

The study followed a 3 × 2 mixed experiment design with one between- and one within-subject independent variable. Age group served as a grouping variable and all drivers were exposed to both levels of visual/cognitive workload, including the 6- and 9-panel logo sign configurations (see Fig. 1). Exposure to the various signage conditions was replicated for each participant. Test trial order was randomized within participant.

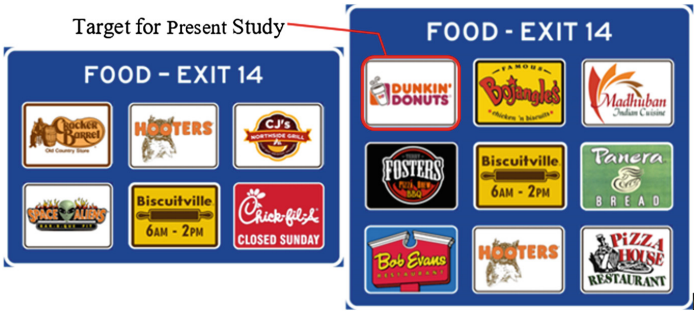


Fig. 1. Example 6- and 9-panel logo signs.

Dependent Variables. There were three dependent variables in this study, including vehicle speed deviation, lane deviation, and blink duration. Each type of measure is detailed in the following subsections.

Driving Measures. Previous studies have shown that elderly drivers often adopt a conservative style, which includes driving slower than average [15] or slowing down to compensate for an increase in reaction time to external stimuli [16]. Consequently, vehicle speed control is not a good indicator of driving performance as slower speeds may not indicate driver impairment but rather safety precautions or personal habits in response to complicated driving situations. In this study, driver reductions in vehicle speed (under workload) to levels below posted limits were considered indicators of adaptation for roadway conditions rather than performance decrements. The absolute value of deviations was recorded for statistical analysis purpose. Here, it is important to note that in Zahabi et al. [13] study, the speed control measure included deviations both above and below posted limit, and was used as performance indicator.

Previous research has found lane deviations to be a more sensitive measure of driving impairments than speed variability [17]. In addition, other work has demonstrated significant differences in the standard deviation of lateral vehicle positions among various driver age groups [18]. Therefore, lane deviation was selected as a driver performance decrement indicator in the current study. Participants were instructed to remain in the same lane throughout experiment trials; absolute deviation from the center of the lane was recorded.

Mental Workload Measures. Unlike secondary task performance measures or subjective rating of workload, remote eye-tracking measures are non-distracting, non-obtrusive to

performance, and objective in nature. Consequently, various eye tracking measures were selected as a basis for determining driver cognitive workload. Here, we note that eye-tracking measures of cognitive load were not analyzed by Zahabi et al. [13]. Eye activity measures, including blinks, fixations, saccades and pupil dilation, have been used extensively in previous studies as estimates of driver mental workload [19]. Blink frequency, blink duration, and mean pupil diameter (in particular) were found to systematically change in response to target density in target identification and tracking tasks [20]. In general, increases in cognitive workload are marked by increases in pupil diameter and decreases in blink duration (i.e., time the eyes are closed). Mixed results on blink rate have been reported in other studies due to the fact that visual demand and mental workload have opposite effects on blink rate; the former leads to blink-rate inhibition with processing of visual stimuli and latter leads to blink-rate increases in memory tasks [19, 21]. Consequently, in research manipulating visual demands in order to infer cognitive workload, blink rate is not an ideal measure of workload. Pupil dilation is said to be sensitive to cognitive demand, however, reflections of light on participant glasses can interfere with pupil detection when using an eye-tracking system. In the present study, over half the participants made use of visual prosthetics and, therefore, pupil dilation was not applied.

Based on the existing literature, blink duration is more sensitive to cognitive workload (induced by visual stimuli) than blink rate [21]. Since the purpose of present study was to assess driver workload associated with visual search and target identification, blink duration was selected as indicator of general cognitive workload. The physiological response measure of blink duration during signage exposure was normalized to account for individual differences in blinking behavior (see Eq. 1).

$$Normalized\ Blink\ Duration\ (ratio) = \frac{Observed\ Duration - Max\ Individual\ Duration}{Max\ Individual\ Duration} \tag{1}$$

Within each of the four experiment trials analyzed as part of this study, segments of roadway were identified in which drivers were not exposed to on-road signage or any other novel visual features, such as curves. Driver maximum blink duration was identified for each of these segments and the max of max durations was found across trials. This value was used as the normalizing factor in Eq. 1 for each individual driver.

2.5 Procedures

After completion of an informed consent form and demographic questionnaire, participants performed a 5 min training scenario to ensure they were capable of controlling the simulated vehicle. The scenario presented a normal rural Interstate driving situation identical to the experimental trials with the exception that only one interchange was presented with no sign target identification requirement. Drivers had to demonstrate an average lane deviation of less than 1.37' [22] and an average speed deviation from a posted limit of 1.0 mph or less. Training was repeated if participant performance did not meet either criterion. If participant performance remained unacceptable after three trials, the experiment was terminated.

Familiar food business targets were presented to drivers before each of the trials analyzed in this study. The experiment scenarios presented three Interstate interchanges spaced 3 miles apart. Before each interchange, there were three logos panel signs (lodging, food, and gas/attraction). Each food target appeared once within a particular scenario; i.e., the target logo was only present at one interchange. However, participants were informed that they might encounter the business target more than once to ensure focus throughout test trials.

Participants were asked to drive each scenario (the primary task) and verbally indicate whether the target logo was present in specific service signs adjacent to the Interstate (the secondary task). Three SDT outcomes were gathered for each target in every scenario (i.e., a "present" or "absent" response at each interchange). Driver workload and performance observations were recorded during a period when a food service sign became visible (650 ft. before the sign) and when the sign was obstructed from driver view by the rear-view mirror of the simulated vehicle (112 ft. before the sign) [11].

2.6 Statistical Data Analyses

Outliers and abnormal data were removed from response data sets prior to analysis. Abnormal data points were identified for both speed and lane deviation and removed by application of Cook's D Method. Speed deviation values above 10 mph ($n = 5$) and lane deviation values above 3 ft. ($n = 8$) were removed.

Concerning the normalized blink duration ratio data, driver perception of target signage was expected to cause increases in visual-cognitive workload and, therefore, blink duration was expected to decrease relative to the max duration exhibited when no signs were present. Consequently, the normalized blink duration ratio was expected to range in value from -1 to 0. Any positive observations ($n = 84$) were identified as abnormal data and were eliminated as such values were indicative of technical difficulties with the eye tracker or participant failure to follow instructions and attend to roadway signs. For ease of reader understanding, the absolute value of the resulting blink duration ratio is used for presentation with a scale from 0 to 1. A value of 1 indicates that a participant did not blink during the sign observation period and value of 0 indicates that drivers were very relaxed during the period. In general, the higher the absolute blink duration ratio value, the higher the cognitive demand.

Diagnostics were conducted on all dependent variables to ensure analysis of variance (ANOVA) assumptions of homoscedasticity (by Bartlett's tests) and residual normality (by inspection of normal probability plots and Shapiro–Wilk's test) were satisfied by the sanitized data sets. A mixed between- and within-subject ANOVA model was used for effects tests. Trial number was added to the model as a co-variate but was removed if insignificant. Duncan's Multiple Range Test (MRT) was applied if the age factor was significant in any response to identify specific differences among groups.

Among all three types of responses, only lane deviation met parametric test assumptions. Speed deviation required exponential transformation following the Box–Cox method with a power of 0.3. The normalized blink duration ratio was subjected to average rank transformation. (These transformations were necessary for inferential statistics; however, all graphical results and condition means are presented in raw units

for reader comprehension.) Finally, nonparametric Spearman's Rank correlation analyses were conducted to identify any relationship between driving performance, driver adaptation behavior and cognitive workload. A significance level of $p \leq 0.05$ was set as significance criterion for the study.

3 Result

3.1 Adaptation Behavior

Trial number was not significant in effect on speed deviation ($p > 0.05$) and was subsequently removed from the statistical model. Logo panel count (the workload manipulation; $F(1, 378) = 0.0615$, $p = 0.8043$, $1 - \beta = 0.0570$) and the interaction with age group ($F(2, 378) = 1.8673$, $p = 0.1560$, $1 - \beta = 0.3882$) were both found to be insignificant. Age group was, however, significant in effect ($F(2, 378) = 43.3069$, $p < 0.0001$, $1 - \beta = 1.00$). Duncan's MRT was applied and results revealed middle-aged and young drivers to exhibit similar speed deviations, while elderly drivers exhibited significantly greater adaptation behavior (higher speed reductions during sign exposure). Figure 2a shows that elderly drivers appeared to slow down much more than the other two age groups when encountering target signs, as an adaptation response to create additional sign detection time.

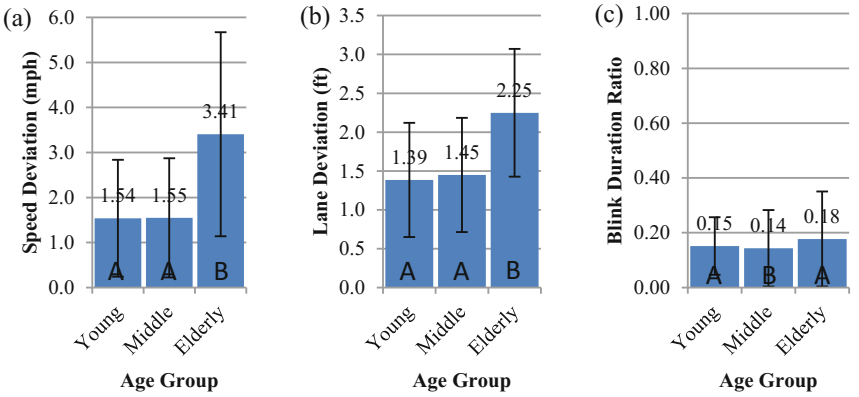


Fig. 2. Effect of age group on (a) speed deviation, (b) lane deviation, and (c) absolute normalized blink duration ratio relative to max.

3.2 Driving Performance

Trial number was found to be significant in statistical model of lane deviation ($F(1, 567) = 14.1517$, $p = 0.0002$, $1 - \beta = 0.9637$) while logo panel count ($F(1, 567) = 0.0039$, $p = 0.9504$, $1 - \beta = 0.0504$) and the interaction with age group ($F(2, 567) = 0.2019$, $p = 0.8173$, $1 - \beta = 0.0815$) were insignificant. It is possible that driver vehicle lateral control may have continued to improve across the experiment trials

following the training session. The main effect of age group was found to be significant ($F(2, 567) = 74.3759$, $p < 0.0001$, $1 - \beta = 1.00$). Duncan's MRT was applied and results revealed middle-aged and young driver performance to be similar, while elderly drivers exhibited significantly worse lane deviations. Figure 2b shows that elderly drivers produced significantly worse vehicle control (highest performance cost) during target sign search.

3.3 Cognitive Workload

Application of an ANOVA to the rank-transformed blink duration ratio constituted a non-parametric analysis. Although cognitive workload has been previously indicated by a decrease in raw blink duration, an increase in the blink duration ratio measure was indicative of an increase in driver cognitive load. The experiment trial number was significant in the model ($F(1, 491) = 12.8791$, $p = 0.0004$, $1 - \beta = 0.9476$) while the logo panel count ($F(1, 491) = 0.1103$, $p = 0.7400$, $1 - \beta = 0.0627$) and the interaction with age group ($F(2, 491) = 0.2677$, $p = 0.7653$, $1 - \beta = 0.0923$) were insignificant. It is possible that driver experience of cognitive workload from target identification may have continued to decrease throughout the experiment following training. The main effect of age group was significant ($F(2, 491) = 3.6559$, $p = 0.0265$, $1 - \beta = 0.6730$). Duncan's MRT was applied and results revealed significantly higher blink duration ratios for elderly and young drivers, as compared with middle-aged drivers. The two former groups exhibited signs of significantly higher cognitive load in response to the target identification task than middle-aged drivers. Figure 2c reveals large response variations within age group.

3.4 Correlation Analysis

Spearman correlation analyses revealed driving performance (lane deviation) to have a positive relationship with adaptation behavior (speed deviation) ($\rho = 0.1925$, $p < 0.0001$). In general, as driving performance degrades, drivers exhibit greater adaptation behavior. A positive correlation was also found between blink duration ratio and lane deviation ($\rho = 0.1066$, $p = 0.0129$). As drivers experience higher cognitive workload, performance degradation increases. However, no correlation was found between cognitive workload and adaptation behavior ($p > 0.05$). In summary, as driver cognitive load increased, performance degradation (lane deviation) was found to increase while adaptation behavior (speed deviation) remained unaffected. Related to this, elderly drivers appeared consistently conservative in vehicle control when loaded with the target detection task.

4 Discussion

Returning to the expectations of this research study, we contended that greater density of external visual information would increase driver task workload and compromise performance. In specific, more logo panels on freeway service signs were expected to

increase mental workload (H1), driver adaptation behavior (H2), and performance degradations (H3) across all ages group. Surprisingly, these particular expectations were not supported by results. There was no main effect of logo panel count on driver performance, adaptation or workload responses. However, Zahabi et al. [13] also found that the 9-panel specific service sign configuration did not have a significant effect on driver behavior when simultaneously taxed with familiar target detection. It is possible that the experimental workload manipulation may not have been sufficient to cause changes in driver visual attention. In the particular case of elderly drivers, it is also possible that both 6 and 9 panel signs posed an attentional challenge.

We also hypothesized that elderly drivers would demonstrate higher mental workload (H4), adaptation behavior (H5) and driving performance degradations (H6) during successful target detections relative to the middle-aged group and possibly comparable outcomes to younger drivers. All of these hypotheses were supported by results. Elderly drivers were observed to have comparable cognitive workload to young drivers and both groups exhibited significantly higher load than middle-aged drivers. Elderly drivers exhibited significantly greater adaptation behaviors (greatest speed deviations) and performance degradations (greatest lane deviations) during target identification by comparison with other age groups. In general, as age increases, adaptation behavior increases and vehicle control degrades.

The age-related effects observed in this study should be considered in conjunction with findings of Zahabi et al., as summarized in Table 1. They found that when a logo target was present in a sign, elderly drivers were only 54–57% correct in identification. That is to say that nearly half of the elderly licensed drivers in Zahabi et al. study could not identify familiar targets on freeway signs even with extra response time gained through adaptive control behaviors. Although the adaptation behaviors and performance degradations of elderly drivers observed in this study were significantly greater than for other drivers, the differences between age groups were not as extreme as one might expect based on the Zahabi et al. target detection accuracy results. Together, these findings indicate that elderly drivers are more likely to abandon secondary task goals (like target search) to focus on primary driving tasks, as compared to other age groups.

The correlation analyses revealed that as driver cognitive workload increases, performance degradations (lane deviation) also increase. No previous study has examined the correlation between driver cognitive workload, based on real-time eye activity measures in the use of on-road signage, with lane deviations. A previous study involving driver navigational challenges (search for landmarks) measured the standard deviation of lane position and found greater variability with a challenge vs. no navigation demand [16]. In general, these results indicate that higher driver visual demand leads to an increase in lateral vehicle position deviations. In regard to the absence of a relation between cognitive load and driver adaptation behavior, there were no apparent effects of the logo count manipulation on driver visual behavior and vehicle control. It is also possible that cognitive load does not vary systematically with adaptation behaviors. Unfortunately, there are few, if any, existing driving studies on adaptation behavior that have also assessed cognitive workload. Additional research is needed in this area.

5 Conclusion

On the basis of our findings, there is some evidence that even normal freeway driving and external visual stimulus detection may be challenging for a significant portion of driving population, specifically elderly drivers. Increased density of visual information presented in specific service signs (6 vs. 9 logo panels) does not appear to have a significant negative impact on driver behavior and workload when performing familiar target detection. Although our signage manipulation did not lead to changes in driver performance and workload, age-related differences were found. Both young and elderly drivers experience significantly higher cognitive workload than middle-aged drivers when attempting to detect familiar sign targets. Furthermore, 65+ elderly drivers appear to require additional time to effectively monitor the roadway environment due to age-related declines in ability. They exhibit significantly greater adaptive vehicle control behaviors than younger drivers. Even when adapting vehicle control, elderly driver target identification accuracy is very low, including 54.29% for 6 logo panels.

Our results raise concerns about the design of existing highway facilities for accommodation of elderly driver capabilities. With the knowledge that driving capabilities decline with age, roadway signage needs to be designed to accommodate the elderly segment of driving population rather than relying on driver self-awareness of such limitations and/or avoidance of challenging roadway environments. Our findings emphasize the need for agencies to consider the capabilities of all driver age groups during design and evaluation of signs. All roadway signage and facilities should be designed according to proper guidelines that reduce driver distraction and accommodate for potential performance decrements.

Although this study did not review specific statistics on elderly driver involvement in freeway accidents, our results show that driver visual behavior and performance vary with age, which are potential roadway incident contributors, according to prior research [1]. Additional driving simulator studies on elderly driver target information acquisition under normal roadway conditions are recommended to determine if accuracy results agree with the present findings. State transportation agencies may want to consider further examining the capability of elderly drivers to effectively make use of existing highway facilities designs through driving simulator based testing and/or to ensure that signage designs are conducive to elderly driver use. The overarching objective of such research is to mitigate the potential for freeway driving incidents among elderly.

Acknowledgments. This study was supported by a Grant from the North Carolina Department of Transportation (NCDOT; Grant No. 2015-46). Special thanks go to Maryam Zahabi, Carl Pankok, Patricia Machado, Yi-Fan Liao, Yulin Deng and Drs. Joseph Hummer and William Rasdorf for their work on the Zahabi et al. [13] study, which served as a basis for the present work.

References

1. Kaber, D.B., Zhang, Y., Jin, S., Mosaly, P., Garner, M.: Effects of hazard exposure on driver situation awareness and performance and the interaction with roadway complexity and driver age. Transp. Res. F **15**, 600–611 (2012)

2. Mcknight, A., Mcknight, A.: Multivariate analysis of age-related driver ability and performance deficits. Accid. Anal. Prev. **31**(5), 445–454 (1999)
3. Tefft, B.C.: Motor Vehicle Crashes, Injuries, and Deaths in Relation to Driver Age: United States, 1995–2010. Technical report. AAA Foundation for Traffic Safety (2012)
4. Kass, S.J., Cole, K.S., Stanny, C.J.: Effects of distraction and experience on situation awareness and simulated driving. Transp. Res. F **10**(4), 321–329 (2007)
5. Motor Vehicle Crashes: Overview. Publication DOT-HS-812-101. Technical report. National Highway Traffic Safety Administration (2014)
6. Newgard, C.D.: Defining the "older" crash victim: the relationship between age and serious injury in motor vehicle crashes. Accid. Anal. Prev. **40**(4), 1498–1505 (2008)
7. Hummer, J.E.: Supplemental interchange signing and driver control behavior. Transp. Res. Rec. **1213**, 56–63 (1989)
8. Ho, G., Scialfa, C.T., Caird, J.K., Graw, T.: Visual search for traffic signs: the effects of clutter, luminance, and aging. Hum. Factors J. Hum. Factors Ergon. Soc. **43**(2), 194–207 (2001)
9. Cantin, V., Lavallière, M., Simoneau, M., Teasdale, N.: Mental workload when driving in a simulator: effects of age and driving complexity. Accid. Anal. Prev. **41**(4), 763–771 (2009)
10. Dagnall, E.E., Katz, B.J., Bertola, M.A., Shurbutt, J.: Business logo signing evaluation of the number of logo panels on specific service signs. Transp. Res. Rec. **2365**, 82–89 (2013)
11. Zhang, Y., Harris, E., Rogers, M., Kaber, D.B., Hummer, J., Rasdorf, W., Hu, J.: Driver distraction and performance effects of highway logo sign design. Appl. Ergon. **44**, 472–479 (2013)
12. Kaber, D., Pankok, C., Corbett, B., Ma, W., Hummer, J., Rasdorf, W.: Driver behavior in use of guide and logo signs under distraction and complex roadway conditions. Appl. Ergon. **47**, 99–106 (2015)
13. Zahabi, M., Machado, P., Pankok, C., Lau, M.Y., Liao, Y., Hummer, J., Rasdorf, W., Kaber, D.B.: The Role of Driver Age in Performance and Attention Allocation Effects of Roadway Sign Count, Format and Familiarity. Appl. Ergon. 63, 17–30 (2017)
14. Manual on Uniform Traffic Control Devices for Streets and Highways (MUTCD). FHWA, U.S. Department of Transportation (2009)
15. Zhang, Y., Kaber, D.B.: An empirical assessment of driver motivation and emotional states in perceived safety margins under varied driving conditions. Ergonomics **56**(2), 256–267 (2013)
16. Trick, L.M., Toxopeus, R., Wilson, D.: The effects of visibility conditions, traffic density, and navigational challenge on speed compensation and driving performance in older adults. Accid. Anal. Prev. **42**(6), 1661–1671 (2010)
17. Verster, J.C., Roth, T.: Effects of central nervous system drugs on driving: speed variability versus standard deviation of lateral position as outcome measure of the on-the-road driving test. Hum. Psychopharmacol. Clin. Exp. **29**(1), 19–24 (2013)
18. Fofanova, J., Vollrath, M.: Distraction while driving: the case of older drivers. Transp. Res. F **14**(6), 638–648 (2011)
19. Marquart, G., Cabrall, C., Winter, J.D.: Review of eye-related measures of drivers' mental workload. Procedia Manuf. **3**, 2854–2861 (2015)
20. Van Orden, K.F., Limbert, W., Makeig, S., Jung, T.: Eye activity correlates of workload during a visuospatial memory task. Hum. Factors J. Hum. Factors Ergon. Soc. **43**(1), 111–121 (2001)
21. Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., Montanari, R.: Driver workload and eye blink duration. Transp. Res. F **14**(3), 199–208 (2011)
22. Horrey, W.J., Wickens, C.D.: Driving and side task performance: the effects of display clutter, separation, and modality. Hum. Factors **46**(4), 611–624 (2004)

Comparison Between Young Male Drivers' Self-assessed and Objectively Measured Driving Skills

Laila M. Martinussen[1(✉)], Mette Møller[1], and Carlo G. Prato[2]

[1] Department of Management, Technical University of Denmark, Diplomvej, Bygning 372, 2800 Kongens Lyngby, Denmark
laima@dtu.dk

[2] School of Civil Engineering, The University of Queensland, St Lucia 4072, Brisbane, Australia

Abstract. Self-assessment of skills is a self-generated feedback process that contributes to confidence in one's skills. The higher one's self-assessed skills, the more likely one is to feel competent a particular domain thereby influencing the related behaviors. Drivers' self-assessed driving skills are not always accurate, which may cause serious problems such as underestimation of risk, reckless driving and accidents. Most previous research on self-assessment of driving skills did not compare self-reported skills to objectively measured driving skills, so the aim of this study was to test the accuracy of young male drivers' self-assessments of driving skills using a driving simulator, and to examine whether self-assessment accuracy varied with driving skill, experience or sensation-seeking propensity. Results showed that the drivers' self-assessments were inconsistent with their driving performance, and this inconsistency varied with driving skill, driving experience and sensation-seeking propensity in a safety-critical way.

Keywords: Self-assessed driving skills · Driving skills · Driving experience · Sensation seeking · Driving simulator · Young male drivers

1 Introduction

Self-assessment is a self-generated feedback process that can contribute to confidence in one's skills. When the self-assessed skill level is high, one is likely to feel more competent and this influences behavior [1]. Accurate self-assessment is essential because it ensures an appropriate level of confidence in one's skills [2]. Underestimation of skills may lead to unnecessary, self-imposed restrictions on activity whereas overestimation may lead individuals to participate in activities that are beyond their capability [1]. Self-assessments of driving skills have shown to not always be accurate [3, 4], which may cause serious problems such as underestimation of risk and thus unintended sensation seeking and reckless driving and accidents [4–7]. Sensation seeking is related to self-efficacy. Self-efficacy is one's belief in own capacity and competence in a given domain [8], which in turn is closely related to self-assessed competence [2]. High sensation seekers may be more prone to develop positive

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_75

perceptions of their self-efficacy through successful performance in the thrilling situations they expose themselves to because arousal-level is one of the means persons assess their self-efficacy through. Behavioral performance and self-assessed skills are in constant interaction with crucial implications for the development of safe driving behavior.

Young male drivers are over-represented in accident statistics [9–11]. Young drivers have a tendency to overestimate their driving skill [3] and this inaccurate self-assessment of skill may be a contributing factor. There is evidence from other fields that the accuracy of self-assessments vary within sub-groups of people [12–14]. For example, it has been shown that low-skilled individuals are less accurate in their self-assessments than high-skilled individuals [12]. In a driving context, drivers who passed their driving test made more accurate self-assessments of driving skills than drivers who failed the test did [15].

Given that driving behavior and skill are known to vary according to individual factors such as gender, age, driving experience, personality etc. [5, 16–19] it is important to establish whether such factors also influence the accuracy of self-assessments. This study contributes to this by investigating the influence of driving skill, driving experience and sensation-seeking propensity on the accuracy of driving self-assessments.

Several previous studies have found that drivers tend to overestimate their driving skills [for an overview, see 20]. In the majority of these studies, drivers were asked to compare their driving skills to the skills of the average driver using self-report measures. Other methods used in the context of self-assessment of driving skills include comparison of a driving license examiner's assessment of a driver's skills to the driver's own self-assessment [e.g. 3, 15]. Studies have found that about 40% of the drivers overestimated their skills [15], and that young drivers overestimated their driving skills [3]. The use of self-reports as a method of assessing driving skill has generally been criticized due to the lack of an objective measure of driving skill [20]. Although assessments including driving license examiners professionally based assessments do not rely completely on subjective assessments made by the drivers, subjective assessment of driving skills is still the key element and may therefore possibly also be subjected to human flaws as well as inter-rater variability. Another way to validate self-reports could include the use of naturalistic driving. However, issues related to the safety of the involved drivers make this method problematic. In addition, it is impossible to expose all participants to identical driving scenarios. The use of a driving simulator in studies on self-assessment of driving skills is a suitable way to overcome some of the challenges of subjectivity. Further the use of a virtual, controlled experimental setting ensures that drivers' skills and behavior can be assessed systematically, safely and objectively [21]. Finally, the use of a driving simulator enables the researcher to measure the drivers' performance of specific driving skills for comparison with the driver's self-assessment of the same driving skills.

Based on the above, the aim of this study was to test the accuracy of young male drivers' self-assessments of driving skills using a driving simulator, and to examine whether self-assessment accuracy varied with driving skill, experience or

sensation-seeking propensity. We hypothesized that driving skill level, driving experience and sensation seeking would affect self-assessment accuracy. Specifically, we hypothesized inaccurate self-assessment among less skilled drivers and drivers with little experience, and we expected that thigh sensation seekers would have high confidence when self-assessing their skills.

2 Method

2.1 Participants

The participants were male students from the Technical University of Denmark (DTU). Participants were aged between 18 and 31 years old (mean age 23.10, SD 2.54) and were recruited on campus or via a DTU Facebook site. Informed consent was obtained by all participants. All participants had a driving license for cars (type B).

Table 1. Description of the scenarios

Scenario	Environment	Scenario description
1	City	As the driver approaches a parked bus (and other parked cars) a pedestrian runs out from behind the front of the bus into the street
2	City	As the driver approaches a traffic light it turns from green to amber, then red
3	Rural	A car approaches from an adjacent road on the left side and merges in front of the driver. The car accelerates and decelerates in front of the driver
4	Rural	A moped is travelling slowly ahead of the driver
5	City	A car is approaching rapidly on the right side of the driver at a crossing where the driver has green and the approaching car red; the approaching car stops just before moving into the driver's lane
6	City	A car is approaching rapidly on the right side in front of the driver and crosses the road when the driver has green and the crossing car red

2.2 Equipment and Materials

Driving Simulator. The experiment was conducted using a fixed-base driving simulator with a 180º visual field with rear and side view mirrors. The simulator software used was 'ScanEr Studio'. Six different scenarios were scripted in ScanEr Studio in order to measure driving skills specified in the DSI (see Tables 1, 2). The six scenarios were set in a city environment, a representation of an urban area or a two-way rural road. A Latin Square procedure was used to counterbalance the order in which participants

experienced the six scenarios. Driving data were extracted at 20 Hz. Additionally a qualitative measure of overtaking performance was used. Each driver's overtaking manoeuvre was judged by the experimenter in terms of location of the overtaking and passing gap allowed [e.g., 22, 23]. Based on this, the skill level of each overtaking manoeuvre was scored (0 = overtook in dangerous circumstances (equivalent to 'well below average'); 2 = did not overtake ('about average'); 4 = overtook safely ('well above average').

Table 2. The scenarios with the DSI items measured in each and corresponding behavioural measures

Scenario	Behavioral indicators		
	DSI item	Skill type	
1	Prediction of situation ahead	P-M: Hazard prediction	Speed at event time (ET) -5 s minus speed at ET -1 s
	Performance in critical situations	P-M: Hazard detection	Latency to braking (after start of the event)
2	Avoiding unnecessary risk	S: Hazard prediction	Speed at ET -5 s minus speed at ET -1 s
3	Fast reactions	P-M: Hazard detection	Latency to braking
4	Fast reactions	P-M: Hazard detection	Latency to braking
5	Overtaking	P-M: Overtaking	Whether the driver overtakes, and if so whether it is done safely
	Keeping sufficient distance	S: maintenance of safe gap to car in front	Minimum time to collision
6	Keeping sufficient distance	P-M: Overtaking	Whether the driver overtakes, and if so whether it is done safely
	City	S: maintenance of safe gap to car in front	Minimum time to collision

Driving Skill Inventory (DSI). The DSI is a questionnaire used to capture drivers' self-assessed perceptual-motor skills and safety skills. Perceptual-motor skills are technical driving skills and car-handling skills; safety skills are accident avoidance skills and careful driving practices. Perceptual-motor skills are based on information-processing and motor skills, whereas safety skills are related to attitudes and personality factors [24]. Participants were asked to report their driving skills on a five-point scale ranging from 0 = 'Well below average' to 4 = 'Well above average' compared to the average driver of their age and gender group. Table 2 specifies the DSI items used in this study and the behavioural indicator used to assess them in the simulator.

Brief Sensation Seeking Scale (BSSS). The BSSS is a questionnaire used to measure sensation seeking in adolescents and young adults [25]. Sensation-seeking propensity was assessed by asking participants to assess whether eight statements relating to thrills, adventure seeking, social disinhibition, susceptibility to boredom and experience seeking were true of them or not.

2.3 Procedure

The simulator task consisted of one introductory drive (rural environment) and six experimental drives with a pause after the third drive (see Table 1, 2). Participants were instructed to remain on the main road and drive as they normally do. After completing the driving simulation task participants filled out the BSSS. Half of the participants completed the DSI before the simulator task and the remainder completed the DSI one week after the simulator task. This procedure was used to control for potential order effects [for more information see 26].

2.4 Data Preparation and Analysis

All behavioural measures can be seen in Table 2. All behavioural data were re-coded using a five-point scale corresponding to the DSI scale (0 = 'Well below average' to 4 = 'Well above average'). For continuous variables the scale was based on the percentile distribution of the performance of the sample, thus the least skilled participant and the most skilled participant were used as thresholds. Data on overtaking were categorical and were therefore mapped as follows: 0 = overtook in dangerous circumstances (equivalent to 'well below average'); 2 = did not overtake ('about average'); 4 = overtook safely ('well above average').

To test whether self-assessments of driving skill varied according to objectively measured driving performance, driving experience and sensation-seeking propensity the sample was divided into two groups with respect to each factor. Participants were assigned to the high- or low-driving skill group on the basis of whether their total score on the behavioural measures was above or below the median for the sample. A similar procedure was used to create low- and high-experience groups (on the basis of weekly mileage) and high- and low-sensation seeking groups (on the basis of BSSS score). Accuracy of self-assessed driving skills was tested in the three groups (driving skill, driving experience, sensation seeking) using paired-sample non-parametric Wilcoxon Signed Ranks t-tests (see Table 3).

Table 3. Comparison of driving performance and self-assessed driving skill in the high and low driving skill, -, driving experience, - and sensation seeking groups

Scenario	Driving skill		Driving experience		Sensation seeking	
Behavioural measure: hazard prediction						
1	Low	High	Low	High	Low	High
Z	−1.754c	−2.116c	−1.138c	−2.848c	−2.366c	−1.327c
p-value	0.079	0.034*	0.255	0.004*	0.018*	0.185
Behavioural measure: hazard detection						
Z	Low	High	Low	High	Low	High
p-value	−2.622c	−0.535c	−1.976c	−1.561c	−1.069c	−2.303c
	0.009*	0.593	0.048*	0.119	0.285	0.021*
Behavioural measure: hazard prediction						
2	Low	High	Low	High	Low	High
Z	−2.098c	−0.366c	−0.292c	−1.897c	−1.051b	−2.631c
p-value	0.036*	0.714	0.770	0.058	0.293	0.009*
Behavioural measure: hazard detection						
3	Low	High	Low	High	Low	High
Z	−3.019c	−1.259c	−2.201c	−2.565c	−2.506c	−1.964c
p-value	0.003*	0.208	0.028*	0.010*	0.012*	0.050*
Behavioural measure: hazard detection						
4	Low	High	Low	High	Low	High
Z	−3.443c	−0.288b	−1.787c	−2.095c	−0.792c	−2.754c
p-value	0.001*	0.773	0.074	0.036*	0.428	0.006*
Behavioural measure: overtaking						
5	Low	High	Low	High	Low	High
Z	−1.029c	−0.660c	−0.112c	−2.157c	−0.322b	−0.364c
p-value	0.304	0.509	0.911	0.031*	0.747	0.716
Behavioural measure: safe gap to car in front						
	Low	High	Low	High	Low	High
Z	−2.587c	−0.482c	−2.517c	−0.073b	−0.395c	−2.732c
p-value	0.010*	0.630	0.012*	0.942	0.693	0.006*
Behavioural measure: overtaking						
6	Low	High	Low	High	Low	High
Z	−0.226b	−0.660c	−1.024b	−1.730c	−1.024b	−1.730c
p-value	0.821	0.509	0.306	0.084	0.306	0.084
Behavioural measure: safety gap to car in front						
	Low	High	Low	High	Low	High
Z	−1.165c	−1.267c	−2.110c	−0.090b	−0.247c	−2.178c
p-value	0.244	0.205	0.035*	0.928	0.805	0.029*

Notes. c, positive difference between performance and self-assessment;
b, negative difference performance and self-assessment
* $p < .05$.

3 Results

See Table 3 for the results. There were discrepancy between self-assessed skills and objective behavioral measures in the simulator. The low-skilled drivers, experienced drivers and high sensation-seeking drivers had the largest discrepancy. Self-assessments of hazard prediction, hazard detection and maintaining a safe distance to car in front were particularly inaccurate in the low-skilled drivers. More experienced drivers were most inaccurate in their self-assessments of hazard prediction, hazard detection and overtaking skills, whereas less experienced drivers were most inaccurate in their self-assessments of hazard detection skill and ability to maintain a safe gap to the car in front. Drivers with high sensation-seeking propensity were inaccurate in their self-assessments of hazard prediction, hazard detection and ability to maintain a safe gap to the car in front. Drivers with low sensation-seeking propensity were inaccurate in their self-assessments of hazard prediction and hazard detection.

4 Discussion

The aim of this study was to test how accurate young male drivers are when they self-assess specific driving skills, and to test whether the accuracy of self-assessments varied according to driving skills, driving experience and sensation-seeking propensity. The results indicated that young male drivers' self-assessments of driving skills are inconsistent with their driving performance; mostly pronounced for hazard prediction and hazard detection skills. Also, this inconsistency varies with driving skill, driving experience and sensation-seeking propensity. In line with results in other fields [12, 27] we found that self-assessments of less-skilled participants (low driving skills) were more inaccurate than high-skilled drivers. In addition, experienced drivers, and high sensation-seeking drivers were inaccurate in their self-assessment. Hazard prediction, hazard detection and ability to maintain a safe gap to the car in front were particularly inaccurate. Because of the small sample size, the results should be read with care. Nonetheless, the findings suggest that developing preventive policies and interventions may benefit from taking into account factors affecting self-assessments of skill such as skill level, experience, and sensation-seeking propensity. An example could be awareness campaigns that focus on the importance of accurate self-assessment of driving skills.

Overestimation of driving skills among young drivers have been found before [3]. However, the overconfidence is not seen in all driving situations [28] indicating that the relation is complex. Studies also show that the confidence depends on the method (average versus peer comparison) used to assess young drivers' self-assessment [3]. De Craen et al. [3] also found that young drivers are more confident in their self-assessment than experts who indicate their opinion about the participants' performance. The current study present specific driving skill measures, unlike previous studies [including; 3] that presented general driving skill measures. One would expect that the results might be different and probably more comprehensive when looking at specific, -rather than general driving skill.

It has been suggested that self-assessment is based on a top-down approach in which pre-conceived beliefs about one's skills are given great weight with the result that less skilled drivers mistakenly assess their driving skills as good [29]. Thus, the inaccuracy of less skilled performers' self-assessments is compounded by the fact that the skills needed to perform proficiently are similar to those needed to evaluate one's own performance. Less-skilled performers are unable to recognize the discrepancy between skilled and unskilled performance, leading to inaccurate self-assessment [14]. The result that less skilled drivers' self-assessments was more inaccurate than high skilled drivers' self-assessment is consistent with this account. Alike, self-assessments of driving skill have been found to be more accurate among learner drivers who passed their driving test than those who failed [15]. The results for the more skilled drivers were surprising, as it has been shown in other domains that highly skilled performers assess their own skills accurately but tend to overestimate the skills of others and thus underestimate their own relative skill [14].

Less skilled drivers' inaccuracy of self-assessments might be affected by the feedback drivers receive whilst driving. Accident involvement might be considered the ultimate form of negative feedback on driving skill, but since accidents are rare [3.4 deaths per 1 billion vehicle-km in Denmark in 2012; 30] less skilled drivers' overestimation of their competence may be reinforced by the fact that they have not been involved in an accident. This logic might also contribute to the inaccuracy of more experienced drivers' self-assessments, as they too are unlikely to be involved in an accident in spite of their driving frequency [31].

The high sensation-seeking drivers' inaccurate self-assessments is of particular importance as it suggests that high sensation-seeking drivers pose a dual risk because of high risk acceptance plus overestimation of own driving skill. Reports that high sensation seekers are more likely to be involved in accidents than low sensation seekers [4, 32] are consistent with this argument. Sensation seekers enjoy risky situations [33] and the thrill they experience in such situations further reduces their ability to self-assess accurately. It has been suggested that stress could mistakenly be interpret as a sign of low skill, which might lead to a lack of confidence when engaging in an activity [8, 34]. In contrast, thrill might mistakenly be interpreted as a sign of high skill, leading to high confidence in one's capability [8, 34]. Thus, sensation seekers may interpret the thrill that they experience in risky situations as an indication of high skill. Future research could profitably explore the relationship between sensation seeking, driving behavior and self-assessment of driving skill in more depth by larger sample sizes and also by measuring other variables related to sensation seeking such as testosterone and cortisol levels.

Self-assessments of hazard prediction and detection were found to be particularly inaccurate in the present sample. Perception and prediction of hazards are related to perception of risk and self-assessed ability to cope with the hazard, and young drivers have shortcomings in both these areas [35, 36]. Perception and prediction of hazards and self-assessment of driving skill is guided by higher-order self-regulatory skills such as general skills for identifying task demands, constructing and evaluating different courses of action, setting suitable goals, creating appropriate incentives to sustain one's effort in the face of setbacks, and coping with stress and negative thoughts [37]. These skills are relevant to skillful driving performance [38]. However, the neurobiological

systems which support higher-order self-regulatory cognitive processes such as evaluation, decision making and impulse control are not fully developed until about 25 years [39, 40], making 87.2% of the present sample subjected. Young drivers' neurocognitive immaturity may explain why their self-assessments were particularly inaccurate in relation to hazard perception and detection.

Limitations to the study should be acknowledged: (1) Because of the sample size caution towards the generalization of the results is suggested. However, the findings are robust [see, 41] and the sample was chosen as homogeneous to reduce possible confounders' effect (gender, age). Although the sample was recruited on campus and via a DTU Facebook site and was not representative of the population, the findings provide a clear indication of the accuracy of self-assessment and provide a robust basis for further discussion on the subject. (2) The skills measured in a simulator might be different from the actual driving skills (ecological validity). To overcome this limitation, the results could be verified by natural driving data, but this would generate an ethical issue because of the possible risk issues related to measuring driving skills in a natural driving setting. (3) The validity of the scales used. Participants might have various perceptions of what an average driver skill is, making it difficult to compare across the sample. This issue was overcome by informing the participants to compare against their own gender and age group i.e. against the same standards. One cannot be sure that participants actually do compare with the same average group nor have the same perception of what average are [42].

5 Conclusion

The results indicate that the accuracy of self-assessment of driving skills among young male drivers vary in a safety critical way, particularly with regard to perception and prediction of hazards. Measures to improve drivers' self-assessment are relevant, but further studies are needed to allow preventive policies and interventions to take factors affecting self-assessments of skill such as skill level, experience, and sensation-seeking propensity into account.

Acknowledgments. Financial support from by the Innovation Fund is greatly appreciated.

References

1. Bandura, A.: Self-Efficacy. The Exercise of Control. W. H. Freeman and Company, New York (1997)
2. Coronado-Aliegro, J.: The Effect of Self-Assessment on the Self-Efficacy of Students Studying Spanish as a Foreign Language. Graduate Faculty, School of Education, University of Pittsburgh (2006)
3. De Craen, S., Twisk, D.A., Hagenzieker, M.P., Elffers, H., Brookhuis, K.A.: Do young novice drivers overestimate their driving skills more than experienced drivers? Different methods lead to different conclusions. Accid. Anal. Prev. **43**, 1660–1665 (2011)

4. Gregersen, N.P.: Young drivers' overestimation of their own skill—an experiment on their relation between training strategy and skill. Accid. Anal. Prev. **28**(2), 243–250 (1996)
5. Özkan, T., Lajunen, T.: What causes the difference in driving between young men and women? The effects of gender roles and sex on young drivers' behavior and self-assessment of skills. Transp. Res F Psychol. **9**, 269–277 (2006)
6. Özkan, T., Lajunen, T., Chliaoutakis, J., Parker, D., Summala, H.: Cross-cultural differences in driving skills: a comparison of six countries. Accid. Anal. Prev. **38**, 1011–1018 (2006)
7. Sümer, N., Özkan, T., Lajunen, T.: Asymmetric relationship between driving and safety skills. Accid. Anal. Prev. **38**, 703–711 (2006)
8. Bandura, A.: Regulative Function of Perceived Self-Efficacy. Lawrence Erlbaum Associates, Hillsdale (1994)
9. Hansen, A.S. Jensen, C.: Risiko I trafikken 2007–2010 [Risk in traffic 2007–2010]. DTU Transport, Report 4 (2012)
10. OECD: Young Drivers—The Road to Safety. Transport Research Centre. ECMT. OECD Publishing, Paris (2006)
11. Twisk, D.A., Stacey, C.: Trends in young driver risk and countermeasures in European countries. J. Safety Res. **38**, 245–257 (2007)
12. Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., Kruger, J.: Why the unskilled are unaware: further explorations of (absent) self-insight among the incompetent. Organ. Behav. Hum. Dec. **105**(1), 98–121 (2008)
13. Haun, D.E., Zeringue, A., Leach, A., Fole, A.: Assessing the competence of specimen-processing personnel. Lab. Med. **31**, 633–637 (2000)
14. Kruger, J., Dunning, D.: Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J. Pers. Soc. Psychol. **77**, 1121–1134 (1999)
15. Mynttinen, S., Sundström, A., Vissers, J., Koivukoski, M., Hakuli, K., Keskinen, E.: Self-assessed driver competence among novice drivers–a comparison of driving test candidate assessments and examiner assessments in a Dutch and Finnish sample. J. Safety Res. **40**, 301–309 (2009)
16. Lajunen, T., Corry, A., Summala, H., Hartley, L.: Cross-cultural differences in drivers' self-assessments of their perceptual-motor and safety skills: Australians and Finns. Pers. Indiv. Differ. **24**, 539–550 (1998)
17. Lucidi, F., Giannini, A.M., Sgalla, R., Mallia, L., Devoto, A., Reichmann, S.: Young novice driver subtypes: relationship to driving violations, errors and lapses. Accid. Anal. Prev. **42**, 1689–1696 (2010)
18. Rimmö, P.A.: Aberrant driving behaviour: homogeneity of a four-factor structure in samples differing in age and gender. Ergonomics **4**(8), 569–582 (2002)
19. Zuckerman, M.: Sensation seeking and risky driving, sports, and vocations. In: Zuckerman, M. (ed.) Sensation Seeking and Risky Behavior, pp. 73–106. American Psychological Association, Washington (2007)
20. Sundström, A.: Self-assessment of driving skill—a review from a measurement perspective. Transp. Res. F Psychol. **11**(1), 1–9 (2008)
21. Boyle, L.N., Lee, J.D.: Using driving simulators to assess driving safety. Accid. Anal. Prev. **42**, 785–787 (2010)
22. Farah, H., Bekhor, S., Polus, A., Toledo, T.: A passing gap acceptance model for two- lane rural highways. Transportmetrica **5**, 159–172 (2009)
23. Farah, H., Yechiam, E., Bekhor, S., Toledo, T., Polus, A.: Association of risk proneness in overtaking maneuvers with impaired decision making. Transp. Res. F Psychol. **11**, 313–323 (2008)

24. Lajunen, T., Summala, H.: Driving experience, personality, and skill and safety-motive dimensions in drivers' self-assessments. Pers. Indiv. Differ. **19**(3), 307–318 (1995)
25. Hoyle, R.H., Stephenson, M.T., Palmgreen, P., Lorch, E.P., Donohew, R.L.: Reliability and validity of a brief sensation seeking scale. Pers. Indiv. Differ. **32**(3), 401–414 (2002)
26. Ozby, P.C.: Methods of Behavioral Research, 10th edn. McGraw Hill, New York (2009)
27. Dunning, D., Heath, C., Suls, J.M.: Flawed self-assessment. Implications for health, education and the workplace. Psychol. Sci. Public Interest. **5**(3), 69–106 (2004)
28. Mayhew, D.R., Simpson, H.M.: The role of driving experience. In: Implications for the Training and Licensing of New Drivers. Insurance Bureau of Canada, Toronto (1995)
29. Dunning, D., Johnson, K., Ehrlinger, J., Kruger, J.: Why people fail to recognize their own incompetence. Curr. Dir. Psychol. Sci. **12**(3), 83–87 (2003)
30. OECD/ITF: Road Safety Annual Report 2014 (official report). International Traffic Safety Data and Analysis Group, International Transport Forum, Paris (2014)
31. Elvik, R.: Exploratory study of mechanisms by which exposure influences accident occurrence. Transp. Res. Rec. **2148**, 7–82 (2010)
32. Elander, J., West, R., French, D.: Behavioral correlates of individual differences in road traffic crash risk: an examination of methods and findings. Psychol. Bull. **113**, 279–294 (1993)
33. Zuckerman, M.: Sensation seeking. In: London, H., Exner, J. (eds.) Dimensions in Personality, pp. 487–559. Wiley, New York (1978)
34. Bandura, A.: Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice Hall, Englewood Cliffs (1986)
35. Deery, H.A.: Hazard and risk perception among young novice drivers. J. Safety Res. **30**(4), 225–236 (1999)
36. McKenna, F.P., Crick, J.: Experience and expertise in hazard perception. In: Grayson, G.B., Lester, J.F. (eds.) Proceedings of the Behavioral Research in Road Safety, pp. 39–46. Transport and Road Research Laboratory, Crowthorne (1991)
37. Bandura, A.: Guide for constructing self-efficacy scales. In: Pajares, F., Urdan, T. (eds.) Self-Efficacy Beliefs of Adolescents, pp. 307–337. Information Age Publishing, Greenwich (2006)
38. Hatakka, M., Keskinen, E., Gregersen, N.P., Glad, A., Hernetkoski, K.: From control of the vehicle to personal self-control: broadening the perspectives to driver education. Transp. Res F Psychol. **5**, 201–215 (2002)
39. Giedd, J. N.: Adolescent brain maturation. In R.E. Tremblay, R.G. Barr, R. DeV., Peters, M. Bolvin (Eds.), Encyclopedia on early childhood development (pp. 1–5). Montreal: Center of Excellence for Early Childhood Development (2010)
40. Reyna, V.F., Farley, F.: Risk and rationality in adolescent decision making: implications for theory, practice and public policy. Psychol. Sci. Public Interest **7**, 1–44 (2006)
41. Stuart A., Ord, J.K., Arnold S.: Kendall's Advanced Theory of Statistics: Volume 2A–Classical Inference and the Linear Model, 6th edn. Edward Arnold Publishers Ltd., London (2009)
42. Groeger, J.A.: Understanding driving—applying cognitive psychology to a complex everyday task. Psychology Press Ltd., Hove, East Sussex (2000)

Identification of Criteria for Drivers' State Detection

Nicolas D. Herzberger(✉), Gudrun M.I. Voß, and Maximilian Schwalm

Institute for Automotive Engineering, RWTH Aachen University,
Steinbachstr. 7, 52074 Aachen, Germany
{herzberger,voss,schwalm}@ika.rwth-aachen.de

Abstract. Even though drivers may perform driving-unrelated activities during automated driving, they must be ready to resume the driving task dependent on the level of automation. With the ultimate aim of guaranteeing safe takeovers, automated driving systems at certain levels of automation need to assess drivers' readiness to take over preliminary to a transition based on driver monitoring systems. In order to optimize current methods of drivers' state detection, the present work tries to investigate the idea of a human-based driver observation. It is strived to find first answers to the question, whether and which benefits such an approach would offer, i.e. whether criteria identified by human observation can serve for drivers' state detection systems. In this context, it was attempted to identify and define possible levels of involvement into the driving task. Furthermore, a study was conducted to find first criteria based on human observation for the levels' detection that in the future might be examined using existing technical methods.

Keywords: Automated driving · Drivers' state · Drivers' state monitoring · Driving task involvement

1 Introduction and Theoretical Background

Due to the human innate desire of self-determined, safe and comfortable transportation as well as the ongoing technological advances, the automobile industry continuously strives to bring forth new features and systems. In this context, the emergence of automated driving turned to be a key issue in the past few years. A broad spectrum of systems – including for example parking assistance systems as well as traffic jam pilots – came onto the market and is available in series vehicles by now. In order to conceptualize these systems, different organizations, among which is the SAE, developed nomenclatures that structure the different types of systems according to their capabilities and prerequisites [1]. According to the SAE classification [1], there are five levels of automation as well as an additional level 0 that represents non-automated driving. Level 1 represents 'driver assistance' and signifies that the system carries out either lateral or longitudinal guidance while the driver is in charge of the other as well as that the driver needs to continuously monitor the driving environment and is the fallback option in case of system failure. Within the next level – level 2 – the system takes over both lateral and longitudinal guidance. Comparably to level 1, the driver still needs to monitor the driving environment and is the required fallback option. This changes profoundly in level 3.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_76

Here, the driver does not need to continuously track the driving task and is only required to take over driving in case of system failure after a set take over time (via a takeover request). From level 4 on, the driver is not needed as a fallback option anymore. While level 4 is restricted to specific applications, level 5 covers the complete driving task and might be understood as 'robot taxi'.

What gets clear from the above overview of automation levels is that at a certain point a critical shift occurs. This point can be identified between level 2 and level 3 that on the surface seem to be alike but that reveal crucial differences when inspecting them in detail. While the system executes the same tasks (longitudinal and lateral guidance), the requirements to the driver differ significantly. In the lower automation level 2, drivers have to monitor the driving task and therefore might be said to be in the loop of the driving task. In contrast, the higher automation level 3 does not pose this requirement. Drivers are allowed to leave the loop of the driving task and are permitted to engage themselves with other, non-driving related task. However, they still need to be able to take over the driving task if required.

In the literature, this critical shift often is discussed as 'unsafe valley' of automation [2, 3]. It describes the fact that drivers often misconceive a level 2 system's abilities by confusing it with a level 3 system. They tend to overestimate a level 2 system's abilities in the sense of assuming that they may leave the loop of the driving task. A potential takeover situation would be unsafe in consequence, since a level 2 system only foresees an immediate takeover (without takeover request). In contrast to a level 3 system, drivers are not given any takeover time to jump back into the loop of the driving task. Following, one might conclude that in case of an 'unsafe valley', drivers are in fact in a 'wrong' drivers' state for a safe driving mode transition.

2 Possible Drivers' States

This sketch of the 'unsafe valley' reveals that in the context of automated driving, drivers are not always able to handle the driving task without errors occurring. In consequence, takeovers from automated to non-automated driving will not be safe per se depending on the drivers' current abilities. Bearing this fact in mind, it seems reasonable that automated systems need to check drivers' state and takeover readiness before initiating the takeover itself. Only if drivers are sufficiently involved in the driving task, a takeover maneuver should be started.

Until now, drivers' state detection or measurement mostly is accomplished via driving parameters such as time to line crossing [4], standard deviation of speed, increase of steering wheel movements [5] or decrease/increase of the headway distance [6]. Furthermore, physiological measurements such as heart rate (pulse detection), brain activity (electroencephalography), eye lid closure and blinks (eyelid sensor) or skin conductance (electrodes) [7] are often used. These methods offer a great variety of possibilities for the comprehensive investigation of drivers' abilities and states.

In this context, the present work tries to find access to the question, which entities and values necessarily need to be tracked and evaluated for an adequate drivers' state detection. For this purpose, the present work pursed a three-fold approach: Constituting

that drivers can be differently involved into the driving task depending on the automation level, it seems advisable to define in a first place possible states or involvement levels in which drivers' might generally be in. Second, it seems important to define which level of involvement is necessary in each level of automation for safe vehicle guidance. This knowledge would be a prerequisite and benchmark for automated driving systems against which they could compare and evaluate the current drivers' state. Only if the target level of involvement is met, a transition between driving modes should be initiated. In order to perform a proper trade-off in this context, it would be helpful to narrowly define which criteria can be used for the assessment of drivers' involvement. Therefore, it is, thirdly, tried to find first answers how such criteria could look like. All of these three steps (definition of involvement levels, definition of minimal involvement requirements, and identification of first assessment criteria for the involvement level fully) are described within the next chapters.

2.1 Definition of Involvement Levels

As already introduced, a first important step for a comprehensive drivers' state detection constitutes the definition of possible states in which drivers' might generally be. Based on the above-described SAE levels of automation [1] a total of five drivers' states regarding the involvement in the driving task can be defined:

- *Fully (F):* The active performance of the entire driving task is required, i.e. longitudinal and lateral guidance as well as cognitive monitoring of the driving task.
- *Involvement 1 (I1):* The active performance of either longitudinal or lateral guidance as well as the cognitive monitoring of the driving task is required. Additionally, an immediate takeover readiness (without takeover request) is needed.
- *Involvement 2 (I2):* No active performance of the driving task is required. However, the cognitive monitoring of the driving task as well as an immediate takeover readiness (without takeover request) are needed.
- *Retrievable:* Neither an active performance of the driving task nor an active cognitive monitoring of the driving task are required. A time-set takeover readiness (with takeover request) is needed.
- *Non retrievable (NR):* Driver sleeps (or similar) or a transition from R to I2 is not possible.

2.2 Definition of Minimal Requirements for Drivers' Involvement

For the purpose of estimating drivers' readiness to perform transitions between different automations levels, automated systems need to have knowledge on the minimal requirements of drivers' involvement for each automation level. A first determination of these requirements is made within the current work. Figure 1 depicts a graphical representation of these minimal requirements of drivers' involvement per SAE automation level. Additionally, the differentiation between the three task layers of driving (navigation, guidance, stabilization) according to Donges [8, 9] is taken into account.

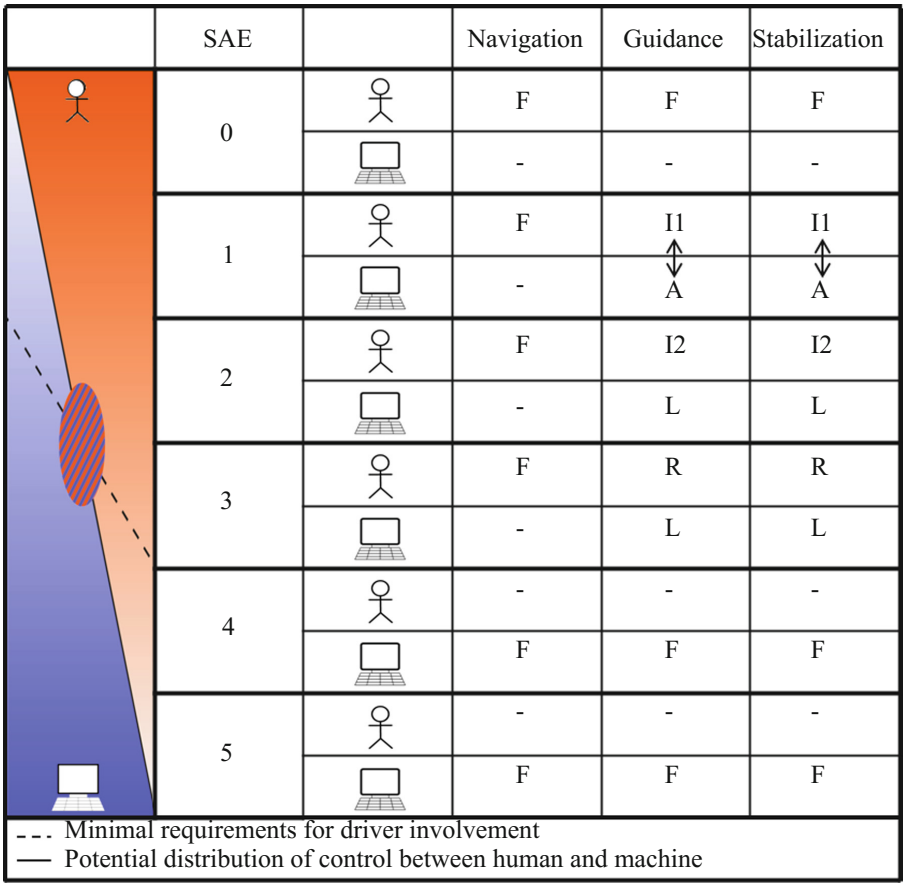


Fig. 1. Minimal requirements of drivers' involvement into the driving task per SAE level of automation.

Within the first column of this figure, the potential control distribution between human and machine (vehicle) is plotted. The orange-shaded area depicts tasks fulfilled by the driver, while the blue-shaded area represents all tasks accomplished by the vehicle. The dotted line within this column indicates the minimal requirements for driver involvement over the automation spectrum. As one can see, the control distribution and the required involvement do not run parallel. As it has been discussed in the forgone chapter, this non-linearity can lead to driver-sided misconceptions illustrated by the 'unsafe valley' of automation. The localization of this valley between SAE level 2 and 3 is plotted as a crosshatched oval in the first column. The three right-most columns of the figure include the minimal required states in the three layers of driving each for the driver and the system. While for the driver, the above-defined states are used, three additional system conditions were defined:

- *Assisted driving (A):* The vehicle is responsible for either longitudinal or lateral guidance
- *Lead driving (L):* The vehicle is responsible for both longitudinal and lateral guidance
- *Fully (F):* The vehicle is responsible for the active performance of the entire driving task.

3 Criteria for Drivers' State Detection

The forgone chapter offers a proposition on the minimal requirements for drivers' involvement depending on the level of automation. As it has already been mentioned, it seems indispensible that automated systems can evaluate drivers' current state and compare it to these target values. In the context of this assessment, it seems to be beneficial to have criteria that could be proofed by existing approaches as e.g. eye tracking. Therefore, the present work tries to provide a first idea of how these criteria, which are ultimately required for a reliable drivers' state detection, could look like.

In this context, a human-based observation is pursed that builds upon the assumption that a human observer is at least to a certain extent capable of estimating another human's current state based on overt observation criteria. An intuitive example that can illustrate this issue is the one a driving instructor, who is capable of estimating whether the learner driver is able to solve a driving situation or whether and to what extent an intervention into the driving task is necessary. The purpose of the current work is to verify whether the implementation of such an observation-based approach seems sensible and should be further pursued in the future.

3.1 Initial Experimental Study

In order to prove that human observers can assess drivers' involvement into the driving task, an initial experimental study comprising two parts (test drives and online study) was carried out.

Methods

As a first approximation to the topic, it was chosen to conduct the experiment in a non-automated driving mode, in which drivers are assumed to be fully involved (see Fig. 1). For reasons of replicability the test drives were video-recorded. In this way, they could be presented afterwards repeatedly to a larger sample within an online study.

The non-automated (SAE level 0) test drives were carried out on an 18-km-long route in the vicinity of Aachen, Germany, which mainly consisted of country-side roads as well as short segments of urban roads. In order to verify the drivers' attention for driving relevant information, i.e. their involvement into the driving task, they had to react to driving relevant stimuli in the environment (traffic signs) by pushing a button on the steering wheel as well as naming the indicated speed limit. Both the drivers and the driving situation were filmed during the test drive. The recordings were later used as video material for the adjacent online study. Figure 2 depicts an exemplary screenshot of a video sequence.

Fig. 2. Exemplary screenshot of a video from the online survey.

This online study served the purpose of investigating whether the assessment of a human observer is a measure of drivers' state and which criteria are used for this assessment. For this purpose the participants were shown the video sequences and were asked, if the drivers saw the driving-relevant information (the traffic sign). In addition, it was the participants' task to generate criteria that they used for their assessment. In total, each participant was presented a randomized set of four video sequences that could be repeated as often as desired. Because not every participant was presented the same four video sequences, the experiment was set up as a between-subject design.

Sample

Overall, $N = 20$ participants took part in the test drive in which the videos were recorded, $n = 7$ of them female and $n = 13$ of them male. The average age was 29 years ($SD = 8$). On average participants held their driver's license for 11 years ($SD = 7$) and the annual mileage as driver of a passenger car was on average 13,225 km ($SD = 13{,}176$).

The answers of $N = 233$ volunteers were included in the analysis of the online survey. Participants, who provided answers to less than 70% of the questions, were excluded from data analysis. $n = 137$ participants were male and $n = 96$ female. The average age was 32 ($SD = 12$). On average the participants held their driver's license for 14 years ($SD = 12$). The annual mileage as driver of a passenger car was on average 14,485 km ($SD = 13{,}500$) and the annual mileage as co-driver of a passenger car was on average 4,601 km ($SD = 7{,}492$).

Results

Over all the videos shown in the online study the correct assessment of the traffic sign recognition by the participants of the online study was on average 73.2%. Additionally to this rating on the recognition of driving-relevant information (traffic sign), the participants of the online study were asked to name the criteria, on which they based their assessment. The results of all instances in which a correct assessment was made are shown in Fig. 3.

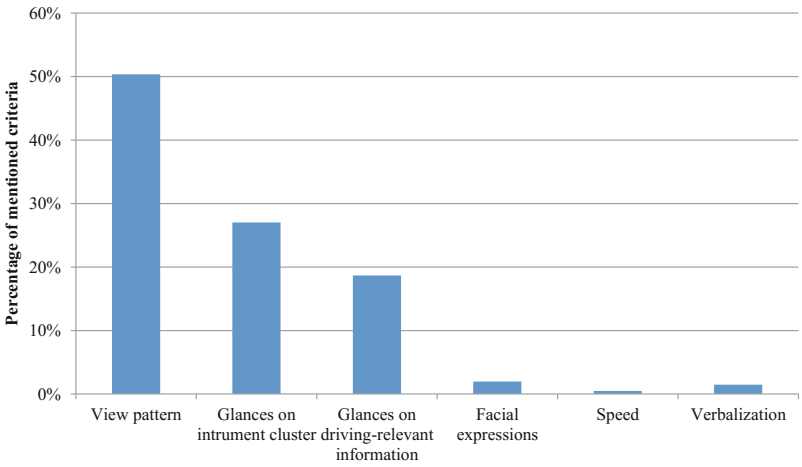


Fig. 3. Percentage of mentioned criteria of correct ratings by driver observation regarding sign recognition.

The cluster 'view pattern includes statements such as 'looks to the top right and then quickly down' or 'looks first to the right and left and then checks the current speed. The cluster 'glances on instrument cluster' includes statements such as 'driver looks regularly at speedometer'. The cluster 'glances on driving-relevant information' includes statements such as 'drivers eyes follow traffic signs and traffic', the cluster 'facial expressions' includes statements such as 'looks very concentrated', the cluster 'speed' includes statements such as 'speed is not adjusted' and the cluster 'verbalization' includes statements such as 'driver is talking'. The results of all instances in which an

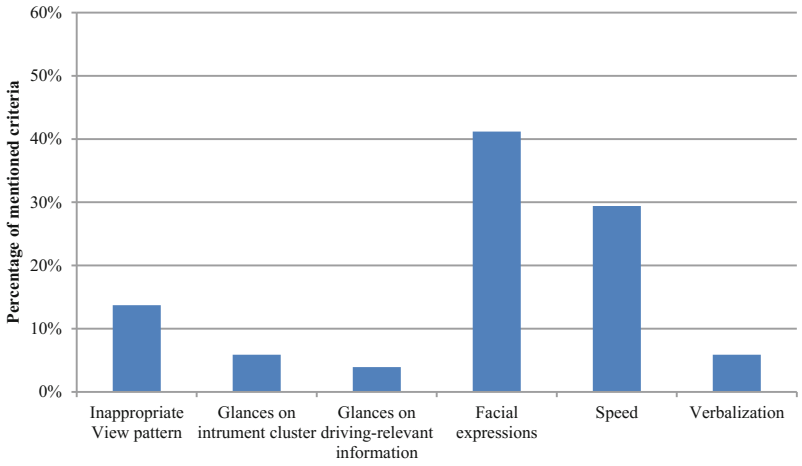


Fig. 4. Percentage of mentioned criteria of incorrect ratings by driver observation regarding sign recognition.

incorrect assessment was made are shown in Fig. 4. As one can see from these two figures, the frequencies of the mentioned criteria differs between correct and incorrect assessments.

4 Summary and Outlook

The present work's goal was a first determination of minimal requirements of drivers' involvement per SAE automation level. Therefore, five possible minimum levels of drivers' involvement were defined. They range from 'fully involved into the driving task' to 'not retrievable to the driving task'. Second, the definition of these states was used to define minimal requirements for drivers' involvement per automation level. These minimal requirements indicate which involvement level is needed in order to guarantee a safe driving performance. Third, it was discussed whether and to what extent a human-based observation method could provide any benefits to drivers' state detection. For this purpose, an initial study for the drivers' state 'fully' was conducted, in which was investigated if human observers are able to assess drivers' involvement in the driving task.

The results of the conducted online survey revealed that human observers are in principle able to infer from the observation of drivers on their attention on driving-relevant information to a fairly high amount (73.2%). Although it was shown that this assessment is possible for the involvement level 'fully', the hit rate was not perfect (100%). The named criteria show that a successful detection of drivers' involvement level 'fully' is especially based on gaze observation (view pattern, glances on instrument clusters, glances on driving-relevant information). However, it appears that not all observable criteria are suitable for the detection of drivers' involvement (e.g. facial expressions).

Overall, participants that correctly assessed the drivers' involvement mostly used view patterns as their basis of assessment. It seems promising to combine this criterion with the objective detection of the same. View patterns, especially patterns adapted to the situation (e.g. looking at the traffic sign and subsequently looking at the instrument cluster or following road users with glances) were very successful.

For this purpose, in a next step, gaze sequences could be analyzed and combined with object detection systems. The system could identify differences in the environmental objects detected by itself and the driver, and use the information for specific warning signals.

Furthermore, these criteria are until now mostly based on implicit behavioral components. For the purpose of developing drivers' state detection systems, it is therefore necessary to define and describe these criteria more explicitly in the future. The current research provided first insight into the drivers' involvement level 'fully' and could reveal that in general such a human-based drivers' state observation seems to be reasonable. In order to carry on this idea, it is indispensible to derive criteria for each potential drivers' state and to develop a measurement concept allowing to detect those criteria.

Acknowledgments. This project was funded within the Priority Programme "Cooperative Interacting Automobiles (CoInCar)" of the German Science Foundation DFG.

References

1. SAE International (2014). http://www.sae.org/misc/pdfs/automated_driving.pdf
2. Flemisch, F., Altendorf, E., Canpolat, Y., Weßel, G., Baltzer, M., Lopez, D., Herzberger, N., Voß, G., Schwalm, M., Schutte, P.: Uncanny and unsafe valley of assistance and automation: first sketch and application to vehicle automation. In: Schlick, C.M., Duckwitz, S., Flemisch, F., Frenz, M., Kuz, S., Mertens, A., Mütze-Niewöhner, S. (eds.) Advances in Ergonomics Design of Systems, Products and Processes. Springer, Cham (2016)
3. Flemisch, F., Altendorf, E., Canpolat, Y., Weßel, G., Baltzer, M., Rudolph, C., Lopez, D., Voß, G., Schwalm, M.: Arbeiten in komplexen Mensch-Automations-Systemen: Das unheimliche und unsichere Tal (uncanny & unsafe valley) der Automation am Beispiel der Fahrzeugautomatisierung. In: 62. GfA-Frühjahrskongress "Arbeit in komplexen Systemen – Digital, vernetzt, human?!", Aachen (2016)
4. Sandberg, D., Åkerstedt, T., Anund, A., Kecklund, G., Wahde, M.: Detecting driver sleepiness using optimized nonlinear combinations of sleepiness indicators. IEEE Trans. Intell. Transp. Syst. **12**(1), 97–108 (2011)
5. Dong, Y., Hu, Z., Uchimura, K., Murayama, N.: Driver inattention monitoring system for intelligent vehicles: a review. IEEE Trans. Intell. Transp. Syst. **12**(2), 596–614 (2011)
6. Bayly, M., Young, K.L., Regan, M.A.: Sources of distraction inside the vehicle and their effects on driving performance. In: Regan, M.A., Lee, J.D., Young, K.L. (eds.) Driver Distraction: Theory, Effects, and Mitigation, pp. 191–214. CRC Press, Boca Raton (2008)
7. Muzet A., Penzel T., Ktonas P., Virkkala T., Anderer P., Dorffner G.: List of Needs and Suggestions for Novel Sensors and Improved Methods for Physiological Data Recording. Deliverable 1.2.2; Sensation EU project no. 507231 (2006)
8. Donges, E.: Aspekte der aktiven Sicherheit bei der Führung von Personenkraftwagen. Automob-Ind 27 (1982)
9. Donges, E.: Fahrerverhaltensmodelle. In: Winner, H., Hakuli, S., Lotz, F., Singer, C. (eds.) Handbuch Fahrerassistenzsysteme. Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort (3. Auflage), pp. 17–26. Springer, Wiesbaden (2015)

A Real Case-Based Study Exploring Influence of Human Age and Gender on Drivers' Behavior and Traffic Safety

Nazha R. Ghadban[1], Galal M. Abdella[1], Khalifa N. Al-Khalifa[1,2], Abdel Magid Hamouda[2(✉)], and Khadija B. Abdur-Rouf[1]

[1] Qatar Transportation and Traffic Safety Center (QTTSC), Qatar University, Doha, Qatar
nghadban@qu.edu.qa
[2] Department of Mechanical and Industrial Engineering, Qatar University, Doha, Qatar
hamouda@qu.edu.qa

Abstract. The demand for understanding the functional relationship between changes in traffic safety level and human-related factors has increased remarkably due to a steady increase in population and traffic volume. Motor vehicle crashes can cause a serious impact on the victims, families, and the community resulting in instability of family and large socio-economic costs to society and healthcare service. Statistical studies of drivers' behavior as a function of some of the human characteristics, in particular, have become increasingly important for establishing an evidence-based framework through which transportation and healthcare authorities can develop innovative solutions to mitigate serious consequences of risky driving. This paper uses a real-world dataset to analyze the relationship between the human characteristics and both the behavior and risk-level. In addition to that, this paper conducts several analytical studies to investigate the extent to which the risky driving behavior may affect the severity of the crash and human injury.

Keywords: Traffic safety · Risky driving behavior · Road crashes · Human factors in traffic safety

1 Introduction

Driving has become an essential part of our day-to-day life. Humans are usually different in nature and have different personalities and behaviors, which can affect the way they drive. Although driving provides ease of movement for people, the different behaviors of the drivers can lead to unfavorable conditions or consequences. Transportation and traffic authorities worldwide have made several endeavors in the direction of mitigating the risky driving behavior through either launching new or modifying existing road traffic regulations. The main purpose of these traffic regulations was to reduce traffic crashes and their severity.

Several publications have been made available over the last years in which the influence of human features such as gender and age on the risky driving behavior was

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_77

investigated. Our investigations in literature have shown that most of these works are essentially based on the police traffic reports as the main source of the research data. However, some of these important research works are reported in this section with detailed description of their data gathering techniques.

Dejoy [1] used a three-part questionnaire to study gender differences in traffic accident risk perception. The first part of the questionnaire was designed for the participants to judge themselves in certain driving characteristics; the second part asked them to rate 15 risky driving behaviors, and the last part collected demographic data. Harré et al. [2] studied gender differences in the driving behaviors and the attitudes of youth. In their study authors relied on the questionnaire as the main tool to collect data relevant to demographics-features, driving status, traffic laws, and crash involvement. Moreover, Özkan and Lajunen [3] used a different questionnaire to understand the effect of gender on driving skills and accident involvement. Authors applied the Bem Sex Role Inventory test and the Driver Skill Inventory test as well as some personal questions. Furthermore, a research published by Jiménez-Mejías et al. [4] developed a questionnaire to carry out an analysis on gender differences. The questionnaire included a demographic section similar to the others. However, the rest of the sections were dissimilar; they focused on collecting data about the frequency of driving, seatbelt usage, and some other selected driving behaviors.

Rhodes and Pivik [5] distributed a questionnaire to establish a clear relationship between age and gender, and twelve driving behaviors. The survey used a scale of five to rate the frequency of displaying twelve different driving behaviors. Moreover, the research also collected responses on the perception of people about the risks involved in doing these risky driving behaviors. Classen et al. [6] tested gender differences in driving behaviors for older drivers. The study was conducted using a questionnaire and other methods such as clinical tests, visual cognition, and motor performance test. The questionnaire collected demographic data, history of the driver, and data about certain driving behaviors.

Another research carried out by Cordellieri et al. [7] studied three aspects as a function of gender and these aspects are: dealing with road safety rules, driving behaviors, and risk perception of male and female drivers. A different questionnaire was designed that collected demographic information and tested each aspect using a scale point. Tao et al. [8] focused on Chinese drivers in investigating three areas: personal characteristics, driving behaviors, and risk of accidents. The research tested the effect of gender and other factors on these areas. A similar research was carried out by Barr et al. [9] who focused on teenagers to study driving behaviors as a function of gender. The research collected responses from male and female teenagers where the teenagers conveyed information about their own driving behaviors.

In addition, some other researchers have analyzed gender differences by collecting data from statistical authorities and police reports or from other researchers who already conducted studies in this field. Welsh and Lenard [10] conducted a research in the UK in which the data used were collected from police records and accident data in the UK. The main aim of the study was to investigate gender differences in collision and injury risk. Lonczak et al. [11] established a study in Washington that endeavored to predict risky and angry driving as a function of gender. In this study, the data were obtained from previous research undertaken in Washington by the Traffic Safety Commission.

In this paper, we aim to explore the gender differences in driving behaviors as well as extend the study to understand age differences in driving behaviors through a case study. Moreover, the correlation of driving behavior at the time crash and the corresponding severity of damage is also studied. Besides, data used was collected from a sample comprising of Qatar University community through an online survey.

The rest of this paper is organized as follows: Sect. 2 explains the methodology used to collect the data. Section 3 highlights the findings of the survey results. In Sect. 4, we discuss in details the results of our analysis. Finally, we summarize the implications of the study and draw conclusions in Sect. 5.

2 Proposed Methodology

We dedicate this section to introduce the methodology followed to achieve the work objectives. In this paper, we designed a questionnaire consisting of two different main sections. The first section is about introducing the survey objectives and identifying the targeted population. The second section of the questionnaire is dedicated to collect data relevant to age, gender, driving behaviors and road crashes history of participants.

The "driving behaviors" section explores how drivers frequently behave. The participants use a scale of 1 to 5 to rate each behavior in which 5 = always, 4 = often, 3 = sometimes, 2 = rarely, and 1 = never, and zero = not applicable. In this study, several behaviors were selected and rated including:

- Not wearing the car seatbelt (DB1)
- Exceeding the speed limit (DB2)
- Driving in the emergency lane (DB3)
- Overtaking from the right side (DB4)
- Changing lanes without using the signs (DB5)
- Not keeping safe distance from the front vehicle (DB6)
- Eating or drinking while driving (DB7)
- Making or accepting phone calls (DB8)
- Reading or sending emails and text messages (DB9)

Furthermore, the "road crashes" section collected information to understand the relationship between car crashes and driving behaviors. The questions asked in this section included the number of crashes that participants were involved in within the last three years, the type of injury that occurred in the two most recent crashes, the type of damage that occurred to the car in the two most recent crashes, and the person who caused the crash. In addition, this section asked the drivers to report the driving behaviors that they were doing when the crash occurred.

3 Survey Results Description

This section describes the data collected for the study as part of a "Driving Behaviors" survey conducted at Qatar University during the month of February of the spring 2017 term. E-mails with the link to participate in a web-based survey were sent to 14,000+

students and 2,000+ faculty and staff. The total number of participants who completed the questionnaire was 500, that is, the sample size was 3.13%. The sample was chosen to be from Qatar University since it represented the diversity of Qatar with regard to gender, age, occupation, and driver characteristics.

3.1 Demographics of Participants

Female (51.89%) respondents for the survey were slightly more than males (48.11%). Besides, the majority of the participants were of the age group 18–22 (43.9%) which reflected the occupation (student) of more than half of the participants (55.7%). Additionally, the driving experience for most of the participants (38%) was less than 2 years. A major percentage of the sample (45.1%) said they drove 1–2 h each day, which reveals both their dependency and affinity towards driving. It is also worth mentioning that 99.5% of the participants graduated from high school. For further detailed information, see Table 1.

Table 1. Summary of selected variables explaining socio-economic demographics of individual participants

Gender	Male	48.1%
	Female	51.9%
Age (years)	18–22	43.9%
	23–30	24.9%
	31–45	20.9%
	46–65	10.3%
Occupation	Student	55.7%
	Working full-time	39.0%
	Temporarily unemployed	2.8%
	Other	2.6%
Driving experience in Qatar (years)	<2	38.0%
	2–5	33.4%
	6–10	14.1%
	>10	14.5%
Hours driven per day	<1	16.3%
	1–2	45.1%
	3–4	32.6%
	>5	6.0%

3.2 Driving Behaviors Analysis

Participants rated how often they practiced the driving behaviors mentioned in Sect. 2. In all cases, the majority of them said they never or rarely practiced risky behaviors such as not wearing car seatbelt. Around 23.0%–29.5% of the participants admitted that

they would sometimes make/accept calls using mobile, not keep a safe distance from the front vehicle, move from one lane to another without using signs, eat/drink while driving or exceed the speed limit. 14.7% reported often using mobile phone followed by 14.5% exceeding the speed limit followed by 11.6% changing lanes without a sign and 10.9% not keeping a safe distance. See Fig. 1 for more details.

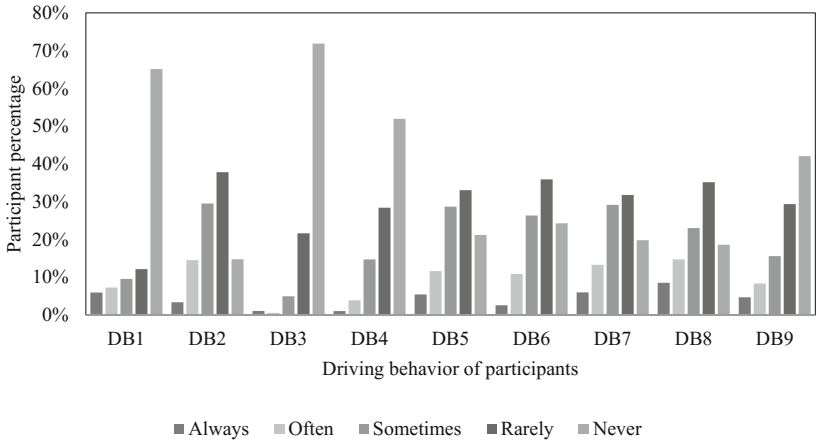


Fig. 1. Distribution of participants' response to driving behavior practices

In addition, around 10.9%–13.3% stated they often would not keep a safe distance, eat/drink, or change lanes without a sign. Among the small percentage who said they always practiced a behavior, the highest percentage was using a mobile phone (8.5%). Small percentages (3.4%–6%) of them would always exceed the speed limit, read/send emails/text messages, change lanes without a sign, not wear a seatbelt, or eat/drink.

3.3 Crash History Analysis

The participants reported how many crashes they were involved in the last three years. Participants involved in a crash identified their driving behavior at the time of the crash. Moreover, they also rated the severity of the crash in terms of injury and car damage.

Frequency of Crash. Out of 500 participants, 34.2% reported being involved in at least one crash in the last three years. Around 37.1% of the male participants said they were in an accident; while a lower percentage (31.5%) of the female participants said they had a crash. Similarly, the males had a higher mean number of crashes (2.39) than the females (1.81). On the other hand, the highest percentage (49.2%) of crashes reported was by the age group 23–30. However, the age group 18 – 22 had the highest mean crashes (2.25) followed by the age group 46–65 (2.09). See Table 2 for the detailed summary.

Table 2. Summary of participants involved in crash in the last three years

		≥ 1 crash	Mean crashes	Std. dev.
Gender	Male	37.1%	2.39	3.15
	Female	31.5%	1.81	0.99
Age (years)	18–22	31.4%	2.25	3.22
	23–30	49.2%	2.00	1.41
	31–45	28.8%	2.03	1.99
	46–65	21.2%	2.09	1.51

Driving Behavior During Crash. Moreover, the majority of the participants (36.5%) involved in their most recent crash reported not keeping a safe distance from the front vehicle at the time of the crash (see Fig. 2). 17.5% of them said they were not wearing the car seatbelt. Around 8%–9.5% stated that they were exceeding the speed limit, changing lanes without using signs, or reading/sending emails and text messages. Also, about 7% were either making/accepting phone calls or overtaking from the right side. Whereas, small percentages (1.5%–3.5%) of them were either driving on the emergency lane or eating/drinking while driving.

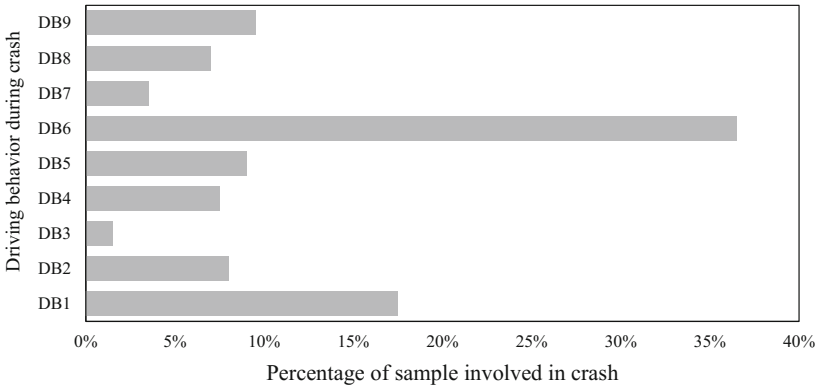


Fig. 2. Driving behavior of participants at the time of their most recent crash

Severity of Crash. The majority of the participants (94%) involved in the crash reported no injury to the driver and the passengers (see Table 3). Only 6% reported slight injury; whereas, none of them reported any severe or fatal injury. Likewise, high car damage was incurred by only a small percentage (6.9%) compared to low car damage sustained by the majority of them.

Table 3. Summary of severity of crashes

Injury	No injury	94.0%
	Slight injury	6.0%
Car damage	Low	72.7%
	Moderate	20.4%
	High	6.9%

4 Results and Discussion

4.1 Driving Behavior Versus Gender

This section uses a non-parametric test called Wilcoxon test. The purpose is to test the difference in terms of seatbelt usage and speeding between two groups: male and female. In this study, we examined the null hypothesis H_0: The two samples follow the same distribution versus H_1: The distributions of the two samples are different. However, the two hypotheses were tested using the Minitab software at level of significance $\alpha = 0.05$. The next sections were dedicated to understand the relationship between the human features, age and gender, and the rate of how frequently the the participants practice the corresponding list of risky driving behaviors.

Not Wearing Car Seatbelt. After applying the Wilcoxon test, the result showed that the p-value is <0.001. Since the computed p-value is lower than the significance level $\alpha = 0.05$, one should reject the null hypothesis H_0, and accept the alternative hypothesis H_1. This shows that there is a significant difference between males and females in terms of car seatbelt usage. Moreover, the mean and standard deviation of the two groups were calculated in order to extract which group wears the car seatbelt more frequent. The male group has higher mean showing that they wear the seatbelt less often than females. The values are shown in Table 4.

Table 4. Seatbelt usage according to gender

Gender	Mean	Std. deviation
Male	2.728	1.062
Female	2.429	0.872

Exceeding the Speed Limit. The results of the Wilcoxon showed that the p-value is <0.001. This reveals that males and females significantly behave differently in terms of exceeding the speed limit. Here as well, the mean of the two groups show that males have higher value illustrating that males exceed the speed limit more often than females. See Table 5 for details.

Table 5. Speeding according to gender

Gender	Mean	Std. deviation
Male	1.870	1.265
Female	1.690	1.190

4.2 Driving Behavior Versus Age

This section uses a test called Kruskal-Wallis test, which is also a non-parametric test. The test will verify if the four age groups behave similarly or differently in terms of seatbelt usage and exceeding the speed limit. We examined the null hypothesis H_0: The two samples follow the same distribution versus H_1: The distributions of the two samples are different. The two hypotheses were examined using the Minitab software at level of significance $\alpha = 0.05$.

Not Wearing Car Seatbelt. The results of the Kruskal-Wallis test provided a *p*-value that is <0.001. According to the level of significance, the null hypothesis will be rejected. This concludes that there is a significant difference between the four age groups in terms of seatbelt usage. From the mean calculations, it can be observed that the age group between 23 and 30 years do not wear the car seatbelt more often than the other age groups. Table 6 shows each group with their corresponding mean and standard deviation.

Table 6. Seatbelt usage according to age

Age	Mean	Std. deviation
18–22	1.642	1.188
23–30	2.127	1.376
31–45	1.870	1.207
46–65	1.237	0.820

Exceeding the Speed Limit. Likewise, the Kruskal-Wallis test gives a *p*-value that is <0.001. It is noted that the difference between the four age groups is significant according to the significance level. Table 7 shows the calculated mean and standard deviation for all age groups. The age group that is between 23 and 30 has the highest mean showing that this age group usually exceeds the speed limits more than the other groups.

Table 7. Speeding according to age

Age	Mean	Std. deviation
18–22	2.617	1.046
23–30	2.804	1.005
31–45	2.403	0.936
46–65	1.921	0.712

4.3 Correlation Analysis: Driving Behavior Versus Injury Severity

Pearson Correlation is used to identify the interrelationship between the severity of an injury caused by a car crash and the different driving behaviors that existed during the crash. The results of the correlation test are shown in Table 8. However, the results obtained from the correlation test show that DB2 and DB1 have the highest values. This proves that seatbelt usage and speeding are the two behaviors that mostly affect the type of injury that can occur to a driver.

Table 8. Correlation between injury severity and driving behaviors

Variables	DB1	DB2	DB3	DB4	DB5	DB6	DB7	DB8	DB9
Injury severity	0.064	0.097	−0.035	0.012	−0.004	0.045	−0.054	0.018	0.075

It is also clear from the results that although DB2 and DB1 have the highest values; their values are still small and not close to 1. This is according to the responses obtained where all crashes were either a crash with no injury or a crash with a slight injury.

4.4 Correlation Analysis: Driving Behavior Versus Car Damage Severity

The interrelationship between the severity of car damage caused by a car crash and the driving behavior that was done during the crash was also tested using *Pearson Correlation*. In this analysis, DB1 was removed because obviously, seatbelt usage has no relation with the car damage severity. Meanwhile, DB2 shows the highest value among all other behaviors. The value represents that the higher the speed, the more severe the car damage. See Table 9 for the summary of the correlation.

Table 9. Correlation between car damage severity and driving behaviors

Variables	DB2	DB3	DB4	DB5	DB6	DB7	DB8	DB9
Car damage severity	0.233	−0.078	−0.063	0.032	−0.139	0.034	0.049	0.086

5 Implications and Conclusion

This study investigates the effect of human characteristics on several driving behaviors that are familiar to most drivers. The analysis of the survey results mainly focused on two famous driving behaviors, which are seatbelt usage and speeding.

The survey results suggest that majority of the drivers in the Qatar University sample were young (18–22 years), had less than 2 years of driving experience in Qatar, and drove 1–2 h per day. Although the sample had almost equal number of males and females, males were found to be reporting more accidents than females with the age group 23–30 being the most vulnerable. On a positive note, the majority of them were involved in no injury and low car damage crashes. In addition, most of them were

rarely practicing unsafe driving behaviors. However, at the time of the reported crashes, most of them were either driving too closely to the vehicle in front, not wearing Seatbelts, or reading/sending text messages/emails. In addition, a small percentage of the sample were found to be always or often involved in unsafe driving practices.

The study concluded that male drivers speed more and wear seatbelt less than female drivers. The other part of the study focused on the effect of several driving behaviors on the injury severity and the car damage severity caused by a car crash. The results analysis revealed that seatbelt usage and speeding have an interrelationship with the injury severity among all other behaviors. Meanwhile, the severity of car damage had the highest interrelationship with speeding.

References

1. DeJoy, D.: Gender differences in traffic accident risk perception. Proc. Hum. Factors Ergon. Soc. Annu. Meet. **34**(14), 1032–1036 (1990). http://dx.doi.org/10.1177/154193129003401416
2. Harré, N., Field, J., Kirkwood, B.: Gender differences and areas of common concern in the driving behaviors and attitudes of adolescents. J. Safe. Res. **27**(3), 163–173 (1996). http://dx.doi.org/10.1016/0022-4375(96)00013-8
3. Özkan, T., Lajunen, T.: What causes the differences in driving between young men and women? The effects of gender roles and sex on young drivers' driving behaviour and self-assessment of skills. Transp. Res. Part F Traffic Psychol. Behav. **9**(4), 269–277 (2006). http://dx.doi.org/10.1016/j.trf.2006.01.005
4. Jiménez-Mejías, E., Prieto, C., Martínez-Ruiz, V., Castillo, J., Lardelli-Claret, P., Jiménez-Moleón, J.: Gender-related differences in distances traveled, driving behaviour and traffic accidents among university students. Transp. Res. Part F Traffic Psychol. Behav. **27**, 81–89 (2014). http://dx.doi.org/10.1016/j.trf.2014.09.008
5. Rhodes, N., Pivik, K.: Age and gender differences in risky driving: the roles of positive affect and risk perception. Accid. Anal. Prev. **43**(3), 923–931 (2011). http://dx.doi.org/10.1016/j.aap.2010.11.015
6. Classen, S., Wang, Y., Crizzle, A., Winter, S., Lanford, D.: Gender differences among older drivers in a comprehensive driving evaluation. Accid. Anal. Prev. **61**, 146–152 (2013). http://dx.doi.org/10.1016/j.aap.2012.10.010
7. Cordellieri, P., Baralla, F., Ferlazzo, F., Sgalla, R., Piccardi, L., Giannini, A.: Gender effects in young road users on road safety attitudes, behaviors and risk perception. Front. Psychol. **7** (2016). http://dx.doi.org/10.3389/fpsyg.2016.01412
8. Tao, D., Zhang, R., Qu, X.: The role of personality traits and driving experience in self-reported risky driving behaviors and accident risk among Chinese drivers. Accid. Anal. Prev. **99**, 228–235 (2017). http://dx.doi.org/10.1016/j.aap.2016.12.009
9. Barr, G., Kane, K., Barraco, R., Rayburg, T., Demers, L., Kraus, C., et al.: Gender differences in perceptions and self-reported driving behaviors among teenagers. J. Emergency Med. **48**(3), 366.e3–370.e3 (2015). http://dx.doi.org/10.1016/j.jemermed.2014.09.055
10. Welsh, R., Lenard, J.: Male and female car drivers - difference in collision and injury risks. In: Proceedings of the 45th Annual AAAM Conference, 24–26 September, Texas (2001)
11. Lonczak, H., Neighbors, C., Donovan, D.: Predicting risky and angry driving as a function of gender. Accid. Anal. Prev. **39**(3), 536–545 (2007). http://dx.doi.org/10.1016/j.aap.2006.09.010

Verification of Installed Position of LED Block Equipped with Projections to Indicate Travel Direction

Tomoyuki Inagaki[1], Hideaki Nagahama[2], Norihiro Ikeda[3], Kazuya Takahashi[4], Kiyohito Takeuchi[3], Hiroshi Ogino[3], Katsuya Sato[2], Sin-Ichi Ito[2], and Shoichiro Fujisawa[2(✉)]

[1] Nipon University, 7-24-1 Narashinodai, Funabashi, Chiba 274-8501, Japan
inagaki.tomoyuki@nihon-u.ac.jp
[2] Institute of Technology and Science, Tokushima University, 2-1 Minamijosan-Jima, Tokushima, Tokushima 770-8506, Japan
{c501202107, katsuyas, s.ito, sfujisawa}@tokushima-u.ac.jp
[3] Kictec Inc., 150 Umegaoka, Agui, Chita, Aichi 470-2295, Japan
{ikeda, ki_takeuchi, h_ogino}@kictec.co.jp
[4] General Support Co. for the Visually Impaired, 2-37-10 Kamiogi, Suginami, Tokyo 167-0043, Japan
saraseina@y3.dion.ne.jp

Abstract. A crosswalk entrance is on the boundary of a roadway and is a dangerous area for visually impaired persons. Tactile walking surface indicators (TWSIs) are installed on the road to support independent travel for blind and low-vision individuals. We developed an LED block equipped with projections to strengthen support at the crosswalk entrance. Blind individuals can sense thiss block through the soles of their shoes. Results indicated that the LED block equipped with projections to locate travel direction effectively supports crossing. The effectiveness of this block is determined using measurement and questionnaire results.

Keywords: LED block · Blind person · Crosswalk · Orientation · Veering distance · Tactile walking surface indicators

1 Introduction

The visually impaired includes blind persons and persons with low vision. When blind individuals walk alone, information on the walking surface is acquired through the soles of their shoes. When low-vision individuals walk alone, information on the walking surface is acquired through the use of residual vision. To obtain information for guidance, these pedestrians use tactual, acoustic, and visual information available from the natural and developed environments [1–10]. Crossing an intersection is an especially risky task. It is thus important to be able to acquire the crossing direction at the crosswalk entrance. Therefore, we developed an LED luminescence block with direction indication as support equipment to use at crosswalks with high risk of an

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_78

accident [11, 12]. Low-vision individuals can recognize the crosswalk entrance by detecting the light of the LED luminescence block. Blind individuals can acquire the direction of the crossing by sensing the projection of the surface of the LED luminescence block through the soles of their shoes. The developed LED luminescence block equipped with projections indicating travel direction is movement-support equipment that guides visually impaired persons to the crosswalk. The present study seeks to determine whether the developed LED block is effective for blind and low-vision persons. Two parallel lines project from the surface of this LED block. These two lines are orthogonal to the direction of the crossing. Blind persons can turn their body in the direction of the crossing by sensing the two lines through the soles of their shoes. This study verified the position for installing the crosswalk entrance of the LED luminescence block equipped with projections to guide the travel of blind pedestrians. The visually impaired person is guided to the crosswalk entrance by the guiding patterns. After the visually impaired person enters the guiding patterns, he/she encounters attention patterns in the crosswalk entrance. Two locations are considered for installing the developed LED block. One is on a connected boundary of the guiding pattern and the attention pattern. The other is on both sides of the boundary of the guiding pattern and the attention pattern. A crosswalk entrance was simulated in the University of Tokushima laboratory, and tactile walking surface indicators (TWSIs) were constructed. The experiment confirmed the difference in discovery times for the two kinds of LED blocks. The experiment result suggested that the walking time (discovery time) is shorter when the LED blocks are constructed on both sides of the guiding pattern than when the LED blocks are constructed on the boundary of the guiding pattern and the attention pattern. The subjects were also asked to complete a questionnaire. The questionnaire results indicated that the discovery of both sides is easier. The walking method was analyzed, and the findings are presented below.

2 Experiment Conditions

2.1 LED Block with Direction Location

Figure 1 presents an outline of the LED block with direction location used in the experiment. The LED block with direction location is travel support equipment that aims to guide the visually impaired person to the crosswalk exit. Two main lines (two bars) project from the surface of the block, and the LED is built into it. The direction location block is constructed at the crosswalk entrance so that projection of the line becomes vertical in the direction of the crossing. Blind persons confirm the indicated direction by the projection of two main lines of the block at the back of the foot, and they locate the crossed correct invasion angle. Low-vision persons discover the crosswalk exit by LED at the early stage, and follow the direction of the crossing.

2.2 Experiment Method

A crosswalk entrance was simulated in the University of Tokushima laboratory, and a block for guiding visually handicapped persons was constructed. The direction location

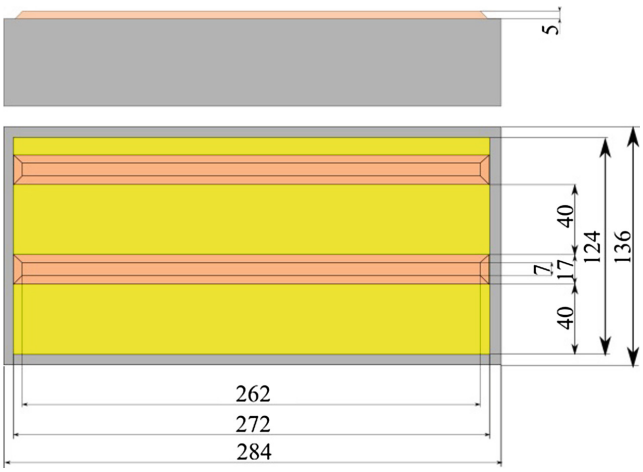


Fig. 1. Spacing and dimensions of LED block with direction location.

block was constructed at two kinds of positions of Installation B that Installation A and the guidance block that had been constructed at the center of the crosswalk entrance constructed right and left and it experimented. Figure 2 indicates the arrangement of each pattern. The experiment was conducted with the cooperation of 20 blind subjects (10 males and 10 females). The average age of the subjects was 62.7 years, and the standard deviation was 9.0 years.

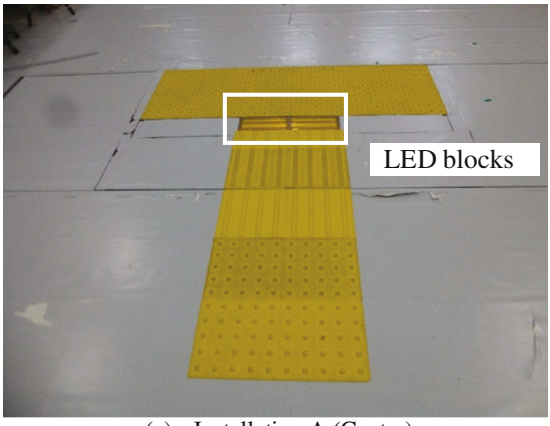
2.3 Experiment Procedure

The starting point of the experiment is an attention pattern at the turning point of a guiding pattern in the pavement and a guiding pattern to guide a person to the crosswalk entrance.

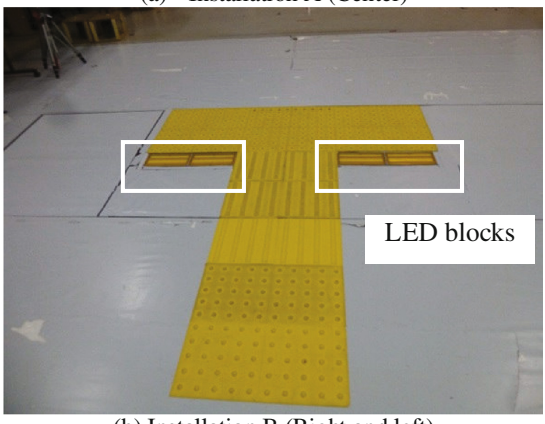
The experiment was conducted as follows.

1. The experimenter guides the subject to the starting point.
2. After walking according to the guiding pattern and discovering the direction location block at the back of the foot, the subject discovers the direction of the crossing.
3. The experimenter records the walking time from the starting point to the direction location block and the angle in the direction of the subject's travel.
4. Steps 1 to 3 are repeated ten times.
5. The subject answers questions about the experiment after the experiment.

This experiment focused on only the sense in the back of the foot, so use of a white cane was prohibited. Therefore, no obstacle other than the guiding pattern affected the subject's walking in the experiment environment. The experimenter always remained close to the subject to prevent the subject from falling and situation outside assumption.



(a) Installation A (Center)



(b) Installation B (Right and left)

Fig. 2. Installation of LED blocks with direction location.

3 Experiment Result

Two subjects assumed that the direction location block discovery would be difficult, and they participated in only the questionnaire response in Installation A. Figure 3 indicates the average walking time and the standard deviation obtained in the experiment. Results confirmed that the average walking time for Installation B was shorter than that for Installation A. Figure 4 plots the distribution of the angle in the direction of travel after the direction of the crossing was discovered. The ratio of both installations and 0° was high. Moreover, the ratio of 0° of Installation B was higher than that of Installation A. After the experiment had ended, participants responded to a questionnaire. Figure 5 presents the results of the questionnaire.

The questionnaire made it answer by three stage evaluation, and answered the reason freely. Results indicated that Installation A ranked lower than Installation B in ease of finding the block and confirming the two bars of the block.

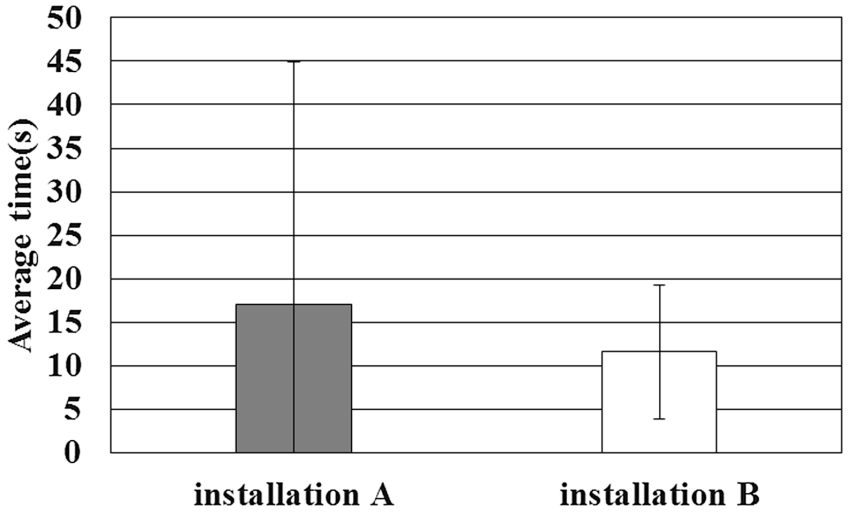


Fig. 3. Average walking time.

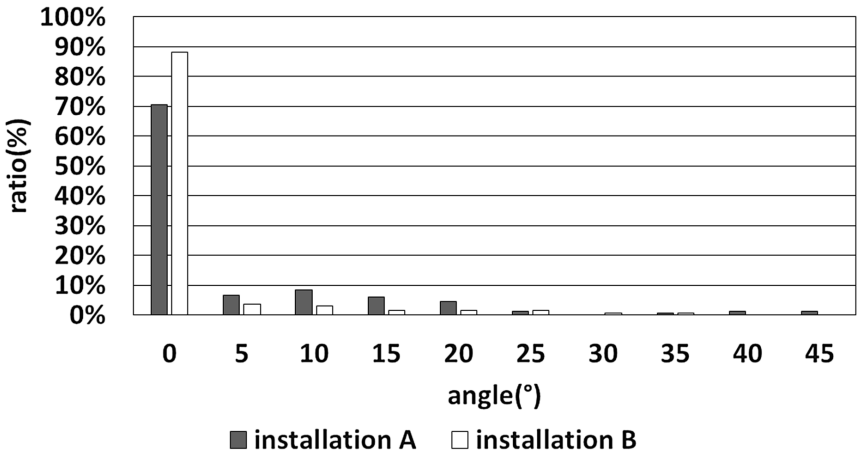
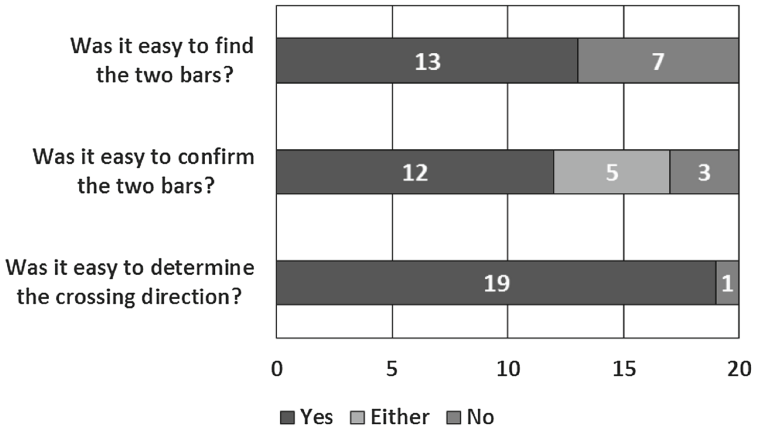


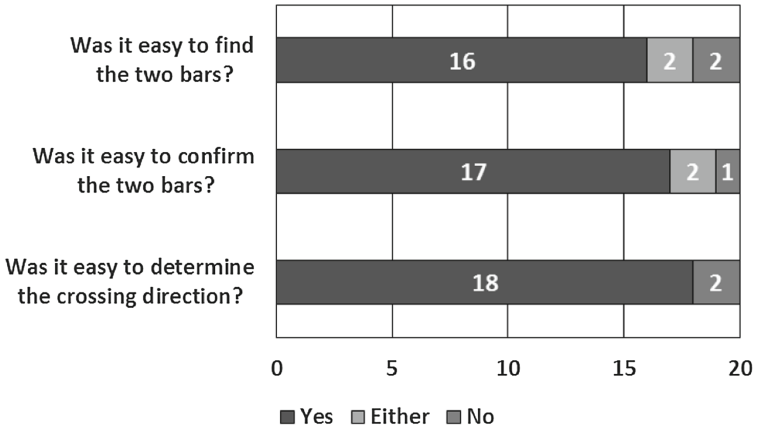
Fig. 4. Distribution of angle in the direction of travel.

4 Considerations

The experiment results indicate that the walking time for Installation B (block constructed next to the guiding pattern) is shorter than that for Installation A (block constructed in the center). It is understood to discover Installation B easily from the item of the easiness of the direction location block of the questionnaire to discover. The subject often passed the direction location block when it was constructed at the center. It is assumed that the direction location block was not confirmed easily because the guiding pattern was behind the direction location block. About 60% of the subjects in



(a) Installation A



(b) Installation B

Fig. 5. Responses to questions after the experiment.

the questionnaire responded that the direction location block of Installation A was found easily. As a result, when the direction location block is constructed at the center, the possibility of taking time in the discovery is suggested. However, the opinion of finding Installation A arrangement easily was stated to the subject who shuffled along the TWSIs. The opinion that it was easy to find the direction location block when stepping on the TWSIs with only one foot in Installation B was stated. It is thus assumed that the evaluation depends on the subject's method of walking. From the experiment results, the angle in the direction of travel becomes 0° in about 70% in Installation A and 90% in Installation B. The ease of the questionnaire in two main lines to confirm was answered that 60% was confirmed easily in Installation A. It was answered that 80% was confirmed easily in Installation B. Thus, direction location is higher for Installation B than for Installation A. In Installation A, the guiding pattern is

constructed behind the direction location block. It is thought that two bars were not discovered easily in the back of the foot because of this. It is thought that the ratio to which the direction of travel shifted from the direction of the crossing because it does not confirm two bars easily increased. Based on these results, it is recommended that the direction location block be constructed to the right and left of the guiding pattern at the crosswalk entrance.

5 Conclusions

This study aimed to determine the effective construction position of an LED luminescence block with direction location at a crosswalk entrance. The experiment was conducted in the University of Tokushima laboratory. The participants were blind persons who performed in a walking experiment. Results confirmed that the average walking time was shorter when the block was constructed at the center of the crosswalk entrance than when it was constructed to the right and left of the guiding pattern, and direction location performance was high. It was able to be confirmed that the block discovery was difficult though the direction location performance of the direction location block was suggested even when constructing it at the center of the crosswalk. Future study will seek to determine whether the use of a white cane affects the discovery of the direction location block. Moreover, the difference of the visibility by the position where the direction location block is constructed is scheduled to be verified to a low vision.

References

1. Loomis, J.M., Golledge, R.G., Klatzky, R.L.: GPS-based navigation systems for the visually impaired. In: Barfield, W., Caudell, T. (eds.) Fundamentals of Wearable Computers and Augmented Reality, pp. 429–446. Lawrence Erlbaum Associates, Mahwah (2001)
2. Ikeda Y., Mori, N.: Research on ITS for pedestrians. In: Proceedings of Tranced 2001, pp. 106–112 (2001)
3. Kooijman, A.C., Uyar, M.T.: Walking speed of visually impaired people with two talking electronic travel systems. Visual Impair. Res. **2**(2), 81–93 (2000)
4. Fanucci, L., Roncella, R., Iacopetti, F., Donati, M., Calabro, A., Leporini, B., Santoro, C.: Improving mobility of pedestrian visually-impaired users. Assist. Technol. Res. Ser. **29**, 595–603 (2011)
5. Loomis, J.M., Philbeck, J.M., Zahorik, P.: Dissociation between location and shape in visual space. J. Exp. Psychol **28**(5), 1202–1212 (2002)
6. Loomis, J.M., Kbapp, J.M.: Visual perception of egocentric distance in real and virtual environments. In: Hettinger, L.J., Haas, M.W. (eds.) Virtual and Adaptive Environments, pp. 21–46 (2003)
7. Kelly, J.W., Loomis, J.M., Beall, A.C.: Judgments of steering without course information. Perception **33**, 443–454 (2004)
8. Matsubara, K., Ito, S., Sato, K., Booka, M., Hidehisa, O., Fujisawa, S.: Difference in road surface recognition by disease and grades of the low vision. Assist. Technol. Res. Ser. **33**, 988–993 (2013)

9. Mitani, S., Yoshida, T., Minato, H., Fujisawa, S., Sueda, O.: Measurement of luminance contrast sensitivity of persons with low visual capability in order to secure the visibility of tactile walking surface indicators. Assist. Technol. Res. Ser. **20**, 326–330 (2007)
10. Mitani, S., Hamada, T., Fujisawa, S., Sueda, O., Tauchi, M.: Measurement of visbility of TWSIs perceived by LVs. Assist. Technol. Res. Ser. **29**, 618–625 (2011)
11. Ikeda, N., Takahashi, K., Yamamoto, G., Kimura, Y., Ito, S., Sato, K., Fujisawa, S.: Verification of LED blocks used at crosswalk entrances for persons with visually impairment. Assist. Technol. Res. Ser. **33**, 982–987 (2013)
12. Ikeda, N., Takahashi, K., Inagaki, T., Sato, K., Ito, S., Seiyama, M., Fujisawa, S.: Emitting LED block at crosswalk entrance for visually impaired persons. Proc. Manuf. **3**, 3774–3778 (2015)

Do End Users Really Have a Place in the Design Arena When Safe Design Is Critical?

Elise Crawford(✉), Yvonne Toft, Ryan L. Kift, and Geoff Dell

School of Health, Medical and Applied Sciences, Central Queensland University, Rockhampton, Australia
e.crawford@cqu.edu.au

Abstract. There is a global trend towards more highly automated traffic management systems. Problematic to increased automation are the higher levels of system integration and associated complexity and uncertainty this brings. This study explored whether end-user participation in new project design and implementation can contribute to the smooth and effective introduction of these new systems for railway control rooms. Results showed that almost all participants agreed that end-user input is important to the success of new technology. However, experiences shared by interviewed controllers highlighted that end-user participation in major projects is underutilized and that certain types of involvement or action can actively hinder the implementation process. The authors anticipate that the experiences of controllers can enhance decision-making associated with end-user participation during the introduction of new traffic management systems. The authors also anticipate that end-user participation will advance improved human system integration.

Keywords: Human factors · Human-systems integration · Rail · Traffic management systems · Participatory design

1 Introduction

Over the last few centuries, the railway industry has benefited from advancements in technology. Today, due to new technical possibilities there is a global trend towards more highly automated control systems. The number of projects underway in the railways is on the rise creating a lot of activity and excitement within the industry. For instance, in the United Kingdom, Network Rail is currently phasing out and replacing their 800 operational signal boxes with 12 highly automated national control centers [1]. Sydney Trains are working on a project that will make them the first fully automated metro system in Australia [2]. Organizations introduce new technologies for many reasons. Business sustainability is particularly vulnerable in this modern competitive world. To remain relevant and competitive, changes often aim to increase efficiency, capability, reliability, customer service, and cost effectiveness, to name a few. Many organizations seek these advantages by introducing new technologies that offer new business possibilities.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_79

As the prevalence of more highly automated systems rise, they become more highly integrated and often lead to the emergence of systems of systems, a situation whereby new systems develop spontaneously [3, 4]. Changes like this initiate new system complexities [4, 5]. In the past, increased automation undermined operational safety in surprising and sometime ironic ways [6, 7]. New system uncertainties introduce new complexities and lead to concerns for safety [8]. The design space is becoming increasingly complex as system integration needs to cater for the needs of humans, software, hardware and a myriad of communication technologies, some of which do not need to be hard wired to operate. The complexity of modern traffic control systems requires a design team that can address each of these components, but also be able to work collaboratively with a diverse set of experts. In support of taking advantage of human strengths, experts call for improved human systems integration and greater teamwork between systems [4, 5, 9–12].

In relation to the design of highly automated control systems, human factors specialists often put forward the end user as a specialist that can help the design team achieve greater human system integration [3, 4, 13–15]. End users in this case are the controllers who interface with the traffic management system. In consideration of the growing complexity of traffic management systems and high priority of safety, a question arises: Is it necessary to involve end users in the design process? If so, how might they best contribute?

The remainder of this paper describes the study methods used to investigate these questions, and sets out to answer them based on study results.

2 Methods

Data used in this study came from a larger mixed methods project that examined factors that influence the successful introduction of new technology into new control-room environments [16]. Population sampling for this study was purposive to capture a deeper understanding of the problem under examination [17, 18]. A total of 438 stakeholders of control room technologies participated in this study. Participants sought were those who had an interest in areas of technology design, human-systems integration, control rooms or industries where safety is a critical element of their business, such as transport systems, power generation and distribution plants, and petro-chemical plants. Respondents fell into four categories, namely Managers, Designers, Evaluators and End users. Recruitment of participants was via an online survey, entitled: *Adoption of New Technology in Control Rooms*. The survey was available to pertinent LinkedIn networking groups, and in paper-form made available at relevant conferences and organizations that met the set sampling criteria. The survey provided both quantitative and qualitative data. In addition, thirty-six semi-structured interviews expanded upon the perspective of controllers (i.e. end users of complex systems such as traffic management systems). The interviews provided an opportunity to give end users a voice on new control-room technologies. Questions concentrated on the new technology project experiences of controllers. Participants were drawn from train traffic control ($N = 12$), air traffic control

($N = 14$) and power production and distribution centers ($N = 10$). In this study, the introduction of new technology included the design and implementation processes.

3 Results and Discussion

Survey data showed that almost all (97.5%, $N = 427$) respondents agreed that input from end users on new technologies was important. This opinion existed across the four stakeholder groups with no statistically significant variance. Of this group 88 respondents (20%) agreed and 339 (77%) strongly agreed. Less than one percent of the respondents (0.9%, $N = 4$) strongly disagreed, while seven (1.6%) neither agreed nor disagreed. These results indicate that stakeholders of safety-critical control systems feel that end-user input is important to the introduction of new technology.

From the survey, 386 (88%) participants responded to the question that asked respondents to rate the value of their experience of obtaining end-user input. A high number (92%, $N = 355$) indicated that end-user input was valuable, one percent ($N = 4$) found end-user input unhelpful, while seven percent ($N = 27$) remained neutral. These findings suggest that, while end-user input is important, not all experiences were valuable. Interview results showed that only one third ($N = 12$) of the controllers had some form of input into the design of new technologies, and even fewer felt that their experience was useful. In some cases, controllers shared experiences that had a negative effect on their desire to adopt. Thus, the successful introduction of new control-room technology has two aspects at stake, the design outcome and its adoption by end users.

3.1 End-User Participation

Survey results suggest that end-user involvement in new technology projects supports introduction success. However, experiences from interview participants suggest that, while the intention to involve end users is high, the knowledge of how to? do this appropriately is low. As expressed by participant controllers, such a situation can leave many projects open to user resistance and lack of end-user buy-in.

A common scenario shared by participants is that the creation or selection of the technology takes place during the design process without the input from end users. Just prior to implementation User Acceptance Testing (UAT)[1] is conducted to ensure the design meets agreed specifications. In some cases, the new technology is introduced small scale via a pilot, at which point end users experience the new system for the first time. Some participants had been involved in screen interface design and room layout, with one participant who was knowledgeable in software logic involved in alarm sensitivity setting. However, overall there was very little involvement of end users during the design. More typically, participants stated that their involvement was to find faults once the design of the technology was complete, often at the point of deployment, at which

[1] Note UAT does not examine whether the end user is accepting of the design. See Cimperman. R.: UAT Defined: a Guide to Practical User Acceptance testing. Addison-Wesley Professional (2006) ISBN: 0321498143 for more details.

time the expectation to adopt as quickly as possible was high. To assist with the technology adoption process, participants reported various types of training regimes, some of which included no formal training, and expectations for controllers to grasp an understanding on their own. While controllers are very capable people and able to adapt to learn a new system while operating 'live' is extremely disconcerting for them. This practice has led to disappointment, distrust in organizational leadership, and calls for a better way.

In search of a better way, interviewed controllers made it quite clear that they want a say in the technologies that influence their work before final specification. At the very least, participants want to have a suitable representative from their particular control room to table their collective viewpoint. One reason argued by controllers for participation is that, at the end of the day, they are responsible for operational safety in their area of authority and hence they need the tools to do their job well. Interviewees requested earlier involvement to ensure the "philosophy" of the new technology is right and to ensure inclusion of their operational needs.

To understand why managers and others do not involve suitable representatives, comments provided by survey participants enlighten the topic by explaining the type of contribution end users bring to the design table. Survey participants explained that experienced controllers develop a 'flair' for their jobs and have 'unique insight' into their area of authority. They are therefore the experts of their work domain. Other disciplines have labeled this type of expertise as *praxical* knowledge. Praxical knowledge goes beyond knowing how to do something, and extends to the 'art of doing' [19, 20]. Experienced controllers know the track topography, effects of weather and other aspects in the local environment. Furthermore, they know how to address issues that arise while at the same time maintaining the smooth running of trains. Praxical knowledge allows controllers to anticipate issues, such as disturbances to tracks, crossings, trains, power, maintenance, communications, and scheduling, to name a few, resolution of which often goes unnoticed by those outside the control room environment. Therefore, praxical knowledge can make a valuable contribution to the design of new technologies.

The literature on user participation links user involvement to greater project success [21, 22]. Some researchers suggest that users have a democratic right to be involved [23]. Others have reported that deeper forms of participation improve design quality and user satisfaction [24, 25], and generate more realistic resourcing as managers and designers become more aware of user needs [26]. These studies further support the end user's call for a better way, through more meaningful involvement.

Although the above studies show the benefits of end-user participation, interviewees explained that not all types of involvement are useful. For instance, interviewees noted that late and superficial involvement, as well as unrealized calls for involvement, create great frustration, can lead to distrust in management intentions and are interpreted by end users as meaningless 'tick and flick' exercises to satisfy organizational ideals but do not reflect reality. This finding aligns with another study that showed how forced or token gestures quickly reduce end-user satisfaction [27]. While many studies have shown the benefits of user participation, study findings suggest that knowing how to involve end users usefully remains elusive. Designers appreciate that end-user involvement can be problematic to achieve and can increase time requirements and costs [28].

However, others offer that such investments incurred during the design process lead to more cost-effective solutions overall due to reduced costs associated with fault finding and training activities [29]. Therefore, the authors offer that end-user participation has a place in the design arena. However, and of equal importance, how end users are involved requires careful consideration.

3.2 The End-User Contribution

With end-user participation deemed useful in the design space, the next step is to determine how they might best contribute. As identified earlier, end users have praxical knowledge for maintaining control. They therefore want to influence the design specifications. However, the achievement of early participation is historically problematic [30]. For instance, the view that human factors and user requirements are common sense concerns is responsible for delayed involvement of end users as well as human factors engineers [15, 31]. This attitude persists, despite being a well-known fundamental fallacy in relation to workspace design [32], and growing concerns across a number of other dimensions [33]. In some cases, leading engineers simply rely on intelligent guesswork [13]. In other cases, use of surrogate users has become popular. Recent studies report that designers consult the "client" who commissions and pays for the new technology as the "user" instead of the actual end user, a problem that leads to design errors [34]. Design errors are latent errors and rarely discovered until the system is field-tested and when the right (wrong) conditions arise [35].

Another concern that inhibits end-user participation is that, in heavy industry including the railways, most projects take on a techno-centric approach to the design of systems. A techno-centric approach frames all activity around the technology, including change management. Typically, end users are engaged in screen or console interface design, workplace layout and training. Involvement like this often leads to substantial changes to correct significant design oversights, if these changes are possible at all [36].

New technology is not infrequently "dropped in the laps" of controllers, a situation that often leads to high levels of stress as controllers frantically grasp how to use the new technology while working at a live desk. Exacerbating this situation is learning that the technology does not resolve a pressing need. End users of safety-critical systems should have confidence that they can do their job without fear of failure. This confidence comes from positive experiences and knowledge about the new system. The practice of leaving faultfinding to the end of the design process does little to engender end-user confidence in the new system. One controller described his experience as "frightening" to encounter so many design issues so close to deployment. Results from past studies have shown that the lack of end-user input during conceptualization leaves the engineering design process unsupported for human systems integration [31, 37], a condition that can lead to design errors and user resistance.

To improve acceptance of new technology, recent research is showing that sensemaking is an underlying factor that influences the users' understanding of new technology [16]. Sensemaking is an emerging perspective on technology adoption, an evolutionary process by which an individual gradually develops greater understanding by acting out projected meaning. These experiences in turn refine the sense

made of the new system [38]. Studies have found that sensemaking takes time and a gradual approach to technology adoption can thus be more fruitful than an approach that relies on end users to learn on their own or through training upon implementation [16, 39]. Thus, participation throughout the design process allows more time for sensemaking for technology adoption and is thus a facilitator of the successful change.

A notable barrier identified by interviewees was the achievement of time off the control desk. In certain cases, invited controllers could not logistically make it to meetings, workshops and the like due to shift conflicts and logistical matters. To be effective, end-user participation needs to be strategic, purposeful and resourced. Thus end-user participation needs the endorsement, support and resources to be made available from management if active participation to be realized. This recommendation is supported by studies that have found increases in new technology success in cases where management have resourced change and have allocated time and opportunity for end-users to get to know new systems [25].

The following are five recommendations to enhance likely end-user design participation.

Recommendation 1. Do not assume that you or others know the needs of end users. Value the contribution that end users can make during the design process. The acknowledgement that one's assumptions about another individual's needs cannot be trusted can increase appreciation for the end-user contribution and open the door for earlier involvement.

Recommendation 2. Involve end users throughout conceptualization. Opportunity for end users to participate in conceptualization will contribute to designs that meet end-user needs and thus support the development of more desirous designs. Regardless of the design approach (traditional or adaptive), end-user involvement should begin at the initial design stage. For traditional approaches, this stage will probably end once design specifications are established. Where adaptive models are used, involvement will continue throughout the design process. Also important to note here, is that end-user participation during conceptualization is likely to reduce pre-deployment testing, fault-finding, and training needs, further reducing the temptation to test on live desks, or expect controllers to learn new technologies on the fly.

Recommendation 3. Provide the necessary resources to support end-user participation. Interviewees expressed considerable frustration when invitations to participate could not be taken up. Two important logistical concerns identified by interviewees are (1) employment of adequate relief staff to give experienced controllers time off the desk, and (2) to schedule participation activities in a way that is conducive to the end-users' shift design.

Recommendation 4. Consider end-user participation as a change management strategy. End-user engagement with the new technology helps to advance sensemaking and technology adoption. Therefore, there is an expectation that participation will expedite the change management process. Appreciation of the connection between participation and change is likely to encourage greater endorsement from management and allow for more participation opportunities.

Recommendation 5. Employ an end-user advocate. In large organizations, it can be difficult for the commissioning manager to advocate for end users from each business center. Therefore, to engage a person who has sufficient power to maintain end-user interests is likely to ensure end-user participation during the design process. Considering their focus on human-centered design, human factors engineers make good end-user advocates.

One final note before the conclusion is the recognition that not all end users will be able to participate in the design process. Some interviewees stated that they would rather a trusted, more experienced controller represent them and thus preferred not to be involved. Thus, for a number of good reasons, many end users will look for other ways to make sense of the pending technology. For these individuals, sensemaking and by extension the adoption process can begin by learning from others, an idea supported by social learning theorists [40, 41]. Communication between work colleagues is likely to promote a smoother change management process because it is very effective for disseminating knowledge that contributes to technology adoption [41, 42]. However, participants noted that communication amongst end users is frequently stifled during the design stage, often in the guise of maintaining intellectual property of the technology company. In some cases, this is true and compliance with contractual agreements is paramount. Contractual agreements aside, communicating with colleagues can provide a social learning opportunity that allows all end users the opportunity to make sense of the new technology before deployment and is thus worthy of consideration.

4 Conclusion

The results from this study showed that while stakeholders of safety-critical systems have high intentions to involve end users in the design process actual involvement is low. Insight into the end users' perspective in regards to end-user participation during the design of new control-room technology revealed a number of practices that work against their desire to adopt new technology. These experiences help to expand current understanding of the underlying issues that give rise to user resistance and poor user buy-in. Therefore, to optimize the introduction of new technology into control rooms, the authors have supported the call for end-user participation during the design of new control-room technology. Recommendations offered aim to combat highlighted barriers and thus extend the body of knowledge on participatory processes and human system integration. The authors anticipate that improved decision-making regarding new technologies can result through the adoption of recommendations as offered. Furthermore, while this study is pertinent to safety-critical situations, knowledge gained has equal transference to less critical work domains.

Acknowledgments. The authors wish to thank Central Queensland University for funding this research project through the Commonwealth Government's Research Training Scheme (RTS) and Australian Postgraduate Award (APA).

References

1. Network Rail. http://www.networkrail.co.uk/publications/better-railway/
2. Sydney Metro. http://www.sydneymetro.info/metro-trains
3. Booher, H.R.: Handbook of Human Systems Integration. Wiley, Hoboken (2005)
4. Pew, R., Mavor, A.: Human-System Integration in the System Development Process: A New Look. The National Academies Press, Washington (2007)
5. Balfe, N., Wilson, J.R., Sharples, S., Clarke, T.: Development of design principles for automated systems in transport control. Ergo **55**(1), 3–54 (2012)
6. Bainbridge, L.: Ironies of automation. Automatica **19**(6), 77–779 (1983)
7. Sheridan, T.B.: Computer control and human alienation. Tech. Rev. **83**, 61–73 (1980)
8. Crawford, E.G.C., Toft, Y., Kift, R.L.: New technology adoption: risky business for the railways. In: Burgess-Limerick, R. (ed.) 46th Annual Conference of the Human Factors and Ergonomics Society of Australia, pp. 10–114. HFESA Inc., Sydney (2010)
9. Harbers, M., Neerincx, M.A.: Value sensitive design of automated workload distribution support for traffic control teams. In: Harris, D. (ed.) EPCE 2014. LNCS, vol. 8532, pp. 12–21. Springer, Switzerland (2014)
10. Inagaki, T.: Adaptive automation: sharing and trading of control. In: Hollnagel, E. (ed.) Handbook of Cognitive Task Design, pp. 169–174. Taylor & Francis, Boca Raton (2008)
11. Lüdtke, A., Javaux, D., Tango, F., Heers, R., Bengler, K., Ronfle-Nadaud, C.: Designing dynamic distributed cooperative human-machine systems. Work **41**(1), 4250–4257 (2012)
12. Miller, C.A., Parasuraman, R.: Designing for flexible interaction between human and automation: delegation interfaces for supervisory control. Hum. Fac. **49**(1), 57–75 (2007)
13. Holstein, W.K., Chapanis, A. (2015). http://www.britannica.com/EBchecked/topic/275693/human-factors-engineering
14. Maguire, M.: Methods to support human-centred design. Int. J. Hum. Com. St. **55**, 587–634 (2001)
15. Wilson, J.R., Sharples, S.: Evaluation of Human Work, 4th edn. Taylor & Francis Group, Boca Raton (2015)
16. Crawford, E.G.C.: Optimising the Introduction of New Technology into Safety-Critical Environments: A Human Factors Study. PhD. Central Queensland University, Rockhampton (2016)
17. Bryman, A.: Social Research Methods, 3rd edn. Oxford University Press, Oxford (2008)
18. Miles, M.B., Huberman, A.M., Saldaña, J.: Qualitative Data Analysis: A Methods Sourcebook, 3rd edn. SAGE, Thousand Oaks (2014)
19. Bolt, B.: The magic is in handling. In: Barrett, E., Bolt, B. (eds.) Practice as Research: Approaches to Creative Arts Enquiry, pp. 27–34. I.B. Tauris, London (2007)
20. Bolt, B. (2014). http://acuads.com.au/wp-content/uploads/2014/12/bolt.pdf
21. Abelein, U., Sharp, H., Paech, B.: Does involving users in software development really influence system success? IEEE Softw. **30**(6), 17–23 (2013)
22. Harris, M.A., Weistroffer, H.R.: A new look at the relationship between user involvement in systems development and system success. Com. Ass. Inf. Sys. **24**, 739–756 (2009)
23. Ehn, P.: Scandinavian design: on participation and skill. In: Schuler, D., Namioka, A. (eds.) Participatory Design: Principles and Practice. Erlbaum Ass, Hillsdale (1993)
24. Palanisamy, R., Sushil, J.L.: User involvement in information systems planning leads to strategic success: an empirical study. J. Serv. Res. **1**(2), 125–157 (2001)
25. Petter, S., deLone, W., McLean, E.R.: Information systems success: the quest for the independent variables. J. Mgt. Info. Sys. **29**(4), 7–62 (2013)

26. Dervin, B.: Sense-making's journey from metatheory to methodology to methods: an example using information seeking and use as research focus. In: Sense-Making Methodology Reader. Hampton Press, Cresskill (2003)
27. Kirsch, L.J., Beath, C.M.: the enactments and consequences of token, shared, and compliant participation in information systems development. Acc., Mgt. Info. Tech. **6**, 221–254 (1996)
28. Lewis, J.R.: Usability: lessons learned … and yet to be learned. Int. J. HCI. **30**(9), 663–684 (2014)
29. Burgess-Limerick, R.: Human Systems Integration is Worth the Money and Effort! The argument for the Implementation of Human Systems Integration Processes in Defence Capability Acquisition. Department of Defence, Canberra (2010)
30. Woods, D.D.: Steering the reverberations of technology change on fields of practice: laws that govern cognitive work. In: 24th Annual Meeting of the Cognitive Science Society, Austin, TX (2002)
31. Norman, D. (2010). http://www.jnd.org/dn.mss/why_human_systems_in.html
32. Pheasant, S.: Bodyspace: Anthropometry, Ergonomics and Design. Taylor & Francis, London (1986)
33. Heeks, R.: Health information systems: failure, success and improvisation. Int. J. Med. Info. **75**(2), 125–137 (2006)
34. Day, R.: Design error—The enemy of safe and robust control systems. In: IV International Confernce on Ultra Modern Telecommunications and Control Systems and Workshops, St. Petersburg, pp. 377–381 (2012)
35. Stoop, J.: No Facts, No Glory. In: Hollnagel, E., Paries, J., Woods, D.D., Wreathall, J. (eds.) Resilience Engineering in Practice: A Guidebook. Ashgate, Farnham (2011)
36. Eason, K.: Ergonomic interventions in the implementation of new technical systems. In: Wilson, J.R., Sharples, S. (eds.) Evaluation of Human Work, 4th edn. Taylor & Francis Group, Boca Raton (2015)
37. Ferreira, P.N., Balfe, N.: The contribution of automation to resilience in rail traffic control. In: Harris, D. (ed.) EPCE 2014. LNCS, vol. 8532, pp. 458–469. Springer, Switzerland (2014)
38. Weick, K.E.: Sensemaking in Organizations, Foundations for Organizational Science. Sage, Thousand Oaks (1995)
39. Seligman, L.: Sensemaking throughout adoption and the innovation-decision process. Euro. J. Inn. Mgt. **9**(1), 108–120 (2006)
40. Bandura, A.: Social cognitive theory. In: Vasta R. (ed.), Annals of Child Development: Six Theories of Child Development. JAI Press, Greenwich (1989)
41. Rogers, E.M.: Diffusion of Innovations, 5th edn. Free Press, New York (2003)
42. Burkhardt, M.E., Brass, D.J.: Changing patterns or patterns of change: the effects of a change in technology on social network structure and power. Adm. Sci. Q. **35**, 10–127 (1990)

Seat Belts Unfastened: Non-seat Belt Use in the Czech Republic

Vit Gabrhel(✉), Petr Zamecnik, Veronika Kureckova, and Pavel Rezac

Transport Research Centre (CDV), Lisenska 33a, 636 00 Brno, Czech Republic
{vit.gabrhel, petr.zamecnik, veronika.kureckova, pavel.rezac}@cdv.cz

Abstract. More than 56 thousand drivers lost their licenses due to traffic offences in the Czech Republic between the years of 2006 and 2013. In almost half of the cases the ban from driving associated with the non-compliance to the mandatory use of seat belts. At the same time, non-seat belt use annually leads to hundreds of drivers and passengers killed in traffic accidents in the Czech Republic. Thus, non-seat belt use represents one of the most frequent and severe traffic offences. However, we do not know how widespread is the non-seat belt usage in the population. Therefore, in this study, we explore the latent form of the non-seat belt use in the Czech Republic as well as related socio-demographic or behavioral factors. The quota sample consisted of 1079 drivers (18–89 years, 44% women). Almost one-third of the drivers admitted driving without fastened seatbelts during the last year. Women and more educated respondents had lower probability of driving without a seat belt than their male and less educated counterparts. At the same time, drink-driving, speeding, and perceived inadequacy of the non-seat belt use sanction led to the higher probability of non-seat belt use.

Keywords: Non-seat belt use · Risky behavior · Latent criminality · Survey · Czech Republic

1 Introduction

Traffic accidents represent a tragic part of everyday mobility. However, in many cases, injuries or even casualties related to traffic could be prevented by relatively small merits. For example, according to the National Highway Traffic Safety Administration (NHTSA) operating in the USA, seat belts have saved almost 150 thousand lives between the years of 1975 and 2001 [1]. As a result of the positive effect of the seat belt use, many countries over the world have implemented mandatory seat belt use along with penalization of non-compliance to these policies during the last 40 decades. However, there is still room for the harm reduction, even in the developed countries. For example, 55% up to 58% passenger vehicle occupants between ages of 16 to 44 years were killed in traffic accidents while unrestrained in the USA between 2003 and 2012

The original version of the book was revised: Missed out corrections have been incorporated. The erratum to the book is available at 10.1007/978-3-319-60441-1_108

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_80

[2]. Thus, more thousands of lives could have been saved by using seat belts and other safety measures such as airbags [1]. Also, apart from the tragic nature of traffic accidents, their economic costs are substantial, e.g. up to $60 billion dollars annually [3, 4].

In Europe, many countries were gradually introducing the so-called demerit point systems to provide negative feedback for risky driving behaviour and to improve road safety. The Czech Republic is among those countries which introduced the demerit point system in the 2000s within the commitment of the EU to reduce road fatalities by 50% by 2010 in comparison with 2000. The points have been recorded since 1 July 2006. According to the evaluations, a demerit point system leads to the reduction of the number of road accidents and a significant reduction of fatalities from road accidents [5, 6]. However, the dramatic improvement of road safety within one year of operation decreases with the following years and has a lower effect on fatality on motorways [6, 7]. Although the effectiveness of demerit point system is the highest in one or two years from its introduction [8], it has, generally speaking, a long-term positive effect on the use of safety belts and child restraint systems [9].

Despite the employed measures, the situation regarding traffic safety is no less severe in the Czech Republic as in other developed countries in Europe or in the USA. According to the data provided by the Ministry of Internal Affairs, 56 thousand drivers lost their driving licenses due to traffic offences between the years of 2006 and 2013. While 24 551 (44%) of the drivers did not use a safety belt at least once, 13 016 (23%) did not comply with this rule twice or more. At the same time, according to the Road Safety Department of the Ministry of Transport of the Czech Republic (BESIP), 545 passenger vehicle occupants died in traffic accidents in 2016 and non-use of seat belts have led to hundreds of killed in traffic accidents annually in the past years [10, 11].

All of the previously mentioned information relate to the revealed non-seat belt use. However, undocumented non-seat belt use could be assumed as even more prevalent than its revealed form. At the same time, the latent form of non-seat belt use has received only limited attention. Thus, this study aims to explore the latent form of non-seat belt use among the population of drivers. Also, it focuses on the socio-demographic and other factors related to non-seat belt use that was identified by previous inquiry. The body of existing research pointed out that persons who are male, young, less educated, have a lower income, and reside in rural areas are less likely to wear safety belts than their counterparts [12–15]. Moreover, risky behavior such as drink-driving or speeding tends to be related to non-seat belt use [16]. Last but not least, compliance with the mandatory traffic safety measures relates to the perceived effectivity of law-enforcing system [17]. As a result, we use these variables to assess their relevance regarding latent non-seat belt use within the Czech context.

2 Methods

2.1 Design

Quota sampling was used to obtain data representative on the socio-economic basis. The quotas consisted of age, gender, educational background, and spatial characteristics such as the adequate proportion of respondents hailing both from urban and rural areas of the Czech Republic.

The data were collected via computer-assisted personal interview (CAPI) during the spring of 2016. No respondents had participated in a market research or an opinion poll at least 6 months before the data collection.

2.2 Methods

The data were collected via standardized questionnaire. It consisted of total 57 items divided into three sections.

The first part included indicators of the risky behavior–non-seat belt use, drink-driving (both right after the intoxication and in its delayed or 'residual' form), speeding less than 10 km/h, and speeding more than 10 km/h. In each of the variables, respondents stated that certain risky behavior either occurred or did not occur during the past year. The categories of answer were intentionally broad in order to minimize social desirability as well as cognitive burden.

The perceived effectivity of traffic police was measured by variables such as experience with traffic police control during the last year (that either occurred or did not occur), predicted number of drivers who change their behavior after traffic police control (on the scale from 0 to 10), and perceived adequacy of the legal penalization related to the non-seat belt use (either adequate or non-adequate).

The last group of variables consisted of socio-demographic characteristics such as age, gender, educational background or size of the town or the city that respondents live in (i.e. the number of inhabitants). These variables are covered further in the Sample section of the text.

The questionnaire was assessed as a whole in a pilot test–its completion time, the general intelligibility of the questionnaire, and the like. Five individuals (members of a general population with various age, gender, and educational and economic background) participated in the pre-testing, which revealed no problematic item or aspect of the questionnaire.

2.3 Sample

There were 1482 participants in the original sample. However, 403 of them declared that they did not drive a car at all in the past year. Since in this study we focus only on the active driver population, these respondents were omitted from further analyses. The final sample, therefore, consisted of 1079 respondents that had driven a car at least once during the last year prior to the survey.

In the sample, men (N = 606, i.e. 56%) were represented slightly more than women (N = 473, i.e. 44%). The mean age in the sample equalled to 45 (SD = 15). The youngest respondent was 18 years old, the oldest 89 years old. The age distribution is described further in Fig. 1:

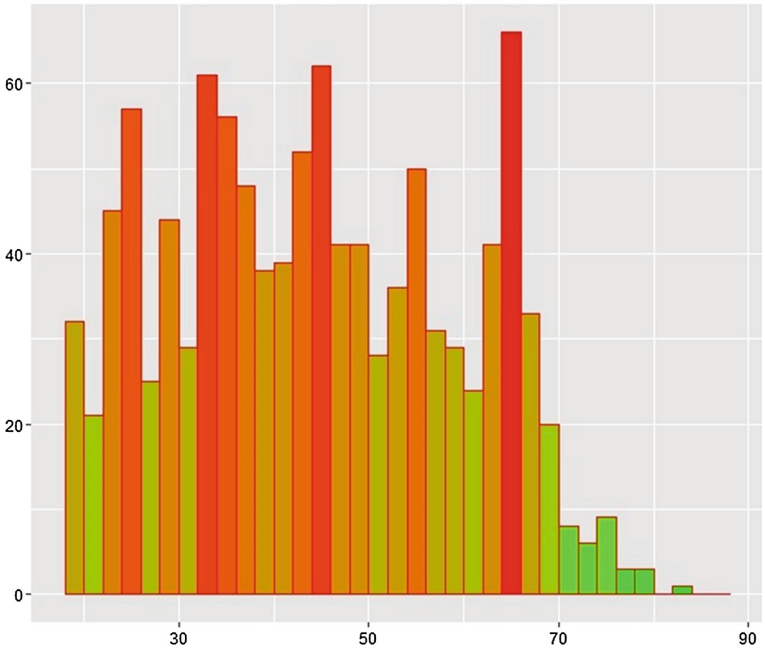


Fig. 1. Histogram showing the age distribution of the participants. The x-axis represents the age, whereas the y-axis the number of respective cases.

As for the educational background of the respondents, based on the International Standard Classification of Education or ISCED [18], 12% (N = 127) of respondents finished their education at primary level of education (ISCED1), 37% (N = 401) at general secondary education (ISCED34), 36% (N = 392) at vocational secondary education (ISCED35), and 15% (N = 159) finished their studies at tertiary level of education (ISCED5-8). From the income point of view, most of the respondents (821, i.e. 76%) refused to state their net monthly household income. As a result, this variable was omitted from the further analyses. Thus, the level of education served as the proxy of respondents' socio-economic status.

Last but not least, most of the respondents (N = 432, i.e. 40%) lived in towns with the maximum of 4999 inhabitants, 20% of participants (N = 213) lived in towns between 5000 and 19,999 inhabitants, and 24% (N = 262) respondents stated that they live in cities between 20,000 and 99,999 inhabitants. Lastly, 16% of the respondents (N = 172) lived in the cities with more than 100 thousand inhabitants.

3 Results

The data were analyzed via programming language and environment for statistical computing R [19].

3.1 Descriptive Statistics

When it comes to the traffic-related risky behavior, 30% of respondents (322) admitted that they had not used safety belts at least once during the last year. Speeding was even more prevalent as almost 70% (N = 719) of respondents stated that they breached the speed limit by more than 10 km/h at least once during the last year. Last but not least, 7% (N = 73) of drivers admitted drink driving. Table 1 contains detailed information regarding drivers' risky behavior:

Table 1. The occurrence of traffic-related risky behavior during the last year

Risky behavior	Never		At least once	
	N	%	N	%
Non-seat belt use	757	70	322	30
Drink-driving (actual intoxication)	1006	93	73	7
Drink-driving (residual)	914	85	165	15
Speeding (up to 10 km/h)	151	14	928	86
Speeding (more than 10 km/h)	360	33	719	67

Note: N = 1079

As for the indicators of the perceived effectiveness of the law-enforcing system, the mean value of drivers who, according to the participants, change their behavior as a result of a traffic-police control, ranging from 0 (actual minimum within the sample to 10 (actual maximum within the sample), equaled to 4.4 (SD = 2.6). Moreover, almost half of the participants (N = 503) declared that they had experienced traffic-police control during the last year. Last but not least, 74% of participants (N = 803) consider penalization for non-seat belt use as adequate.

3.2 Model Predicting Non-seat Belt Use

To assess factors related to the non-seat belt use, we conducted a logistic regression with non-seat belt use (never versus at least once during the last year) as the binary outcome variable. In every categorical variable, we used indicator coding with the first category as the referential one. The total of 1079 respondents entered the model. Results are presented in Table 2.

According to the model, age of the respondents, size of the town or the city that respondents live in (i.e. the number of inhabitants), drink-driving with residual alcohol, experience with traffic police control, and predicted the number of drivers who change their behavior as a result of traffic police controls did not account for non-seat belt use category membership. On the other hand, women and more educated respondents had a significantly lower probability of driving without a seat belt than male and less educated counterparts. At the same time, drink-driving in terms of actual intoxication, speeding in the both observed categories, and perceived inadequacy of the non-seat belt use sanction led to the higher probability of non-seat belt use.

Table 2. Final model of a hierarchical logistic regression

Predictor	B	SE	p	Exp (B)	95% C.I. for Exp (B)	
					Lower	Upper
Gender						
Women versus men	−.32	.16	.04	.72	.53	.99
Age	−.01	.01	.19	.99	.98	1.00
Education						
Vocational secondary versus primary	−.16	.23	.51	.86	.54	1.36
General secondary versus primary	−.64	.24	.01	.53	.33	.85
Tertiary versus primary	−1.04	.30	<.00	.35	.20	.63
Size of the town/city in number of inhabitants						
5000–19,999 versus 4999 and less	.31	.20	.11	1.37	.93	2.01
20000–99,999 versus 4999 and less	−.19	.19	.32	.82	.57	1.20
More than 100,000 versus 4999 and less	−.09	.22	.70	.92	.60	1.41
Drink-driving (actual intoxication)						
At least once versus never	1.17	.32	<.00	3.21	1.73	5.98
Drink-driving (residual)						
At least once versus never	.26	.22	.24	1.30	.84	2.01
Speeding (up to 10 km/h)						
At least once versus never	.72	.34	.03	2.06	1.07	3.97
Speeding (more than 10 km/h)						
At least once versus never	1.15	.21	<.00	3.14	2.09	4.72
Traffic-police control						
No versus yes	−.17	.15	.25	.84	.62	1.13
Adequacy of the non-seat belt use sanction						
Adequate versus not adequate	.32	.16	.05	1.38	1.00	1.89
Number of drivers who change their behavior	−.02	.03	.40	.98	.92	1.03

Note: N = 1079, N = 1165; R^2 = .15 (Cox and Snell), .22 (Nagelkerke). Model χ^2 (8) = 3.26, p = .92

4 Discussion

This study had two aims. The first was to provide an insight into the latent criminality in non-seat belt use in the Czech Republic. From this point of view, non-seat belt use was quite present in the sample of drivers since almost one-third of them admitted this kind of risky behavior at least once during the last year. In addition, speeding was even more prevalent as 86% of all respondents declared that they exceeded the speed limit by less than 10 km/h and 67% by more than 10 km/h. Even though the number of drivers who admitted drink-driving was proportionately lower, its value was itself alarming. Overall, respondents declared a violation of the traffic rules quite openly. Their honesty may suggest that respondents did not identify themselves with the setting of the legal system in traffic or even internalize the legal system. If these values are to be truth in the population, much more effort has to be put into dealing with this issue.

However, this effort has to focus on the socialization of drivers, i.e. on psychological and cultural aspects of driving [20].

The second goal was to assess socio-demographic and other factors related to the non-seat belt use in the Czech Republic in its latent form. In this regard, the results are consistent with the previous findings in terms of men being more likely not to use seat belts than women or less educated persons more than those with a higher level of education. On the other hand, rural versus urban environment did not influence the non-seat belt use [12–15]. At the same time, the trend present in the educational level did not apply in our sample in general. Respondents with vocational secondary level education did not differ from those with a primary level of education in non-seat belt use. Both groups may have similar behavioral patterns based on similar lifestyles that negatively project into their attitudes towards risky behavior as well as into the actual behavior. Last but not least, despite the fact that almost half of the respondents declared experience with traffic police control during the last year, it did not play any role when it comes to the (non)seat belt use. Again, this may relate to only formal acceptance of the traffic laws by the respondents. The same principle is applicable to the perceptions of adequacy of the sanctions related to the non-seat belt use, which did not play a significant role when predicting non-seat belt use.

However, our study has several limits. First of all, our sample did not represent the general population as we focused solely on the active drivers. This fact may bias the external validity of our findings regarding socio-demographic characteristics such as gender or age of the participants, although the original sample reflected the general population in these aspects. On the other hand, overrepresentation of men and other socio-economic groups in comparison to the general population could be attributed to nature of the studied phenomena. At this point, however, we do not have the certainty. Moreover, quota sampling does not control for the systematic influence of variables that did not serve as quotas. In addition, measures used in this study provide only limited level of precision. As a result, it is not possible to differentiate between respondents in a more detailed way. For example, speeding by more than 10 km/h could include both 11 km/h and 50 km/h.

As for the further directions, measures assessing risky behaviour in greater detail are needed. At the same time, the further research should include psychological measures tailored for the context of traffic offenses, such as attitudes and perceptions towards non-seat belt use [21].

Acknowledgments. This article was produced with the financial support of the Ministry of Education, Youth and Sports within the National Sustainability Programme I, a project of Transport R&D Centre (LO1610), on the research infrastructure acquired from the Operation Programme Research and Development for Innovations (CZ.1.05/2.1.00/03.0064).

References

1. Glassbrenner, D.: Estimating the lives saved by safety belts and air bags. Age **5**, 132–137 (2016)
2. National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811892

3. U.S. Centers for Disease Control and Prevention. http://www.cdc.gov/motorvehiclesafety/index.html
4. Blincoe, L., Seay, A., Zaloshnja, E., Miller, T., Romano, E., Luchter, S., Spicer, R.: The Economic Impact of Motor Vehicle Crashes, 2000. National Highway Traffic Safety Administration, Washington, DC (2002)
5. Pulido, J., Lardelli, P., De la Fuente, L., Flores, V.M., Vallejo, F., Regidor, E.: Impact of the demerit point system on road traffic accident mortality in Spain. J. Epidemiol. Commun. Health **64**(3), 274–276 (2010)
6. Butler, J.S., Burke, J.P., Healy, D.G., Stephens, M.M., McManus, F., McCormack, D., Poynton, A.R.: Trends in RTA related spinal injuries: the post penalty points era. Ir. J. Med. Sci. **175**(1), 20–23 (2006)
7. Farchi, S., Chini, F., Rossi, P.G., Camilloni, L., Borgia, P., Guasticchi, G.: Evaluation of the health effects of the new driving penalty point system in the Lazio Region, Italy, 2001–4. Injury Prevent. **13**(1), 60–64 (2007)
8. Castillo-Manzano, J.I., Castro-Nuño, M., Pedregal, D.J.: An econometric analysis of the effects of the penalty points system driver's license in Spain. Accid. Anal. Prev. **42**(4), 1310–1319 (2010)
9. Zambon, F., Fedeli, U., Visentin, C., Marchesan, M., Avossa, F., Brocco, S., Spolaore, P.: Evidence-based policy on road safety: the effect of the demerit points system on seat belt use and health outcomes. J. Epidemiol. Commun. Health **61**(10), 877–881 (2007)
10. Centrum služeb pro silniční dopravu. http://www.cspsd.cz/storage/files/nehody_2016.pdf
11. BESIP. http://www.ibesip.cz/cz/aktivity/archiv-kampani/bezpecna-obec/komunikacni-aktivity/vlastni-medialni-projekty/bezpecnostni-pasy
12. Nelson, D.E., Bolen, J., Kresnow, M.J.: Trends in safety belt use by demographics and by type of state safety belt law, 1987 through 1993. Am. J. Public Health **88**(2), 245–249 (1998)
13. Glassbrenner, D.: Safety Belt Use in 2004—demographic results. US Department of Transportation, Washington, DC (2005)
14. Boyle, J.M., Vanderwolf, P.: Motor Vehicle Occupant Safety Survey, Volume 2: Safety Belt Report. US Department of Transportation (No. HS-809 789), Washington, DC (2004)
15. Lerner, E.B., Jehle, D.V., Billittier, A.J., Moscati, R.M., Connery, C.M., Stiller, G.: The influence of demographic factors on seatbelt use by adults injured in motor vehicle crashes. Accid. Anal. Prev. **33**(5), 659–662 (2001)
16. Wilson, R.J.: The relationship of seat belt non-use to personality, lifestyle and driving record. Health Educ. Res. **5**(2), 175–185 (1990)
17. Vito, G.F., Maahs, J.A., Holmes, R.M.: Criminology: Theory, Research and Policy. Jones and Bartlett, Sudbury (2001)
18. International Standard Classification of Education. http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf
19. R: A language and environment for statistical computing. https://www.R-project.org/
20. Cauzard, J.-P. (ed.): European Drivers and Road Risk: Project on Social Attitudes to Road Traffic Risk in Europe SARTRE 3. Institut National de Recherche sur les Transports et leur Sécurité, Arcueil (2004)
21. Simsekoglu, O., Lajunen, T.: Social psychology of seat belt use: a comparison of theory of planned behavior and health belief model. Transp. Res. Part F: Traffic Psychol. Behav. 11 (2008)

The Effect of See-Through Truck on Driver Monitoring Patterns and Responses to Critical Events in Truck Platooning

Bo Zhang[1(✉)], Ellen S. Wilschut[2], Dehlia M.C. Willemsen[3], Tom Alkim[4], and Marieke H. Martens[1,2]

[1] University of Twente, 7522 NB Enschede, The Netherlands
b.zhang@utwente.nl
[2] TNO Human Factors, 3769 ZG Soesterberg, The Netherlands
[3] TNO Integrated Vehicle Safety, 5708 JZ Helmond, The Netherlands
[4] Rijkswaterstaat, 3526 LA Utrecht, The Netherlands

Abstract. Automated platooning of trucks has its beneficial effects on energy saving and traffic flow efficiency. The vehicles in a platoon, however, need to maintain an extremely short headway to achieve these goals, which will result in a heavily blocked front view for the driver in a following truck. Monitoring surrounding traffic environment and foreseeing upcoming hazardous situations becomes a difficult, yet safety-critical task. This exploratory study aims to investigate whether providing platoon drivers with additional visual information of the traffic environment can influence their monitoring pattern and increase awareness of the upcoming situation. 22 professional truck drivers participated in the driving simulator experiment, either following a see-through lead truck (i.e., with projection of forward scene attached to the rear of the lead truck), or a normal lead truck until the automation system failed unexpectedly in a critical situation. Results showed that when provided with front view projection, the participants spent 10% more time monitoring the road, and responded less severely to a critical situation, suggesting a positive effect of the "see-through" technology.

Keywords: Automated driving · Truck platooning · Human machine interaction · Takeover · Eye tracking · Response time

1 Introduction

The technological progress in intelligent transportation systems (ITS) and communication technologies enables rapid development of automated vehicles. Some vehicles already have systems such as Traffic Jam Assist that work at low speeds, which can keep a distance to the vehicle in front (e.g., based on Adaptive Cruise Control) and take control of the lateral position of a car within its lane based on road markings (Lane Keeping Assist System). Drivers are thus relieved from dynamic control tasks in specific situations. There are systems available or under development that work on motorways at higher speeds as well. One example is cooperative platooning, which can be defined as a group of automated vehicles travel closely together for the benefits of

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_81

energy saving, traffic flow efficiency and safety [1, 2]. In 2016, the Netherlands initiated a European Truck Platooning Challenge where partially automated trucks of various brands traveled in a platoon to the Netherlands from different European cities (e.g., DAF Trucks participated under the name of the 'EcoTwin' project [3]), showing the possibility and readiness of trucking platooning to be implemented on public roads in the near future [4].

Despite the promising perspective, considerable efforts still have to be made to realize these high levels of automation where human intervention and supervision are rarely needed. The current commercially available systems are not yet able to cooperate with human drivers/supervisors at an optimal level [5], and the drivers are still required to be prepared to take over immediate manual control in case of system boundaries and failures. Maintaining sufficient situation awareness about the surrounding driving environment and system status therefore becomes safety critical. However, this can be challenging even for drivers who are constantly monitoring because this monotonous task can easily cause boredom and vigilance decrement, which can cause slow and inadequate responses [6–8]. The situation is especially disadvantageous for the drivers in a truck platoon when the headways to the lead truck have to be extremely short (between 0.2 and 0.3 s) in order to effectively save up fuel consumption by reducing air resistance [1, 9]. On the one hand, the blocked and monotonous front view may increase difficulties for the platoon drivers to stay "in the loop", and anticipate upcoming traffic situations. Consequently, platoon drivers may not be able to anticipate critical situations until being notified by the system. In extreme cases where timely warning cannot be guaranteed, the drivers are forced to respond immediately and an unsafe situation is most likely to occur due to a lack of preparation. Although it is a highly exceptional situation and should in principle not occur, it is interesting to study how critical this will be and if additional measures may be helpful.

One additional measure may be to capture real-time video streams by a camera installed in the front of the vehicle, and broadcast the image to the following vehicle either through an in-vehicle device [10–13] or on a large screen attached to the back of the lead vehicle such as implemented in Samsung Safety Truck [14]. These see-through systems (STS) provide drivers with very intuitive traffic information and facilitate safety assessment during overtaking maneuvers.

Seeing the potential benefits of implementing STS in truck platooning systems, we aimed to investigate whether providing platoon drivers with additional visual information of the forward traffic scene can influence their monitoring pattern and increase awareness of the upcoming situation. A driving simulator experiment was conducted in the present study to analyze driver gaze behavior and responses to a sudden and time-critical event when following a see-through lead truck (i.e., with a screen at the rear part, inspired by the Samsung Safety Truck). A normal lead truck served as the control condition. Accordingly, the hypotheses were formulated as follows:

1. When driving behind a see-through truck, the forward scene becomes less monotonous and drivers will allocate more time monitoring the road, especially the see-through screen area.
2. More information about the situation by means of a see-through truck helps a driver to respond more quickly and less severely in case of an emergency situation.

2 Method

2.1 Participants

22 participants took part in the experiment. They all held a truck driver's license for at least 8 years and drove at least 10,000 km per year. On average participants were 47.4 years old, with a standard deviation of 11.5. Minimum age was 27 and maximum 64 years old. The group consisted of 20 male and 2 female drivers.

2.2 Apparatus

The experiment was carried out in a high fidelity moving base driving simulator consisting of a DAF truck mock-up mounted on a 6DOF moving base. The road and traffic environment were projected on cylindrical screens around the vehicle with a horizontal viewing angle of 180° and two external rear view mirrors. Vehicle and situational parameters such as pedal positions, speed, steering wheel angle and button reaction time were recorded at a sampling rate of 50 Hz. Synchronously, driver eye movements were recorded using a SmartEye nonobtrusive remote eye tracker (SmartEye AB, Gothenburg, Sweden) at a sampling rate of 60 Hz.

2.3 Two-Truck Platooning System

In this experiment, an automated two-truck platooning system was simulated that allows a truck to follow a lead vehicle at a relatively short following distance, controlling both the longitudinal and lateral motion of the vehicle. The first truck is intended to be driven by a human operator (but is controlled by the simulator in the experiment). Once engaged, the second truck is controlled by the automated system. The automated system was modelled as a combination of a Cooperative Adaptive Cruise Controller (CACC) and a Lane Keeping Assist System (LKA). The driver could push a button normally used for cruise control on the right side of the steering wheel to switch the automated system on/off. To be able to switch the system on, the driver had to drive in an 'activation zone' behind the lead truck at a distance close enough to hook on. After activating the system by pushing the 'on/off' button the system would take over both longitudinal and lateral control. To deactivate the automated system, the driver had to push the 'on/off' button again. The automated system would then transfer both longitudinal and lateral control back to the driver. A tablet display located at the right side of the dashboard indicated status change of the automated systems.

2.4 Experimental Design and Test Procedure

Participants drove on the right-hand lane (the slower lane) of a two-lane motorway behind a lead truck that was driving with an average speed of 80 km/h. The participants were instructed to follow this lead truck and not to change lanes. There were no entries or exits on the route the participants drove. Slight curves and surrounding traffic were

included in the drive. All participants started with a training session to get familiar with the driving simulator and the automated driving system. During training they were asked to perform the coupling and decoupling procedures at least three times, until they felt comfortable with it. After training, all participants started with a baseline run, i.e. without the automated system, hence normal manual truck driving (MD). This was a normal drive on the same road as they would drive on in the conditions with the system. After the baseline trial, the experimental session testing automated driving conditions (AD) began. In total, each participant conducted eight AD trials under different conditions. In this paper, we only report results about the last trial where the see-through truck condition and a critical event were included.

In the critical event trials, 10 participants drove behind a normal lead truck and 12 drove behind a lead truck which had a simulated screen attached to the rear of the truck showing the images of the road before the lead truck, thus giving the impression of a see-through truck (see Fig. 1). Shortly after the beginning of the trial, the participants pressed the button to activate the automation system and follow the lead truck at a short distance of 0.3 s time headway. In order to have the largest chance of a proper driver response, the participants were instructed to be attentive of the traffic during this scenario. After 6500 m (approx. 5 min after the activation of the system) a visual/auditory warning was issued that the control was switched back to manual control. Immediately afterwards, the lead truck made an emergency maneuver either to the left lane (50% of the cases) or to the emergency lane (50%). A stationary vehicle was positioned 200 m from the onset of the warning (approx. 6.5 s to collision if no reaction was taken). After pressing the button to reclaim manual control, the drivers could either brake, or steer to the adjacent lane to avoid an accident. An overview of the experimental conditions is presented in Table 1.

Fig. 1. Screenshot of a normal lead truck (left) and a see-through lead truck (right) in the simulation.

2.5 Dependent Variables

Gaze Behavior. The analysis of visual behavior was conducted comparing three conditions: following a see-through truck (AD), following a normal truck (AD) and baseline drive (MD). For both AD conditions, the time interval for analysis was the period when the automation system was activated. An equivalent time interval was

Table 1. Overview of the experimental conditions

	Evading to the right	Evading to the left	N (participants)
Normal truck AD	x		5
Normal truck AD		x	5
See-through truck AD	x		6
See-through truck AD		x	6
Baseline MD			22

chosen for the manual driving condition to allow the comparison with the AD conditions. Four relevant Areas of Interest (AOIs) were defined: windshield, left mirror, dashboard and secondary screen (where system status and warning were displayed). Glances outside the defined AOIs are labeled "Others". Three basic metrics according to ISO 15007 [15] were utilized to gain an overview about drivers' information acquisition: *percent time on AOI*, *mean glance duration* and *glance rate*. These measures facilitate the understanding of driver monitoring pattern and enable quantitative comparison between the conditions. Since the windshield is the largest and most important AOI representing a driver's forward view, and one glance onto this area may contain multiple fixations, a further analysis was conducted for the fixations located within this area for a deeper insight. A three dimensional figure was plotted for each condition combining the positions of the fixations and the corresponding fixation durations, aiming to show an intuitive picture of drivers' gaze patterns when monitoring the traffic environment.

Response to Critical Event. Analysis for the takeover behavior and quality was only conducted for two AD conditions. To assess the takeover behavior, we analyzed the reaction types for each condition analog to the previous study of Gold et al. [16], namely braking, changing lane without deceleration, and changing lane with deceleration. We assessed timing and quality aspects of the takeover based on a number of dependent variables. Regarding the timing aspects, driver response time (RT) was defined as the time interval from the warning onset until the first conscious intervention was taken, either by pressing the brake pedal for more than 10% or changing the steering angle more than 2° compared to the mean of the steering angle 10 s before the warning. Whichever happened first counted.

Quality aspects of the takeover were evaluated by the maximum brake accelerations, maximum steering angle and the minimum occurring time to collision (min TTC) within each condition. It was assumed that less harsh braking and smaller steering angles are equivalent to a less dynamic maneuver. Min TTC is an objective measure to assess the criticality to the situations. Larger min TTC indicates longer time left to avoid a collision with the obstacle, and therefore a safer handling. In summary, early intervention, less severe maneuvers and larger minimum TTC indicate better performance.

3 Results and Discussion

3.1 Gaze Behavior

Two participants in the see-through truck condition and five participants in the baseline condition were omitted from the analysis due to missing data in eye-tracking recordings. To gain an overview about drivers' attention allocation, the percent time on AOI was first presented in Fig. 2, based on total glance times for all participants within each condition. Results showed when driving manually, the drivers devoted the largest percentage of their attention on the road (85% within all conditions). When following the normal lead truck, the participants showed the lowest percentage of time of eyes on the road (62%), and spent around one third of the total time looking at undefined areas. By checking video recordings, we concluded that they were predominantly looking outside the left window, which was out of the recording range of the eye-tracking system. When following the see-through truck, 73% of their time was spent monitoring the road, which is 11.3% more compared to driving behind a normal truck. In both AD conditions, the participants spent slightly more time checking the mirror, but less time monitoring the dashboard.

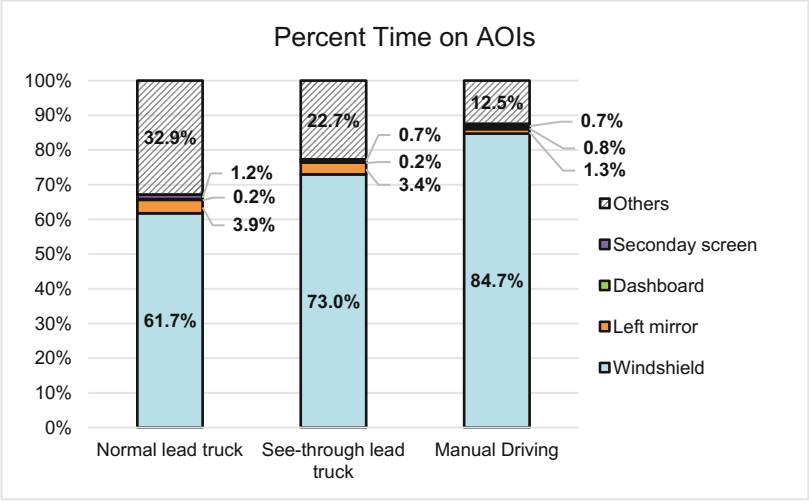


Fig. 2. Percent time on AOIs compared across groups

Figures 3 and 4 show mean glance durations and glance rates regarding different AOIs. Due to the fact that not all participants looked at the left mirror, dashboard or secondary screen throughout the experiment, statistical tests were only performed for glances onto the windshield for the three conditions. Two one-way ANOVA tests were conducted, showing significant differences between groups in terms of mean glance duration ($F(2,33) = 3.359$, $p = .047$). A Bonferroni post hoc test revealed that when driving behind a see-through truck, the mean glance duration onto the windshield was

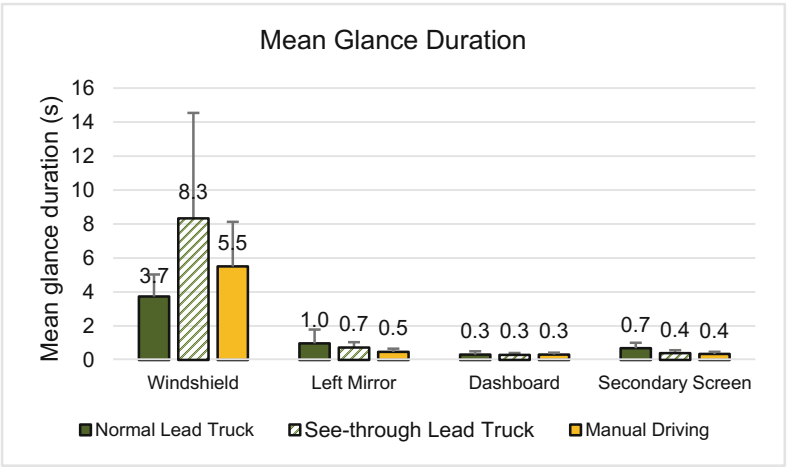


Fig. 3. Mean glance duration on AOIs compared across groups. Error bars represents the standard deviations.

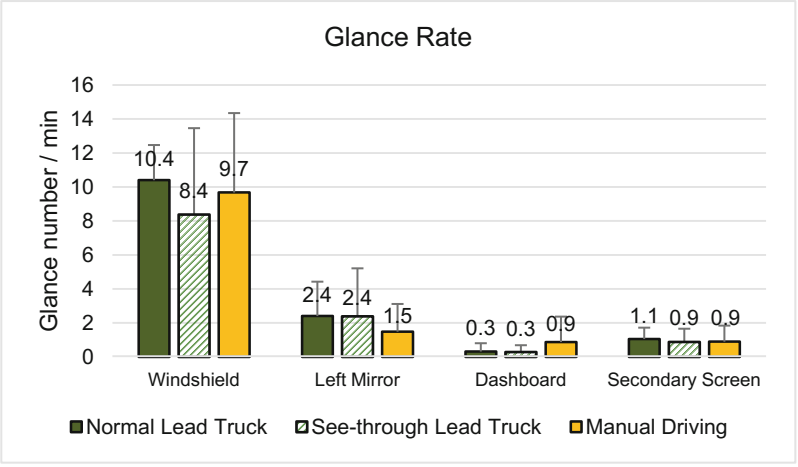


Fig. 4. Glance rate on AOIs compared across groups. Error bars represents the standard deviations.

significantly longer than following a normal truck ($p = 0.043$). Differences between other conditions were not significant. Regarding glance rate, thought the tendency of less frequent attention switching between the windshield and other areas can be seen for the see-through truck group compared to the other two groups, no significant differences were found.

Figure 5 shows how the drivers distributed their attention within the windshield area. The position of each fixation point was illustrated on the x/y plane according to their

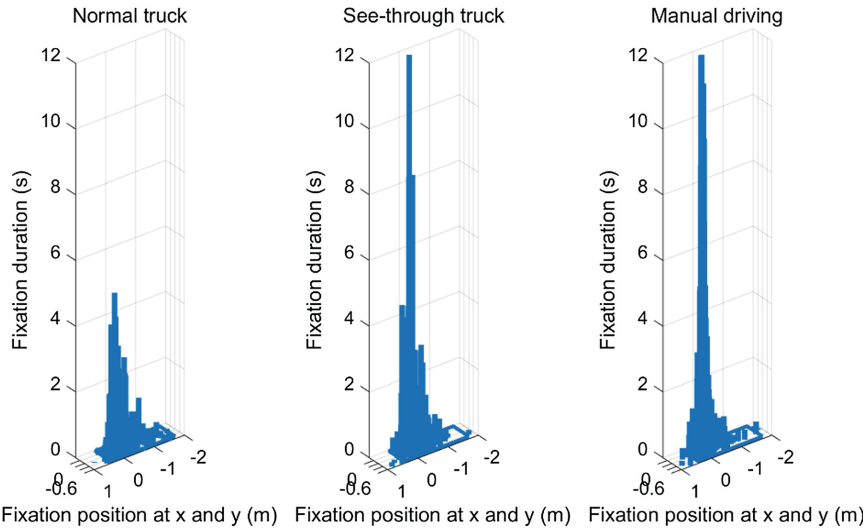


Fig. 5. Fixations on the windshield. The positions of each fixation point are illustrated on an x/y plane according to their world coordination. The values on the z-axis indicate the corresponding fixation durations.

world coordination in the eye-tracking system. The blue rectangle on this plane represented the configuration of the windshield. The duration of each fixation was also visualized on the z-axis. As can be seen from the 3D figure, when driving manually and following a see-through truck, the participants concentrated more on the road center area (the left part of the rectangle) with longer fixation durations. Some extreme values were even above 6 s. When following a normal truck, drivers showed the similar tendency to focus more on the road center, but the fixations were shorter and more dispersed.

Combining the results above, the implementation of the see-through truck significantly increased the amount of attention allocated on the road, especially the center area where the screen was located. These differences are mainly caused by longer single forward glances rather than more frequent fixations of the eyes. Furthermore, it can be inferred that when following a see-through truck, a driver's gaze behavior is more similar to that during manual driving. Since the drivers were not instructed to do any secondary tasks, the behavior of shifting attention away from the road implied a vigilance decrement when conducting the monitoring task. The shorter eyes-off-road time of the drivers in the see-through truck group indicated a higher level of attentiveness and more active monitoring. The first hypothesis is therefore supported and the results suggest a positive effect of the see-through truck on driver monitoring behavior.

3.2 Response to the Critical Event

We measured driver performance metrics within 10 s after the warning onset. All participants successfully avoided a collision with the stationary truck. Table 2 gives an descriptive overview of driver reaction types in response to the critical event. Seven out

Table 2. Numbers of participants who conducted three reaction types in each condition

	Brake	Lane change without deceleration	Lane change with deceleration
Normal truck	7/10	2/10	1/10
See-through truck	8/12	2/12	2/12

of 10 drivers stayed in the right lane and braked in order to avoid the stationary truck when following a normal truck (70%). One of the three drivers who did a lane change braked to full stop in the adjacent lane and didn't pass the obstacle. For the see-through truck, four out of 12 drivers changed lanes (33%).

Table 3 contains the means and standard deviations for the takeover quality measures, and the results of the independent t-test for the comparison between groups. For the three vehicle related measures, results are reported only for the participants who did not conduct evasion maneuvers in order to avoid biases, because the operations on the brake pedal and steering wheel during lane changing are not comparable to those during braking in the ego lane. The only exception is to include the participant that conducted a harsh brake after a lane change in the analysis of max. braking acceleration.

Table 3. Means and standard deviations for response times, minimum TTC, maximal brake acceleration and maximum steering wheel angle (SWA), and t-test results for the comparison between groups

	Normal Truck	See-through Truck	t-test results
Response time (s)	M = 2.70 SD = 0.55	M = 2.89 SD = 0.64	$t(20) = -0.728$, $p = 0.475$
min TTC (s)	M = 4.42 SD = 0.79	M = 4.13 SD = 0.66	$t(13) = 1.059$, $p = 0.309$
max. brake acc. (m/s^2)	M = 6.14 SD = 0.27	M = 5.28 SD = 0.96	$t(14) = 2.280$, $p = 0.039^*$
max. SWA (deg.)	M = 14.65 SD = 3.58	M = 12.62 SD = 8.03	$t(13) = 2.439$, $p = 0.030^*$

Contrary to our expectation, the drivers in the see-through truck condition did not intervene faster in response to the takeover event. Similarly, no significant differences were found regarding min TTC. Significant effects of the see-through truck were found when comparing maximum brake acceleration and maximum SWA, with smaller values for both measures in the see-through group. Therefore, a mixed answer was obtained when testing our second hypothesis. It seems that the implementation of see-through lead truck could be associated with less dynamic response maneuvers, but did not evoke earlier intervention responses. Seen from the min TTC and the fact that no collision occurred, the situation never became really critical for the participants so a faster response might not be necessary. Another explanation can be that the participants in the see-through truck group already detected the obstacle before the warning, so they could better prepare themselves for a moderate response maneuver. Previous work of

Gold et al. [16] suggested a faster reaction within a short time budget could be accompanied with worse quality. The extra time budget gained from monitoring the see-through screen might have encouraged the participants to take slower, but less severe responses. These assumptions need further elaboration to be tested.

4 Conclusion and Outlook

In summary, despite the limited number of participants, a positive effect of the see-through truck can be suggested, with the participants allocating more time monitoring the traffic and responding less severely when encountering a critical event within comparable safety boundaries. This study also provided insight in possible safety issues in case of unforeseen scenarios. Even though drivers showed to be capable of responding appropriately when actively monitoring the environment, there are large individual differences in response times and type of response. Further research will need to focus on the individual differences between truck drivers, on other types of critical situations and on responses when a driver has been out of the loop when distracted by secondary tasks which will be a normal situation in future truck platooning concepts. In addition, the see-through truck tested in the current study has its drawbacks such as high cost and vulnerability to damage. It would be of interest to investigate the effects of in-vehicle see-through displays implemented in truck platooning in future studies.

Acknowledgments. This study was funded jointly by the TNO Early Research Program (ERP) Human Enhancement: Adaptive Automation, and Rijkswaterstaat. Ministry of Infrastructure and the Environment (in the context of the knowledge agenda automated driving, the full report can be found and downloaded at the corresponding website: http://knowledgeagenda.connekt.nl). The first author and the last author are involved in the Marie Curie Initial Training Network (ITN) project HFAuto – Human Factors of Automated Driving (PITN-GA-2013-605817).

References

1. Bergenhem, C., Shladover, S., Coelingh, E., Englund, C., Tsugawa, S.: Overview of platooning systems. In: Proceedings of the 19th ITS World Congress, 22–26 October, Vienna, Austria (2012)
2. Alam, A., Besselink, B., Turri, V., Mårtensson, J., Johansson, K.H.: Heavy-duty vehicle platooning for sustainable freight transportation: a cooperative method to enhance safety and efficiency. IEEE Control Syst. Mag. **35**(6), 34–56 (2015)
3. 'EcoTwin' participating in the European Truck Platooning Challenge. https://www.daf.com/en/news-and-media/articles/global/2016/q1/22032016-ecotwin-participating-in-the-european-truck-platooning-challenge
4. European Truck Platooning Challenge (2016). https://www.rijkswaterstaat.nl/english/about-us/international-cooperation/the-netherlands-eu-presidency-2016/european-truck-platooning-challenge-2016.aspx

5. Martens, M.H., Van den Beukel, A.P.: The road to automated driving: dual mode and human factors considerations. In: Intelligent Transportation Systems-(ITSC), 16th International IEEE Conference, pp. 2262–2267 (2013)
6. Bainbridge, L.: Ironies of automation. Automatica **19**(6), 775–779 (1983)
7. Mackworth, N.H.: The breakdown of vigilance during prolonged visual search. Q. J. Exp. Psychol. **1**(1), 6–21 (1948)
8. Körber, M., Schneider, W., Zimmermann, M.: Vigilance, boredom proneness and detection time of a malfunction in partially automated driving. In: 2015 International Conference, Collaboration Technologies and Systems (CTS), pp. 70–76 (2015)
9. Willemsen, D., Stuiver, A., Hogema, J.: Automated driving functions giving control back to the driver: a simulator study on driver state dependent strategies. In: 24th International Technical Conference on the Enhanced Safety of Vehicles (ESV), No. 15-0109 (2015)
10. Olaverri-Monreal, C., Gomes, P., Fernandes, R., Vieira, F., Ferreira, M.: The see-through system: a VANET-enabled assistant for overtaking maneuvers. In: 2010 IEEE, Intelligent Vehicles Symposium (IV), pp. 123–128 (2010)
11. Gomes, P., Vieira, F., Ferreira, M.: The see-through system: from implementation to test-drive. In: 2012 IEEE Vehicular Networking Conference (VNC), pp. 40–47 (2012)
12. Belyaev, E., Vinel, A., Egiazarian, K., Koucheryavy, Y.: Power control in see-through overtaking assistance system. IEEE Commun. Lett. **17**(3), 612–615 (2013)
13. Patra, S., Arnanz, J.H., Calafate, C.T., Cano, J.C., Manzoni, P.: EYES: a novel overtaking assistance system for vehicular networks. In: International Conference on Ad-Hoc Networks and Wireless, pp. 375–389 (2015)
14. The Safety Truck Could Revolutionize Road Safety. https://news.samsung.com/global/the-safety-truck-could-revolutionize-road-safety
15. ISO, E. 15007-1: 2002. Straßenfahrzeuge-Messung des Blickverhaltens von Fahrern bei Fahrzeugen mit Fahrerinformations-und Fahrerassistenzsystemen-Teil 1: Begriffe und Parameter. Europäisches Komitee für Normung (2002)
16. Gold, C., Damböck, D., Lorenz, L., Bengler, K.: "Take over!" How long does it take to get the driver back into the loop? In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 57, no. 1, pp. 1938–1942, Sage CA: Los Angeles (2013)

Road and Rail: Vulnerable Road Users

Driver–Cyclist Interaction Under Different Bicycle Crossroad Configurations

Francesco Bella(✉) and Manuel Silvestri

Roma TRE University, Vito Volterra n. 62, 00146 Rome, Italy
{francesco.bella, manuel.silvestri}@uniroma3.it

Abstract. Objective of the present study was to analyze the effectiveness of countermeasures at bicycle crossroads during the driver–cyclist interaction by the use of a driving simulator. Three countermeasures were examined: baseline condition (absence of treatment), colored paved markings and raised island. Forty-two participants drove a suburban scenario with several bicycle crossroads having difference configurations and presence/absence of cyclist. 378 speed profiles were plotted from 150 m before each crossroads to obtain the explanatory variables of the driver behavior. A set of ANOVAs was used to analyze the driver behavior. Results showed that countermeasures affected in statistically significant way the distance at which the driver ends the braking maneuver, highlighting the better effectiveness of the colored pavement. A general improvement of the driver speed behavior was also observed for this countermeasure. The outcome of the questionnaire confirmed that it was considered as the most effective in terms of driving aid.

Keywords: Road safety · Simulation · Vulnerable road users · Bicycle crossroad · Safety countermeasures

1 Introduction

The use of the bicycle as valid transport alternative is continuously increased in last years. All around the world public agencies promote this kind of transport because of the potential environmental and health benefits [1, 2]. This needs to be taken into account in the mobility and traffic studies, as well as the potentialities and the risks related to the presence of these road users have to be considered. One of the main problems in the bicycle usage is the lack of the network accessibility and adequate paths, which should ensure high safety levels during the travel. More specifically, a critical point for cyclists' safety is the crossroad. At this location, the interactions with the motorized vehicles are more dangerous because of the high severity of the impact in case of accident, due to the limited protections and the low tolerance to biomechanical forces [3, 4]. European accident statistics [5] show that 31% of cyclist fatalities occur at intersections and about 56% of these fatalities are recorded at crossroads. Despite the problem is relevant the efforts to make safer the driver–cyclist interaction seem to be not enough. It is widely known that reduce the risks connected to the driver–cyclist interaction is linked to a good cycling infrastructure [e.g. 6, 7], which includes the bicycle crossroads. However, the problems concerning the driver–cyclist interaction that occur at the crossroads are

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_82

attributable both to driver's behavior and to the poor layout of the crossroad, which does not often induce adequate drivers' yielding behavior.

Räsänen and Summala [8] found that the driver yielding decision process is mainly affected by the timely detection and vision of the cyclist. They suggested that adequate vehicles approaching speed to intersection can improve the probability of a correct yielding behavior. Several other studies [e.g. 9–11], revealed that the speed is a crucial factor for the yielding decision process and main cause of serious and fatal outcomes in vehicle–bicycle accidents.

Therefore, according to the literature, interaction between vehicles and cyclists at bicycle crossroads are critical situations, in which the driver has to be influenced in order to adapt his speed if the cyclist is present, avoiding the occurrence of accidents and improving the cyclist safety. For these reasons, this paper aims at analyzing the driver behavior at bicycle crossroads under several countermeasures and at identifying the most effective treatment for bicycle crossroads.

Two countermeasures were analyzed: colored paved markings and raised island. The first one is bicycle crossroad in which the pavement is painted in red and its aim is to focus the attention of the approaching driver on it, contributing at adequate speed and yielding behavior [12]. The second treatment is a physical facility placed between the vehicle lanes, determining a relative narrowing and inducing the driver to adapt his speed [13]. Accomplishing these aims was possible by the use of a driving simulator that, mainly, allows risk avoidance for the experimenters and full control of the experimental conditions.

2 Method

The present study was carried out using the advanced fixed – base driving simulator of the Department of Engineering – Roma Tre University. It's widely known that that driving simulators are useful tools for the assessment of the driver's behavior as induced by the road configuration and road traffic conditions [e.g. 14–18]. More specifically, several studies [19–22] showed the high effectiveness of driving simulators to provide useful recommendations to improve the interaction between driver and road vulnerable users (pedestrians and cyclists).

2.1 Road Scenario

The simulated road scenario was a two-lane suburban road about 7.6 km long in which also the cyclist that crosses the road was simulated. 9 bicycle crossroads (3 for each one of the 3 types represented in Fig. 1) were placed along the alignment. For ensuring the same approaching conditions, a signalized intersection was present in advance of each bicycle crossroad. Each driver was obligated to stop at the signalized intersection, due to the red light that turned on when the driver was at approximately 100 m from the intersection. After the signalized intersection, the driver could then reach a consistent speed for the simulated scenario due to the 400 m between the intersection and the bicycle crossroad. The posted speed limit was 50 km/h while the cross-section was

10 m wide, formed by two 3.00 m wide lanes, two paved shoulders 0.50 m wide and two curbs 1.50 m wide, according to the Italian road design guidelines [23]. According to the Italian Highway Code [24] the bicycle crossroad was 1.50 m wide and three vertical signs were posted: one at 150 m in advance of the bicycle crossroad to signal the presence of it. One at 10 m, to advice the driver to yield if the cyclist was present and one in correspondence of the bicycle crossroad. This configuration represents the baseline condition (Fig. 1a). In addition to the baseline condition, two other possible configurations of the bicycle crossroad were placed in the simulated scenario: colored paved markings and raised island. The first was the bicycle crossroad provided with a pavement red painted (Fig. 1b), while the second was a physical facility 40 m long, 0.50 m wide and 0.05 m height (Fig. 1c), which causes the narrowing of the two vehicle lanes up to 2.75 m in the last 40 m in advance of the bicycle crossroad.

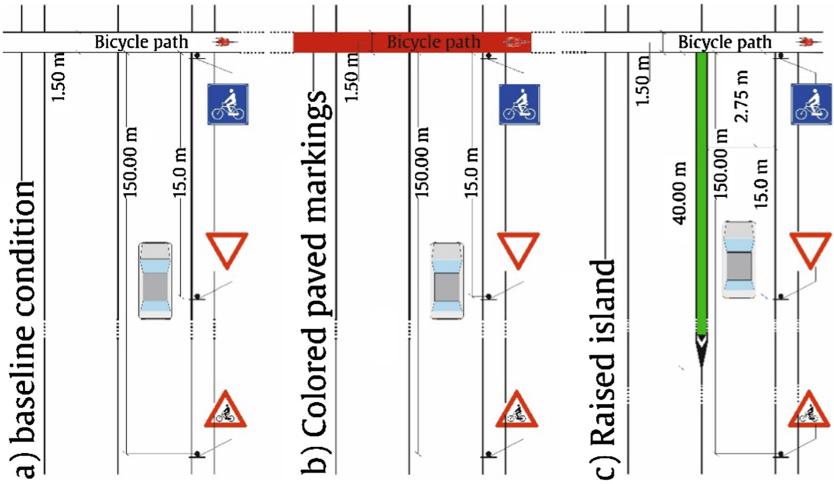


Fig. 1. (a) baseline condition (b) colored paved markings (c) raised island

At six bicycle crossroads (2 for each countermeasure) a cyclist coming from driver right side crossed the road. The cyclist was set to start the crossing at 20 m from the collision point when the driver was at 50 m from it. The speed of the cyclist was set equal to 20 km/h. Assuming the driver speed equal to 50 km/h, in such condition the driver and cyclist have the same time to arrive (3.6 s) to the collision point. It should be noted that this condition is representative of a theoretical driver–cyclist interaction, which occurs only if the driver reaches the hypothesized speed value at 50 m in advance of the bicycle crossing.

To avoid a potential order effect on the driver's behavior, 3 road scenarios that have a different sequence of the 9 combinations of bicycle crossroad (baseline condition, colored paved markings and raised island) x cyclist (cyclist absence and two conditions in which the cyclist was crossing the road) were implemented in the driving simulator. Each scenario was driven by one of the 3 groups into which the participants were divided (see next section on participants).

2.2 Apparatus

The interactive fixed-base driving simulator of the Department of Engineering – Roma Tre University was used for accomplishing the aims of the present study. The driving simulator was previously tested, calibrated, and validated [25, 26] as a reliable tool for the study of the driver's speed behavior. The hardware interfaces (wheel, pedals and gear lever) are installed on a real vehicle. The driving scenario is projected onto three screens: one in front of the vehicle and one on either side, ensuring a 135° field of view. The resolution of the visual scene is 1024 × 768 pixels with a refresh rate from 30 to 60 Hz. The vehicle of the driving simulator is also equipped with an audio system that reproduces the sounds of the engine.

2.3 Procedure

The driver did not encounter other vehicles in its own driving lane. To induce the driver not to drift to the incoming traffic lane, a slight amount of traffic was implemented along the opposing lane. The simulated vehicle was a standard medium-class car with automatic gears. The data recording system recorded all of the parameters at spatial frequency of 2 m. The experiment procedure consisted in: (a) communicating to the driver about the duration of the driving and the use of the simulated vehicle tools; (b) training at the driving simulator on a specific alignment approximately 5 km long to became familiar with the steering wheel, accelerator and brake pedal; (c) filling in a form with personal data, years of driving experience, average annual distance driven; (d) driving one of the three road scenarios with a specific bicycle crossroads-cyclist sequence; (e) filling in of a questionnaire about the perceived discomfort during the test, to eliminate from the sample driving performed under anomalous conditions, which could affect the data analysis. This questionnaire consisted of 5 questions, in which each question addressing a type of perceived discomfort: nausea, giddiness, daze, fatigue, other. A score from 1 to 4 in proportion to the level of experienced discomfort could be selected, equivalent to null, light, medium, and high. The null and light level for all four types of discomfort is considered to be the acceptable condition for driving; (f) filling in of a questionnaire about how the driver perceived the countermeasures. For this question, the participant could choose between a) driving aid and b) obstacle; for each of them, the driver could associate a score from 0 to 3 in proportion to the perceived effectiveness. Drivers were instructed to drive as they normally would in the real world.

2.4 Participants

Forty-two drivers (24 men and 18 women), whose ages ranged from 24 to 59 (mean = 29.3; SD = 8.5) and who had regular European driving licenses for at least three years were selected to participate at the experiment (mean = 11.0; SD = 8.0). According to outcomes of the questionnaire about the perceived discomfort, no participant was excluded from the sample. Thus, the number of participants used for the analysis consisted of all 42 drivers, which were divided into 3 groups; the 3 groups drove different scenarios, which were each characterized by a specific encounter order of bicycle crossroad/cyclist.

3 Data Processing

For the analysis of the divers' behavior under the presence/absence of cyclist condition and the countermeasure conditions the speed profiles of the last 150 m in advance of each one of the 9 bicycle crossroads along the alignment were plotted. Overall, 378 speed profiles were plotted (42 drivers × 9 bicycle crossroads), from which the following variables were determined:

- V_i: initial speed, is the speed identified at the moment when the driver starts to decrease his speed, in response to the cyclist that is crossing;
- L_{Vi}: initial speed distance, is the distance from the bicycle crossroad where V_i is recorded;
- V_m: minimum speed, is the minimum value of the speed during the driver braking maneuver;
- L_{Vm}: minimum speed distance, is the distance from the bicycle crossroad where V_m is recorded;
- d_{av}: average deceleration rate, is the average deceleration rate adopted by the driver during the entire braking maneuver obtained from the following equation

$$d_{av} = \left(V_i^2 - V_m^2\right) / \left[2(L_{Vi} - L_{Vm})\right]. \quad (1)$$

In addition, also the cases in which the driver did not yield to the cyclist were recorded; more specifically, in 12 cases for the baseline condition, 5 cases for the raised island and 6 cases for the colored paved markings the driver missed the yielding maneuver. These cases were excluded from the analysis because the driver did not experience the interaction with the cyclist. Thus, a total of 355 observations were used for the analysis of the drivers' behavior while approaching the bicycle crossroad.

4 Data Analysis and Results

A set of Analysis of Variance (ANOVA) was carried out to verify the effects due to the countermeasure condition (baseline condition, raised island and colored paved markings) and the cyclist condition (presence/absence) on the dependent variables obtained from the drivers' speed profiles (V_i, L_i, V_{min}, L_{Vmin}, d_{av}). In the Table 1 the descriptive statistics of the dependent variables for each combination of countermeasure condition by cyclist condition are reported.

4.1 Drivers' Initial Speed

The effects of the countermeasures on drivers' initial speed was not statistically significant ($F_{(2,349)} = 0.335$, $P = 0.715$); the mean value of the initial speed was 12.11 m/s under the baseline conditions, 11.78 m/s for the colored paved markings and 11.83 m/s for the raised island (Fig. 2a). Also the effect of the cyclist condition on the initial speed was not statistically significant ($F_{(1,349)} = 1.517$, $P = 0.219$). The mean value under the cyclist absence condition was 12.13 m/s while for the cyclist presence condition was 11.69 m/s (Fig. 2a).

Table 1. Descriptive statistics

Dependent variable	Countermeasure	Cyclist condition	Mean	SD
V_i [m/s]	Baseline	Cyclist absent	12.35	2.23
		Cyclist present	11.88	3.03
	Colored paved markings	Cyclist absent	12.17	3.39
		Cyclist present	11.39	2.63
	Raised island	Cyclist absent	11.86	2.58
		Cyclist present	11.80	2.72
L_i [m]	Baseline	Cyclist absent	61.16	15.61
		Cyclist present	40.23	10.40
	Colored paved markings	Cyclist absent	58.72	15.50
		Cyclist present	45.69	10.25
	Raised island	Cyclist absent	55.62	18.75
		Cyclist present	41.47	11.33
V_{min} [m/s]	Baseline	Cyclist absent	8.53	3.31
		Cyclist present	4.76	3.21
	Colored paved markings	Cyclist absent	8.08	3.82
		Cyclist present	4.88	2.82
	Raised island	Cyclist absent	7.52	3.70
		Cyclist present	4.89	3.05
L_{Vmin} [m]	Baseline	Cyclist absent	25.33	11.09
		Cyclist present	20.64	9.91
	Colored paved markings	Cyclist absent	28.29	17.02
		Cyclist present	26.07	10.93
	Raised island	Cyclist absent	26.39	15.22
		Cyclist present	20.84	11.23
d_{av} [m/s^2]	Baseline	Cyclist absent	1.20	1.06
		Cyclist present	2.75	1.56
	Colored paved markings	Cyclist absent	1.37	1.00
		Cyclist present	2.58	1.58
	Raised island	Cyclist absent	1.59	1.60
		Cyclist present	2.64	1.70

4.2 Distance from the Bicycle Crossroad Where the Yielding Maneuver Begins

Results showed a non-significant effect of the countermeasures on the variable L_{Vi} ($F_{(2,349)} = 1.775$, $P = 0.171$). For the baseline condition the mean value was 50.63 m, while for the colored paved markings and the raised island the mean values were 52.20 and 48.75 m, respectively (Fig. 2b). Conversely, the effect of the cyclist condition was statistically significant ($F_{(1,349)} = 97.854$, $P = 0.000$). The pairwise comparison showed that the point in which the driver started to decrease the speed was statistically significantly further from the bicycle crossroad when the cyclist was present (mean value = 58.50 m) than that for the cyclist absence condition (mean difference = 16.04 m; $P = 0.000$) (Fig. 2b).

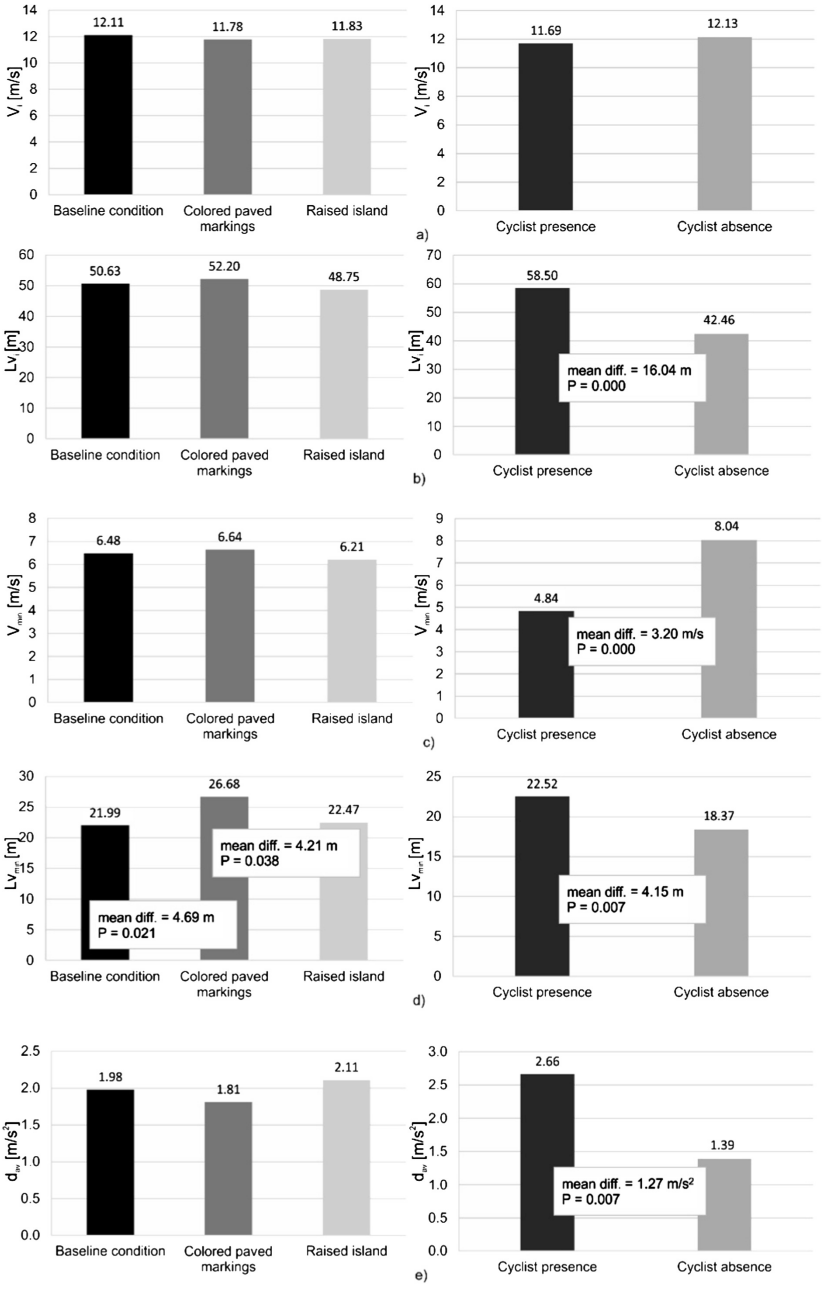


Fig. 2. Effects of countermeasures and cyclist conditions on: (a) Drivers' initial speed; (b) Distance from the bicycle crossroad where the yielding maneuver begins; (c) Drivers' minimum speed; (d) Distance from the bicycle crossroad where the yielding maneuver ends; (e) Average deceleration rate

4.3 Drivers' Minimum Speed

The analysis showed that the effect of the countermeasures on drivers' minimum speed was not statistically significant ($F_{(2,349)} = 0.406$, $P = 0.667$). For the baseline condition the mean value was 6.48 m/s, while for the colored paved markings and the raised island the mean values of the variable V_{min} were 6.64 m/s and 6.21 m/s, respectively (Fig. 2c). The effect of the cyclist condition affected the drivers' minimum speed in a statistically significantly way ($F_{1,349)} = 61.108$, $P = 0.000$); the pairwise comparison showed that for the cyclist present condition (mean value = 4.84 m/s) the minimum speed was statistically significantly lower than that for the cyclist absent condition (mean difference = 3.20 m/s, $P = 0.000$) (Fig. 2c).

4.4 Distance from the Bicycle Crossroad Where the Yielding Maneuver Ends

The effect of the countermeasure condition on the variable L_{Vmin} was statistically significant ($F_{2,349)} = 4.586$, $P = 0.011$); post-hoc analysis revealed that the drivers ended the yielding maneuver farther from the bicyclist crossroad under the condition of colored paved markings (mean value = 26.68 m), while for the baseline condition and the raised island the mean values were statistically significantly less (mean difference = 4.69 m, $P = 0.021$; mean difference = 4.21, $P = 0.038$). No other difference was statistically significant (Fig. 2d). The effect due to the cyclist condition was also statistically significant ($F_{1,349)} = 7.366$, $P = 0.007$). Results showed that for the cyclist present condition (mean value = 22.52 m) the variable L_{Vmin} was statistically significantly higher than that for the cyclist absence condition (mean difference = 4.15 m, $P = 0.007$) (Fig. 2d).

4.5 Average Deceleration Rate

The statistical analysis revealed that the effect of the countermeasure on drivers' average deceleration rate was not statistically significant ($F_{(2,349)} = 0.220$, $P = 0.802$); however it should be noted that the lower mean value was obtained for the colored paved markings (1.81 m/s^2); higher mean values were recorded for the baseline condition (1.98 m/s^2) and for the raised island (2.11 m/s^2) (Fig. 2e). Conversely, the effect of the cyclist condition was statistically significant ($F_{(1,349)} = 42.675$, $P = 0.000$). The pairwise comparison showed that the average deceleration rate when the cyclist was present (mean value = 2.66 m/s^2) was statistically significantly higher than that for the cyclist absence condition (mean difference = 1.27 m/s^2, $P = 0.000$) (Fig. 2e).

4.6 Outcomes of the Questionnaire

The results of the questionnaire showed that 71% of the drivers (30 of 42) selected the highest score of "driving aid" for the colored paved markings, while only 19% of the drivers (8 of 42) reported the same score of "driving aid" for the raised island. The highest

score of "driving aid" for the baseline condition was reported by 40% of the drivers (17 of 42). According to this outcome, the 79% (33 of 42) of the drivers reported the lower score of the "obstacle driving effect" for the colored paved markings, while only the 31% (13 of 42) of the drivers reported the lower score of the "obstacle driving effect" for the raised island. The lower score of the "obstacle driving effect" was reported by 76% of driver (32 of 42) under the baseline condition.

5 Discussion

5.1 Effects of the Countermeasures

Statistical analyses indicated that the variables V_i, L_{Vi}, V_{min} and d_{av} were not affected in a statistically significantly way by the countermeasure condition. However, an improvement of the drivers' speed behavior was observed.

The effects on initial speed was not statistically significant and this result was consistent with the expected driver behavior; the approaching speed could not be affected by the presence of the countermeasure at the bicycle crossroad because the driver did not perceive them when he was far from the bicycle crossroad.

The variable L_{Vi}, which is the distance from the bicycle crossroad where the driver started the yielding maneuver, showed that under the colored paved markings the driver started to decrease the speed 1.57 and 3.45 m before compared to the baseline condition and the raised island. This result is consistent with the outcomes of Hallmark et al. [12] and Campbell et al. [13], which showed that a painted bicycle crossroad was clearly perceived by the drivers, which were better advised about the possible presence of a cyclist that could cross. In this way, the driver anticipated the yielding maneuver making safer the driver–cyclist interaction. In addition, such results are consistent with the study of Leden et al. [27], which reported that the bicycle crossroad provided with colored pavement can influence the drivers' visual search and decrease the probability of unexpected situations, improving the cyclist safety.

The effect due to the advance of the yielding maneuver was clearly observed on L_{Vmin}, which was statistically significantly affected by the countermeasure condition. Results showed that the driver ended the yielding maneuver farther from the bicycle crossroad in presence of colored paved markings (26.68 m) compared to the baseline condition (21.69 m) and the raised island (22.47 m). This means that the distance between the cyclist that was crossing and the driver that was yielding increased when the bicycle crossroad was provided with the colored paved markings. From the cyclist safety point of view this result is consistent with the aim of the countermeasure, which is improving the driver performance and, thus, the driver–cyclist interaction.

Such results confirm the outcome of a previous study [22], which revealed that the colored paved markings at the bicycle crossroads induced the drivers to start to reduce the speed farther from the collision point with the cyclist ending in advance the yielding maneuver.

The effects of the countermeasures on d_{av}, even if not statistically significant, can be also linked to the possibility of the driver to advance the yielding maneuver for the colored paved markings undertaking lower deceleration rates (the average deceleration

rate for the colored paved markings was 1.81 m/s^2, while higher values were recorded for the baseline condition and for the raised island).

5.2 Effects of the Cyclist Condition

Statistical analysis showed that the cyclist condition affects the variables L_{Vi}, V_{min}, L_{Vmin} and d_{av}, while the initial speed was not affected in statistically significantly way. The result on V_i was expected because the adopted approaching speed by the driver was independent from the presence of the cyclist. In other words, the driver did not perceived the presence of the cyclist until he reached a consistent distance from the bicycle crossroad that allowed him to see the crossing cyclist. The results on the variables L_{Vi}, V_{min}, L_{Vmin} and d_{av} were also expected and were fully consistent with the driver behavioral models in literature [28, 29].

For the cyclist absence condition, the driver moves forward the point in which he starts to decrease the speed because no yielding maneuver is needed; when the cyclist was absent, in fact, the lowest value of L_{Vi} was recorded (42.46 m).

For the minimum speed, results showed also that the higher value of V_{min} was reached when the cyclist was absent (8.04 m/s). This means that the driver does not need to slow down as much as when the cyclist is crossing, because he does not perceive the interaction and he has not to yield.

The driver moves forward also the point in which the minimum speed is recorded in the condition of cyclist absence (18.37 m); as for the effect on L_{Vi}, the driver when does not perceive the presence of the cyclist, delays the braking maneuver and ends the deceleration phase nearer to the bicycle crossroad.

The average deceleration rates (d_{av}) confirm how the driver reacts when he perceives the interaction with the cyclist. In the cyclist present condition, the average deceleration rate reached the highest value (2.66 m/s^2). As expected, the need of yield or avoid the potential conflict, induces the driver to adopt a more abrupt maneuver than that of the cyclist absence condition. For this last cyclist condition, the driver does not experience an interaction with the cyclist and, therefore, he performs a smoother maneuver.

6 Conclusions

The present study aimed at the assessment of the potential safety effects of countermeasures at the bicycle crossroads in terms of drivers' speed behavior in response to a cyclist that crosses the road. The safety countermeasures implemented in the driving simulator scenario were the colored paved markings and the raised island.

The results showed that the effect of the countermeasures was statistically significant only for the variable L_{Vmin}. However, the general trend becoming from also the other variables (L_{Vi}, V_{min}, L_{Vmin} and d_{sv}) shows an improvement of the driver speed behavior in the presence of the colored paved markings. In this condition, the driver started the yielding maneuver at a distance that was higher (non-statistically significant) compared to the baseline condition (50.63 m) and the raised island (48.75 m). This

capacity of the driver to advance the yielding maneuver due to the better visibility provided by the colored pavement, allowed him to end the yielding maneuver at a distance from the bicycle crossroad that was higher (statistically significant) than that for the baseline condition (21.69 m) and the raised island (22.47 m), increasing the cyclist safety during the driver–cyclist interaction at the bicycle crossroad. Moreover, these outcomes implied that the driver adopted a smoother maneuver characterized by less deceleration rates.

In addition, also an improvement of the yielding compliance was observed for the countermeasures. Compared to the baseline condition, for the colored paved marking and the raised island only in 2.6% (6 of 229) and in 2.2% (5 of 229) of the driver–cyclist interactions the driver did not yield, while when the countermeasures were absent the failed yielding rate was 5.9% (12 of 229).

The general improvement of the driver speed behavior and the driver yielding behavior due to the presence of the colored paved markings was also emphasized by the outcomes of the questionnaire, which revealed a greater driver acceptance of this countermeasure which did not cause a "physical obstruction", as occurred for the raised island.

Further studies might analyze the driver–cyclist interaction under different cyclist condition as groups of cyclists, with the aim of assessing the effectiveness of the countermeasures in the most common cyclist traffic conditions and highlighting additional effects on driver behavior.

References

1. Edwards, R.D., Mason, C.N.: Spinning the wheels and rolling the dice: life-cycle and benefits of bicycling commuting in the US. Prev. Med. **64**, 8–13 (2014)
2. Fishman, E., Washington, S., Haworth, N.: Barriers and facilitators to public bicycle scheme use: a qualitative approach. Transp. Res. F. **15**, 686–698 (2012)
3. OECD/ITF. Cycling, Health and Safety: OECD Publishing/ITF. (2013)
4. Räsänen, M., Summala, H.: Attention and expectation problems in bicycle–car collisions: an in-depth study. Accid. Anal. and Prev. **30**, 657–666 (1998)
5. ERSO, European Road Safety Observatory: Traffic Safety Basic Facts – Cyclist (2015)
6. Scheepers, P., Hagenzieker, M., Methorst, R., van Wee, B.: A conceptual framework for road safety and mobility applied to cycling safety. Accid. Anal. Prev. **62**, 331–340 (2014)
7. Wegman, F., Zhang, F., Dijkstra, A.: How to make more cycling good for road safety? Accid. Anal. Prev. **44**, 19–29 (2012)
8. Räsänen, M., Summala, H.: Car drivers' adjustments to cyclists at roundabouts. Transp. Hum. Factors **2**, 1–17 (2000)
9. Silvano, A.P., Koutsopoulos, H.N., Ma, X.: Analysis of vehicle bicycle interactions at unsignalized crossings: a probabilistic approach and application. Accid. Anal. Prev. **97**, 38–48 (2016)
10. Kim, J.K., Kim, S., Ulfarsson, G.F., Porrello, L.A.: Bicyclist injury severities in bicycle-motor vehicle accidents. Accid. Anal. Prev. **39**, 238–251 (2007). doi:10.1016/j.aap.2006.07.002
11. Minikel, E.: Cyclist safety on bicycle boulevards and parallel arterial routes in Berkeley, California. Accid. Anal. Prev. **45**, 241–247 (2012). doi:10.1016/j.aap.2011.07.009

12. Hallmark, S.L., Peterson, E., Fitzsimmons, E., Hawkins, N., Resler, J., Welch, T.: Evaluation of gateway and low cost traffic calming measures for major routes in small rural communities. Center for Transportation Research and Education, Iowa State University (2007)
13. Campbell, J.L., Richard, C.M., Graham, J.: NCHRP Report 600 B: Human Factors Guidelines for Road Systems. TRB, Washington, DC (2008)
14. Bella, F.: New model to estimate speed differential in tangent-curve transition. Adv. Transp. Stud. Int. J. **2008**(15), 27–36 (2008). doi:10.4399/97888548186823
15. Bella, F.: How traffic conditions affect driver behavior in passing maneuver. Adv. Transp. Stud. Int. J. (Spec. Issue): 113–126 (2011). doi:10.4399/978885484657911
16. Bella, F.: Driver perception of roadside configurations on two-lane rural roads: Effects on speed and lateral placement. Accid. Anal. Prev. **50**, 251–262 (2013). doi:10.1016/j.aap.2012.04.015
17. Bella, F.: Driver performance approaching and departing curves: driving simulator study. Traffic Injury Prev. **15**, 310–318 (2014). doi:10.1080/15389588.2013.813022
18. Bella, F.: Operating speeds from driving simulator tests for road safety evaluation. J. Transp. Saf. Secur. **6**, 220–234 (2014). doi:10.1080/19439962.2013.856984
19. Duivenvoorden, K., Hogema, J., Hagenzieker, M., Wegman, F.: The effects of cyclists present at rural intersections on speed behavior and workload of car drivers: a driving simulator study. Traffic Injury Prev. **16**, 254–259 (2015). doi:10.1080/15389588.2014.937484
20. Bella, F., Silvestri, M.: Effects of safety measures on driver's speed behavior at pedestrian crossings. Accid. Anal. Prev. **83**, 111–124 (2015). doi:10.1016/j.aap.2015.07.016
21. Bella, F., Silvestri, M.: Driver's braking behavior approaching pedestrian crossings: a parametric duration model of the speed reduction times. J. Adv. Transp. **2016**(50), 630–646 (2016). doi:10.1002/atr.1366
22. Bella F., Silvestri M.: Drivers' behavior at bicycle crossroads. Ed. Springer. Advanced concepts, methodologies and technologies for transportation and logistics (2017, in press)
23. Ministry of Infrastructures and Transports: Decreto Ministeriale del 5/11/2001 Norme funzionali e geometriche per la costruzione delle strade, Roma (2001). (in Italian)
24. Ministry of Infrastructures and Transports: Nuovo codice della strada, D.L. 30/4/1992 n.285, Gazzetta Ufficiale Repubblica Italiana, n. 114 (1992). (in Italian)
25. Bella, F.: Validation of a driving simulator for work zone design. Transp. Res. Rec.: J. Transp. Res. Board **1937**, 136–144 (2005). doi:10.3141/1937-19
26. Bella, F.: Driving simulator for speed research on two-lane rural roads. Accid. Anal. Prev. **40**, 1078–1087 (2008). doi:10.1016/j.aap.2007.10.015
27. Leden, L., Garder, P., Pulkkinen, U.: An expert judgement model applied to estimating the safety effect of a bicycle facility. Accid. Anal. Prev. **32**, 589–599 (2000)
28. Fuller, R.: A conceptualization of driver behavior as threat avoidance. Ergonomics **27**, 1139–1155 (1984)
29. Fuller, R.: Predicting what a driver will do: implication of the threat avoidance model of driver behavior – road user behavior. Theory and research. In: Proceedings of 2nd International Conference on road safety. Groningen, Netherlands (1987)

Operational and Safety Analysis of Signage and Pavement Marking Treatments in Puerto Rico Dynamic Toll Lane Using a Driving Simulator

Bryan Ruiz(✉), Didier Valdés, Benjamín Colucci, Johnathan Ruiz, Ricardo García, and Enid Colón

Department of Civil Engineering and Surveying,
University of Puerto Rico at Mayagüez (UPRM), P.O. Box 9000, Mayagüez, PR 00698, USA
{bryan.ruiz, didier.valdes, benjamin.coluccil, johnathan.ruizl, ricardo.garcia5, enid.colonl}@upr.edu

Abstract. The Puerto Rico Dynamic Toll Lane (DTL) is a managed lane system for light traffic that combines congestion pricing, reversible lanes and a Bus Rapid Transit system within the median of freeway PR-22. This research aims to evaluate if modified signage and the addition of in-lane pavement markings could help improve operational and safety aspects of the DTL using the UPRM driving simulator. The performance measures used to evaluate driving behavior were vehicle average position, average speed and acceleration noise, which has been used as a surrogate measure to evaluate safety. A total of four virtual scenarios were studied using time of day and traffic flow as independent variables to evaluate two configurations: existing conditions (Configuration 1) and proposed conditions using modified signs and pavement marking treatments (Configuration 2). However, results of the study indicates that there were no significant differences between configurations for the performance measurements evaluated.

Keywords: Driving simulator · Human factors · Transportation safety · Traffic control devices · Managed lanes

1 Introduction

Managed lanes have been used over the past decades to address traffic congestion and travel delay in the United States. This type of transportation system strategy is feasible for improving highway capacity and travel time in highways segments, where traffic congestion affects social and economic development of metropolitan regions with limited right-of-way and budget restrictions. The Federal Highway Administration (FHWA) defines managed lanes as highway facilities or a set of lanes in which operational strategies are implemented and managed in response to real-time data, where traffic elements can be variated using pricing, access control and vehicle eligibility [1]. Examples of managed lanes includes High-Occupancy Vehicle (HOV) lanes,

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_83

High-Occupancy Toll (HOT) lanes, reversible lanes and express lanes. One recent managed lane that is operating in the Commonwealth of Puerto Rico is the reversible Dynamic Toll Lane (DTL), located in freeway PR-22.

As a consequence of the increase in traffic volume at the Metropolitan region of Puerto Rico, this reversible DTL was designed and constructed by the Puerto Rico Highway and Transportation Authority (PRHTA) and operated under a Public Private Partnership with Autopistas Metropolitanas de Puerto Rico, LLC (Metropistas). This managed lane system uses congestion pricing to provide an alternative road to avoid the traffic congestion in the main lanes of the PR-22 tollway. The Dynamic Toll Lane, also known as express lane, is located in the median of freeway PR-22 and has two travel lanes that are shared between automobiles under controlled light traffic conditions and a Bus Rapid Transit (BRT) system. This managed lane operates as a reversible lane, with traffic going in the Eastbound (EB) direction between 5:00 AM and 11:00 AM and in the Westbound (WB) direction from 1:30 PM to 9:00 PM [2]. However, high speeds and incorrect use of exit lanes may be affecting operational and safety aspects of this new managed lane system.

The average speed of vehicles inside the system is usually above the posted speed limit, which is 40–45 mph in the westbound direction and 40 mph in the eastbound direction. In addition, a considerable amount of drivers use the incorrect exit by traveling through the BRT exit lane. A recent study analyzed the effect of lane width, posted speed limit and time of day in PR-22 DTL using the UPRM driving simulator and found that subject drivers used the incorrect exit lane in 26% of the evaluated scenarios. This suggest that the existing Traffic Control Devices (TCDs) should be evaluated and compared with potential treatments that may improve drivers understanding and reduce driving confusion prior to the DTL exit. Signage configuration in the EB direction (Express Lane exit to the right) is different than that of the WB direction (Express Lane exit to the left); therefore, the location of the exit only in yellow background sign varies among directions (Fig. 1).

(a) Eastbound Direction (b) Westbound Direction

Fig. 1. Existing signage display for EB and WB directions

The fact that drivers are wrongly using the BRT exit instead of the freeway exit, has raised serious safety concerns. Furthermore, when drivers realize that they have taken the wrong exit and, considering the light traffic conditions, decide to stop and carefully drive in reverse until they reach a point where they can manage to switch into the correct exit lane. This situation present the potential to generate serious issues with the operation of the BRT if an automobile is found stopped or trying to drive in reverse in the bus-only exit lane.

2 Objective

The objective of this study is to evaluate if the proposed signage modifications and the addition of in-lane pavement markings improve driving behavior of subject drivers that were exposed to scenarios of the PR-22 DTL system when compared with the existing TCD's. The addition of pavement markings, such as word and symbols, can complement guidance signs located throughout the managed lanes and reduce the variability in lane position, average speed and acceleration noise prior to the exit. Therefore, a reduction in driving confusion when choosing the exit lane could minimize last moment lane changes and the variability in speed and acceleration, resulting in road safety improvements.

3 Hypothesis

The hypothesis tested in this research project is the following:

> *"Subject drivers exposed to scenarios with the proposed treatments will have a better driving performance, in terms of lane position, average speed and acceleration noise, than those drivers exposed to the existing conditions."*

4 Methodology

The methodology followed in this study is composed by seven tasks. These are the following: literature review, experimental design, development of scenarios, subject drivers, data collection, analysis and conclusions.

4.1 Literature Review

The literature review of this research was divided into three areas of interest, namely, signage, pavement markings and driving simulators.

Signage. The Manual on Uniform Traffic Control Devices (MUTCD) indicates that the purpose of road signs is to provide information of regulations, warnings and guidance to all road users [3]. The uniformity in TCDs is important to improve road safety and reduce potential conflict points in transportation facilities [4]. The design and location of TCDs shall take in consideration the Perception-Response Time (PRT) of drivers, which is defined as the time required by drivers in order to detect, recognize, make a decision and perform an action given a certain situation (MUTCD, 2009). The design of signs included in the MUTCD are provided in the supplementary book Standard Highway Signs and Markings, which includes dimensions, shape, colors, legends, borders and illumination [5]. Guidance signs for applicable preferential lanes such as PR-22 DTL should have white legend with green background whereas regulatory signs should have black legend with white background [3].

Pavement Markings. Pavement markings provide additional guidance and information to road users while reducing driver distraction from the roadway [3]. Markings can complement signs by providing more information of roadway conditions and guidance without creating a major diversion since drivers are supposed to pay attention to the road ahead. For example, transverse markings such as word, symbol and arrows should be proportioned to provide visibility at least equal to longitudinal lines to provide a proper view angle of visibility. Previous studies that have evaluated pavement markings treatments (such as centerlines, edge lines, wider markings, and words for dedicated electronic toll collection lanes) encountered a decrease in lane change, crashes and an improvement in drivers' decision making when choosing the desire lane with anticipation prior to toll plazas [6, 7]. Therefore, the addition of words, symbols and route shields markings for preferential lanes may reduce driving confusion at divergent exit segments similar to PR-22 managed lane.

Driving Simulators. Driving simulators have been used as an innovative and cost-effective research tool to evaluate the behavior of drivers in a wide diversity of research fields such as human factors, transportation, psychology, medicine, computer science, training and other driving activities [8]. There are several types of driving simulation systems, where hardware and software varies according to the visual display and vehicle components. For example, desktop simulators are non-motion simulators composed by a set of screens or monitors, steering wheel, acceleration/brake pedals, gearshift and a driver seat. Contrarily, motion simulators are more complex systems that have a greater visual display and several degrees of freedom that permits vehicle movements similar to those experienced in a real driving experience. However, both type of simulators are useful for evaluating existing and emerging transportation treatments without exposing subject drivers to physical harm in scenarios where a potential of crash may exist. Several driving simulators studies have used lane position, speed and acceleration data as performance measures to evaluate driving behavior. Representative examples of these studies include driver distraction when using cell phones, the evaluation of overhead freeway exit signs and the effect of lane width in mountain highways [9–11]. Furthermore, other transportation studies have adopted the variable acceleration noise as a surrogate measurement for evaluating safety aspects of transportation systems [12].

UPRM Driving Simulator Equipment. The UPRM driving simulator is composed of three main features. The first component is the driving cockpit, which includes of a driver chair, steering wheel, acceleration/brake pedals, gear shift and turning signals. The second element features the visual display provided by three fixed projectors and a set of screens that provides a panoramic view of the virtual scenario with 120 ° of visibility. The last component is the computer hardware of a laptop and desktop computer with an NVidia GTX 980 graphic card, surrounding audio system and the simulation software of Realtime Technologies Inc. (RTI).

4.2 Experimental Design

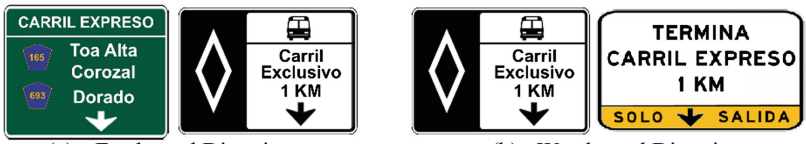
The experimental design of this research study consisted of 6 scenarios that represented the existing traffic control devices of PR-22 DTL (Configuration 1) and another 6 scenarios based on the proposed signage modifications and the addition of in-lane pavement markings (Configuration 2). These scenarios were developed taking into consideration two independent variables, namely, *time of day* and *traffic flow*. As described in Table 1, *time of day* was evaluated at three levels (morning in EB direction and afternoon and nighttime in WB direction) whereas *traffic flow* had two levels (no traffic and low traffic). Am ANOVA analysis was performed to evaluate significant differences among the evaluated variables. The order in which scenarios were exposed to each subject driver was based in the latin square design, with the purpose of reducing the effect of scenario order in the subject drivers.

Table 1. Description of scenarios developed for each configuration

Scenario	Time of day	Traffic flow
1	Morning (EB)	None
2	Morning (EB)	Low
3	Afternoon (WB)	None
4	Afternoon (WB)	Low
5	Nighttime (WB)	None
6	Nighttime (WB)	Low

4.3 Development of Scenarios

A total of six scenarios used for this research were developed for two Configurations of PR-22 DTL. The scenarios of Configuration 1 were created based on the design plans provided by the operating agency and video data that was recorded with a dash camera throughout the managed lane to identify and verify the location of TCDs. The scenarios for Configuration 2, which included the proposed treatments using modified signs and in-lane pavement markings (words and symbols), were developed based on the MUTCD and the supplementary manual of TCDs for Spanish text, signs and symbols in Puerto Rico [3, 13]. The modifications included the addition of the word EXPRESS LANE at the top of the guide sign and the diamond shape symbol used for HOV lanes as illustrated in Fig. 2.



(a) Eastbound Direction (b) Westbound Direction

Fig. 2. Proposed signage configuration for EB and WB directions

Furthermore, BUS ONLY markings were combined with the diamond shape symbol for the bus lane whereas the PR-22 route shield was applied in the express lane exit (Fig. 3). Following the MUTCD recommended design for freeways, three sets of markings were spaced at a 1000 ft interval, prior the BRT exit lane [3].

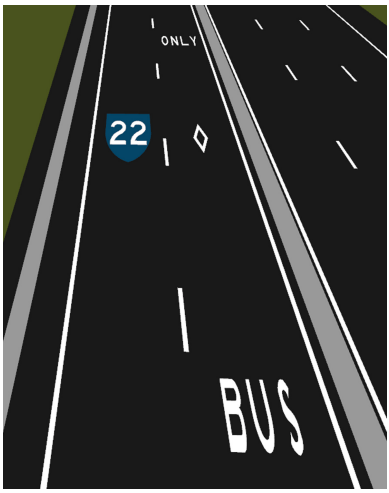


Fig. 3. Proposed in-lane markings for PR-22 DTL

4.4 Subject Drivers

All subject drivers participated in this study as volunteers that had to meet with an established criterion that consisted in having more than 18 years of age with at least 18 months of driving experience, a valid driving license and no health conditions that could affect their driving behavior. As part of the study protocol, participants had to fill an Informed Consent Form (which was approved by the Institutional Review Board of the University) and a pre-study questionnaire that included demographic information and driving history. Each component of the cockpit simulator was explained to each subject driver and a test scenario was provided in order for participants to familiarize with the steering wheel, gearshift and pedals of the simulator. In addition, participants were instructed to assume that the vehicle they were driving was a rental, where the characteristics of the car are different from their own. The experimental scenarios were presented once drivers felt comfortable driving the simulation equipment.

A total of 24 subject drivers (83% males and 17% females) participated in this research. The driving sample was composed of students, faculty members and administrative employees of UPRM. The age range of participants for Configuration 1 was between 19 and 59 years with an average age of 25.92 years and a standard deviation of 10.92 years. Similarly, participants for Configuration 2 had an age range

between 18 and 58 years with an average age of 27.50 years and a standard deviation of 11.04 years. Furthermore, 67% of the participants have not used PR-22 DTL. Nevertheless, 50% of these subject drivers recognized that an express lane exists in PR-22.

4.5 Data Collection

The performance measures taken into consideration for this research were the vehicle average position, average speed and acceleration noise [12, 14]. These dependent variables were determined at different zones of interest that were selected a posteriori, since the data was recorded for all trajectories of each participant. Four zones of interest prior the end of the PR-22 DTL virtual scenario were selected to compare each dependent variable. In addition, subject drivers' demographic information was gather through the pre-simulation questionnaires.

4.6 Analysis of Results

An ANOVA analysis was performed at a 95% confidence level to evaluate driving behavior of participants when exposed to representative scenarios of PR-22 DTL in the Westbound direction. The F-test and the ANOVA analysis have been used in several driving simulator studies, where comparisons of vehicle position, average speed and acceleration noise between scenarios and configurations are performed to identify statistical significant differences [12, 15, 16]. If the variation of the performance measurements is lower for the proposed treatments when compared with the current condition in the PR-22, then a safety improvement has been fulfilled. Therefore, the statistical analysis was performed for the dependent variables average position, mean speed and acceleration noise for four continuous zones of interest. The whole region of analysis starts 1 km prior the Westbound exit lane (Zone 1) and ends at the separation between the express lane exit and the BRT exit lane (Zone 4), where drivers can still manage to change lanes.

The results of this research indicates that there were no significant differences for variables average position and average speed for all zones of interest between subject drivers that drove through scenarios in Configuration 1 in WB direction (existing conditions of PR-22 DTL) and subject drivers that were exposed to Configuration 2 in WB direction (proposed conditions). Vehicle trajectories and average speed of each subject driver are illustrated in Figs. 4 and 5. Still, one participant exited through the BRT lane in Configuration 2 in scenarios that had night conditions (Scenario 5 and 6 in Fig. 4).

In terms of average speed, the driving sample tend to exceed the posted speed limits in all evaluated scenarios. This result is similar to the current situation, where users of the express lane drive above speed limits. However, higher speed profiles could be associated with the fact that the posted speed limit in PR-22 mainline is higher

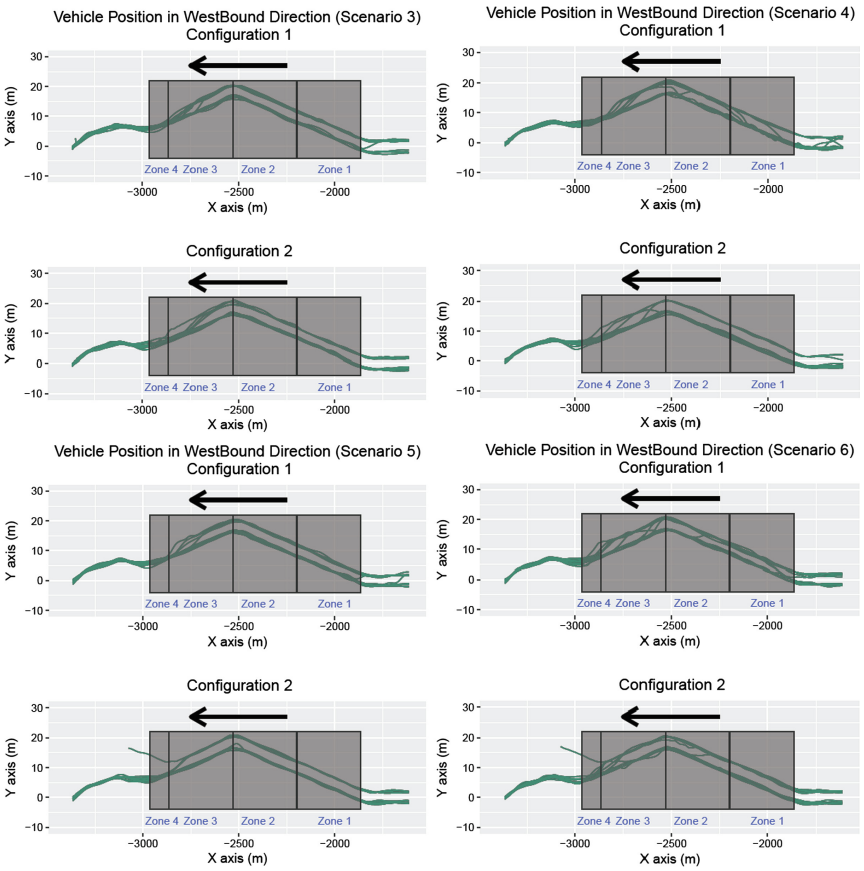


Fig. 4. Vehicle position of subject drivers

(55 mph) than the DTL lanes (40–45 mph). Furthermore, light traffic and geometric aspects of the DTL could influence drivers' perception of safety when traveling at high speeds.

In addition, modified signs and in-lane pavement markings did not had an effect in drivers' lane decision when approaching the express lane end. Therefore, their behavior was similar to participants in Configuration 1, where subjects changed lane almost at the last available moment. This indicates that other modifications that captures driver's attention should be considered for both signs and markings.

The performance measurement acceleration noise proved to be significant for the four zones of interest. However, as illustrated in Table 2, Zone 4 had significant differences with variable traffic flow, the interaction between configuration and time of day (two-way interaction) and configuration, time of day and traffic flow (three-way interaction).

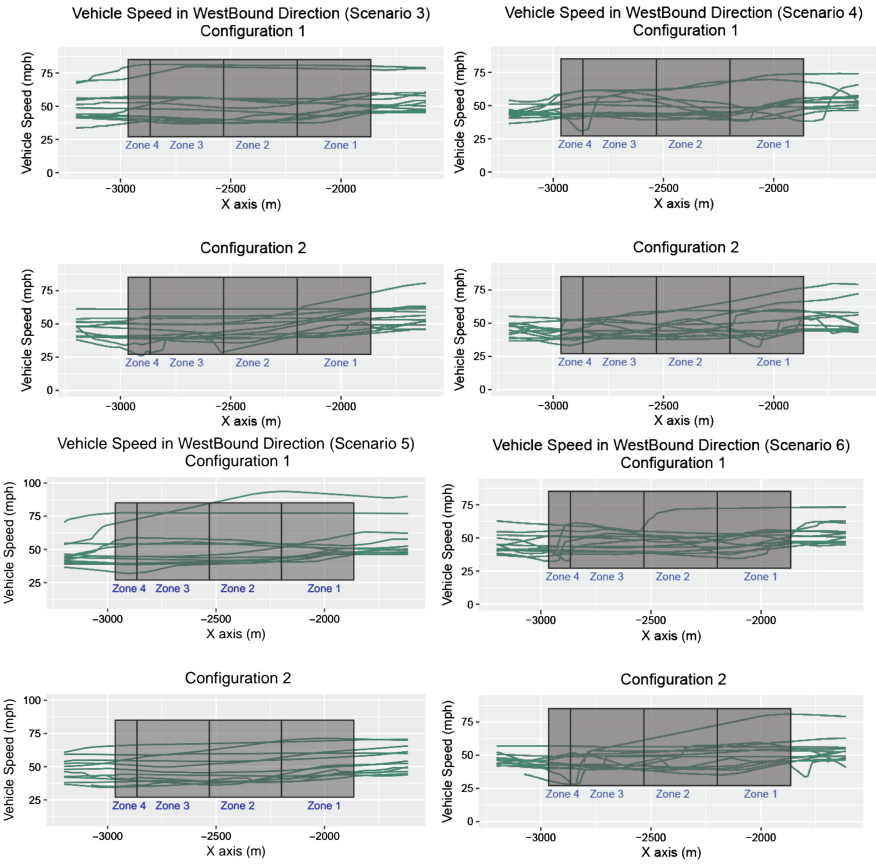


Fig. 5. Vehicle speed of subject drivers

4.7 Conclusions

The following conclusions are based on the results of statistical analysis performed for the WB direction of the DTL.

- The proposed treatments using modified signs and the addition of in-lane pavement markings did not have a significant effect in drivers' lane position. There were subject drivers that drove through the lane with the BUS ONLY marking and change lanes when approaching the divergent segment at the DTL exit. Furthermore, one subject driver exposed to Configuration 2 used the incorrect exit lane in two scenarios, where time of day condition was nighttime.
- There was no significant differences in drivers' average speed between configurations.
- Dependent variable acceleration noise had a significant difference for all zones of interest, especially in Zone 4 where it had two-way and three-way interactions.

Table 2. Analysis of variance for acceleration noise in Zone 4.

Source	DF	Sum of squares	Mean square	F value	*P* value
Model	7	2.07600966	0.29657281	2.87	0.0096
Error	88	9.09877533	0.10339517		
Corrected total	95	11.17478499			

R-Square	Coeff Var	Root MSE
Mean acceleration noise		
0.185776	242.6068	0.321551
0.13254		

Source	DF	Anova SS	Mean square	F value	*P* value
Configuration	1	0.04369353	0.04369353	0.42	0.5173
TimeofDay	1	0.02235755	0.02235755	0.22	0.6431
Configuration*TimeofDay	1	0.57686858	0.57686858	5.58	0.0204*
TrafficFlow	1	0.82469227	0.82469227	7.98	0.0059*
Configuration*TrafficFlow	1	0.03924419	0.03924419	0.38	0.5394
TimeofDay*TrafficFlow	1	0.01190095	0.01190095	0.12	0.7352
Configuration*TimeofDay*TrafficFlow	1	0.55725259	0.55725259	5.39	0.0226*

- Subject drivers could not be paying the required attention to traffic control devices, which results in driving confusion and a safety concern for both users and the operating agency.
- Cultural aspects and the lack of awareness regarding PR-22 DTL could have influenced driving behavior since a considerable percent (67%) of subject drivers have never used the system.

4.8 Future Work

Further research will include the evaluation of subject drivers in EB direction of PR-22 DTL using lane position, average speed and acceleration noise as performance measurements. In addition, a new configuration with additional modifications for signs and pavement markings will be considered for the analysis since the proposed treatments did not had a significant effect that reduced the variability in lane position or average speed variables.

Acknowledgments. The authors of this article extend their gratitude to the University Transportation Centers (UTC) Program for funding and supporting SAFER-SIM investigations over the past years. Also, the authors acknowledge Dr. Pedro Torres, for the technical assistance in the statistical analysis.

References

1. Obenger, J.: Managed lanes. Public Roads **68**(3), 48–55 (2004)
2. Colucci, B.: Dynamic toll lane: a success story as part of the public private partnerships in the commonwealth of Puerto Rico. In: The Proceedings of the 2nd International Conference on Public-Private Partnerships, Austin, Texas (2015)
3. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration (2009)
4. Fendrick, M.: Signing and pavement marking for concurrent-flow high-occupancy vehicle lanes: summary of current practice. Transp. Res. Rec. J. Transp. Res. Board **1603**, 64–71 (1997)
5. Standard Highway Signs: Federal Highway Administration. http://mutcd.fhwa.dot.gov/
6. Brown, L., McDonald Jr., D.R., Myers, E.J.: Developing traffic control strategies at toll plazas: institute of transportation engineers. ITE J. **76**(11), 22 (2006)
7. Carlson, P., Park, E., Andersen, C.: Benefits of pavement markings: a renewed perspective based on recent and ongoing research. Transp. Res. Rec. J. Transp. Res. Board **2107**, 59–68 (2009)
8. Fisher, D.L., Rizzo, M., Caird, J., Lee, J.D.: Handbook of Driving Simulation for Engineering, Medicine, and Psychology. Taylor and Francis Group, LCC, Florida (2011)
9. Dutta, A., Carpenter, R., Noyce, D., Duffy, S., Fisher, D.: Drivers' understanding of overhead freeway exit guide signs: evaluation of alternatives with an advanced fixed-base driving simulator. Transp. Res. Rec. J. Transp. Res. Board **1803**, 102–109 (2002)
10. Schattler, K., Pellerito Jr., J., McAvoy, D., Datta, T.: Assessing driver distraction from cell phone use: a simulator-based study. Transp. Res. Rec. J. Transp. Res. Board **1980**, 87–94 (2006)
11. Wang, P., Fang, S., Wang, J.: Impact of lane width of mountain highway on vehicle lateral position and speed. Adv. Transp. Stud. **32** (2014)
12. Valdés, D., Colucci, B., Fisher, D., Ruiz, J., Colón, E., García, R.: Driving simulation in the safety and operation performance of the freeway toll plaza. Transp. Res. Rec. J. Transp. Res. Board **2602**, 129–137 (2016)
13. Manual de Señales de Tránsito para las Vías Públicas de Puerto Rico, Departamento de Transportación y Obras Públicas (1979)
14. Boonsiripant, S.: Speed profile variation as a surrogate measure of road safety based on GPS-equipped vehicle data. Dissertation. Georgia Institute of Technology, Atlanta, Georgia (2009)
15. Montella, A., Aria, M., D'Ambrosio, A., Galante, F., Mauriello, F., Pernetti, M.: Perceptual measures to influence operating speeds and reduce crashes at rural intersections: driving simulator experiment. Transp. Res. Rec. J. Transp. Res. Board **2149**, 11–20 (2010)
16. Colonna, P., Intini, P., Berloco, N., Ranieri, V.: The influence of memory on driving behavior: how route familiarity is related to speed choice. Saf. Sci. **82**, 456–468 (2016)

Is the Driving Behaviour of Young Novices and Young Experienced Drivers Under Alcohol Linked to Their Perceived Effort and Alertness?

Catherine Berthelon[1(✉)] and Edith Galy[2]

[1] IFSTTAR, TS2, LMA, Salon de Provence 13300, France
catherine.berthelon@ifsttar.fr
[2] LAPCOS, University of Nice Sophia Antipolis, Nice, France

Abstract. The aim is to evaluate effort and alertness perception and objective driving performance of young drivers depending on Blood Alcohol Concentration (BAC) and driving experience. Young novice and young experienced drivers participated in three simulated driving sessions (BACs of 0.0, 0.2 and 0.5 g/L). They had to drive during 45 min. on a simulated highway road. After each driving session, they responded to the Thayer scale and to an adaptation of the NASA-TLX. Results showed that young experienced drivers estimated to make less effort and had better driving performance than young novice drivers. Estimated alertness level was the lowest and speed variation was the highest with BAC 0.5 g/l. It also existed an interaction effect between perceived effort and alcohol and between alertness and alcohol on driving performance. In summary, alcohol degrades driving performance, and especially when the effort is high, alertness is low and drivers lack experience.

Keywords: Young drivers · Alcohol · Driving simulator · Subjective estimation · Performance

1 Introduction

Young novice drivers have a high risk of crashes, notably during the first year of driving [1]. They are also over represented in crashes linked to long period of drive or night driving [2] and it is well known that the major part of sleep (or fatigue)-related crashes takes place during the two periods of physiological decrease of alertness (early morning hours 2:00–6:00 a.m. and afternoon period 1:00–4:00 p.m.) [3]. A monotonous road environment can also influence the level of alertness and the first effects of fatigue and drowsiness can rapidly occur [4].

Workload induced by driving can also vary as a function of environment and driving experience. Thus, when driving in a monotonous road environment, workload and/or effort can be high and associated to a decrease of driving performance, particularly for novice drivers [5]. This is explained by the fact that the drivers have to provide a high effort to overcome the vigilance decrement and effort is more important for novice because they have not completely automatized driving tasks compared to

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_84

experienced drivers [6]. Moreover, researches show that effort of novice drivers is too high when alertness is low and the lack of resources can explain this result [7].

Finally, alcohol is recognized as one major factor of driving impairment and a linear relationship has been demonstrated between blood alcohol concentration (BAC) and crash risk, notably for young drivers [8, 9].

Consequently, the aim of this work is to evaluate the combined effect of these major factors of accident and, more particularly, to examine alcohol effect on driving performance as a function of experience, perceived effort and alertness. The main hypothesis is that, with alcohol, when alertness is low and for long and monotonous driving task, driving impairment is higher for young novice drivers than for young experienced drivers. A secondary hypothesis is that, in this case, an increase of effort due to lack of experience decreases alertness.

2 Experimental Protocol

Fifteen young novice drivers (18 years, less than two months of driving license) and fifteen young experienced drivers (21 years, 3 years of driving license) participated in three simulated driving sessions in which BACs were randomly manipulated (0.0, 0.2 and 0.5 g/l). The order of the session was counterbalanced.

Every experimental session took place between 1:45 and 3:45 pm during post-prandial period, around 1 h after the drink. The task consisted to drive on a circuit representing a typical highway road during 45 min and to maintain a steady speed (110 km/h) and a stable position on the right lane.

After each driving session, participants estimated their workload (NASA-TLX questionnaire) [10] and their subjective alertness and tension (Thayer checklist) [11]. Speed and lane position of their vehicle were continuously recorded.

Only estimated alertness and effort were analyzed in this paper. Generalized linear model (GLM) were applied to data, completed with correlations.

This study was granted ethical approval by the French local ethics committee and by the French Health Products Safety Agency.

3 Results

First, we noted that obtained levels of alcohol before the drive were similar to what was expected (Table 1).

Table 1. Mean BACs measured by breathalyzer in each experimental session prior and following the driving task as a function of group (SD between brackets). Values obtained with breathalyzer must be multiplied by two in order to obtain a value in g/l

	0.2 g/l, prior the drive	0.2 g/l, after the drive	0.5 g/l, prior the drive	0.5 g/l, after the drive
Young novice	0.114 (0.011)	0.064 (0.008)	0.252 (0.011)	0.166 (0.130)
Young experienced	0.108 (0.015)	0.044 (0.015)	0.250 (0.020)	0.160 (0.010)

3.1 Effect of Alcohol

Results of generalized linear models showed that estimated alertness varied as a function of alcohol ($\chi^2(2) = 5.37$, $p < .068$), and was lower with 0.5 g/l of alcohol. Speed ($\chi^2(2) = 12.483$, $p < .002$) and Standard Deviation of speed (SDspeed) ($\chi^2(2) = 9.42$, $p < .007$) also significantly varied as a function of alcohol and were higher with 0.5 g/l than with the two others conditions (0.2 and 0.0 g/l) (Table 2).

Table 2. Global data as a function of BAC (SD between brackets).

	0.0 g/l	0.2 g/l	0.5 g/l
SDLP (cm)	42.14 (12.66)	48.74 (14.63)	55.27 (20.76)
LP (cm)	752.70 (38.46)	750.91 (52.28)	751.81 (43.12)
Speed (m/s)	30.17 (0.81)	30.18 (0.74)	**30.40 (1.06)**
SDspeed (m/s)	0.76 (0.56)	0.93 (0.55)	**1.10 (0.72)**
Estimated effort (Nasa)	11.31 (4.76)	12.27 (3.66)	13.28 (3.81)
Estimated alertness (Thayer)	1.07 (0.69)	0.88 (0.54)	0.86 (0.69)

It also existed an interaction between alertness and alcohol on driving performance ($\chi^2(2) = 13.65$, $p < .001$), thus an increase of subjective alertness produced an increase of speed but only with BAC 0.2 g/l and BAC 0.0 g/l.

Other interactions between alcohol and effort on speed ($\chi^2(2) = 45.20$, $p < .001$) and on SDspeed ($\chi^2(2) = 19.01$, $p < .001$) showed that, without alcohol and with BAC 0.2 g/l, an increase of effort produced speed decrease, conversely with BAC 0.5 g/l an increase of effort was linked to an increase of speed. Concerning SDspeed, an increase of effort produced increase of SDspeed with BAC 0.0 and BAC 0.5 g/l, but decrease of SDspeed with BAC 0.2 g/l.

3.2 Effect of Experience

Results of generalized linear model analyses showed that young novice drivers estimated to make a more important effort than young experienced drivers ($\chi^2(1) = 14.56$, $p < .001$). Their Standard Deviation of Lateral position (SDLP) ($\chi^2(1) = 13.92$, $p < .001$) and SDspeed ($\chi^2(1) = 7.24$, $p < .007$) are also significantly higher than these of young experienced drivers. Conversely, their speed was lower and their position lightly more on the right on the lane compared to more experienced drivers ($\chi^2(1) = 10.52$, $p < .001$) (Table 3).

Correlation analyses indicated that globally alertness estimations were negatively correlated with perceived effort ($r = -.305$; $p < .01$), stability of speed ($r = -.110$; $p < .01$) and of lateral position ($r = -.146$; $p < .01$) et positively correlated with speed ($r = .139$; $p < .01$). This was notable for young novice drivers, indicating that when young novice drivers estimated high alertness, perceived effort was low, and stability of speed and of lateral position was better, however they also increase their speed. In another hand, effort was positively correlated with SDLP ($r = .161$; $p < .01$) and with SDpeed ($r = .131$; $p < .01$).

Table 3. Global data as a function of experience (SD between brackets).

	Young novice	Young experienced
SDLP (cm)	49.72 (20.31)	47.65 (13.34)
LP (cm)	**752.7 (58.6)**	753.4 (24.4)
Speed (m/s)	**30.22 (1.08)**	30.28 (0.68)
SDspeed (m/s)	0.99 (0.77)	0.87 (0.44)
Estimated effort (Nasa)	12.42 (4.33)	12.16 (2.89)
Estimated alertness (Thayer)	0.83 (0.55)	1.04 (0.73)

Complementary analyses also showed that it is notably for young experienced drivers that the increase of speed is verified with 0.5 g/l. (Figure 1), conversely it is notably for young novice drivers that the increase of Sdspeed is verified with 0.5 g/l (Fig. 2).

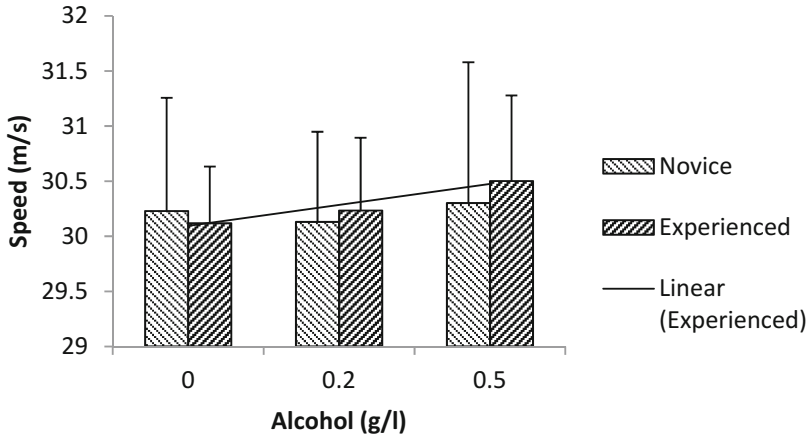


Fig. 1. Mean speed as a function of alcohol and driving experience.

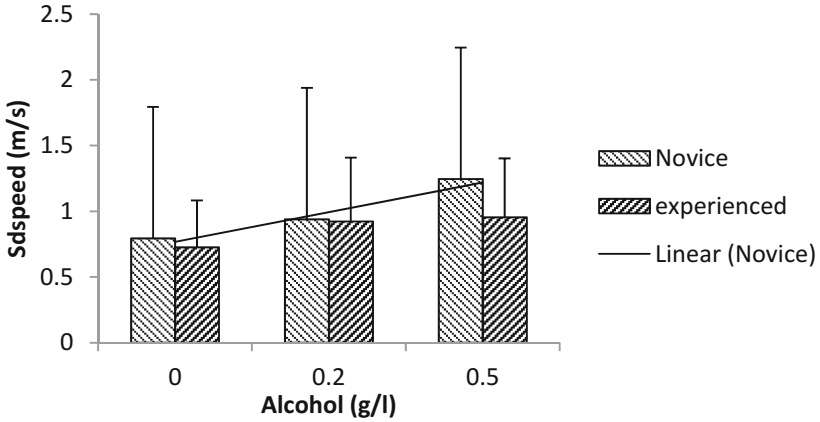


Fig. 2. Mean standard deviation of speed as a function of alcohol and driving experience.

4 Conclusion

Results indicate some balanced effect between estimated effort and estimated alertness. This balanced effect permits to conserve good performance whatever the level of driving experience. Effort production thus compensates lack of alertness until a certain threshold.

In another hand, an increase of subjective alertness produced a speed increase with 0.2 and 0.0 g/l, e.g. when participants estimated to be alert, they increased their speed in the driving sessions without alcohol and with a very low dose of alcohol. Conversely, an increase of effort produced speed decrease in the driving sessions without alcohol and with BAC 0.2 g/l and a speed increase with BAC 0.5 g/l. This implies that effort and alertness have also separated effects on driving parameters.

Globally, an increase of SDLP is correlated to an increase of speed and of SDspeed indicating close links between lateral and longitudinal parameters of the drivers' trajectories.

Results are also coherent with past results on alcohol, with speed and speed variation higher with BAC 0.5 g/l than with the other two conditions. Alcohol could thus have a de-inhibited effect on drivers' speed, but imply a low stability of speed. Moreover, with 0.5 g/l of alcohol, speed variation increased with the effort. Conversely, with BAC 0.2 and 0.0 g/l, an effort increase was linked to speed variation decrease. This signifies that, with 0.5 g/l of alcohol, an increase of effort is not sufficient to maintain good longitudinal performance. All the results obtain with BAC 0.5 g/l thus converge and indicate an increase of risky driving whatever the level of driving experience.

Results concerning driving performance are also coherent with literature concerning driving experience. Thus, driving performance (stability of speed and of lateral position) is determined by the level of experience, with performance better for young experienced drivers than for young novice drivers [12]. Young experienced drivers also estimated lower effort and higher alert compared to young novice drivers. Thus, the driving task seems them easier due to their supplementary experience.

Another interesting result consists in the fact that alcohol has differential effect as a function of driving experience. Only, young experienced drivers increase their speed with the level of alcohol, which is a sign of the de-inhibited effect of alcohol. But, without having increase their speed, young novice drivers have more difficulties to stabilize their speed with 0.5 g/l than young experienced drivers.

First effect of alcohol could thus appear at lower doses for young novice drivers than for young experienced drivers and this result confirms our earlier works. Young novice drivers have not acquire all the competences necessary for a safe driving and to use regulatory mechanisms to preserve their driving performance [13]. In the theoretical field of mental workload, it is also known that the overload threshold is reached earlier for novice drivers than for more experienced drivers [14]. It thus seems that alcohol degrades driving performance, and especially when the effort is high, alertness is low and drivers lack experience e.g. in long monotonous drive at a period of low alertness.

These results highlight the interest to study combined effect of different factors influencing driving performance.

Acknowledgments. This work was funded by the French Road Safety Foundation, Convention 2013/MP/03. The authors thank Laurent Ferrier for making passed this experiment and participated in the preliminary treatments of the data.

References

1. Maycock, G., Forsyth, E.: Cohort study of learner and novice drivers, part 4: Novice driver accidents in relation to methods of learning to drive, performance in the driving test and self-assessed driving ability and behaviour. Report No. TRL 275. Crowthorne: TRL (1997)
2. Williams, A.F.: Teenage drivers: patterns of risk. J. Saf. Res. **34**, 5–15 (2003)
3. Pack, A.I., Pack, A.M., Rodgman, E.A., Cucchiara, A., Dinges, D.F., Schwab, C.W.: Characteristics of crashes attributed to the driver having fallen asleep. Accid. Anal. Prev. **27**, 769–775 (1995)
4. Dunn, N., Williamson, A.: Driving monotonous routes in a train simulator: the effect of task demand on driving performance and subjective experience. Egonomics **55**, 997–1008 (2012)
5. Smith, S.S., Horswill, M.S., Chambers, B., Wetton, M.: Hazard perception in novice and experienced drivers: the effects of sleepiness. Accid. Anal. Prev. **41**, 729–733 (2009)
6. De Craen, S., Twisk, D.A.M., Hagenzieker, M.P., Elffers, H., Brookhuis, K.A.: The development of a method to measure speed adaptation to traffic complexity: identifying novice, unsafe, and overconfident drivers. Accid. Anal. Prev. **40**, 1524–1530 (2008)
7. Paxion, J., Galy, E., Berthelon, C.: Mental workload and driving. Front. Psychol. (2014). doi:10.3389/fpsyg.2014.01344
8. Peck, R.C., Gebers, M.A., Voas, R.B., Romano, E.: The relationship between blood alcohol concentration (BAC), age, and crash risk. J. Saf. Res. **39**, 311–319 (2008)
9. Zador, P.L., Krawchuk, S.A., Voas, R.B.: Alcohol related relative risk of driver fatalities and driving involvement in fatal crashes in relation to driver age and gender: an update using the 1996 data. J. Study Alcohol **61**(3), 387–395 (2000)
10. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (task load index): results of emperical and theoritical research. Adv. Psychol. **52**, 139–183 (1988)
11. Thayer, R.E.: Activation-deactivation adjective check list: current overview and structural analysis. Psychol. Rep. **58**(2), 607–614 (1986)
12. Freydier, C., Berthelon, C., Bastien-Toniazzo, M.: Does early training improve driving skills of young novice French drivers? Accid. Anal. Prev. **96**, 228–236 (2016)
13. McCartt, A.T., Mayhew, D.R., Braitman, K.A., Ferguson, S.A., Simpson, H.M.: Effects of age and experience on young driver crashes: review of recent literature. Traffic Injury Prev. **10**, 209–219 (2009)
14. Paxion, J., Berthelon, C., Galy, E.: Does driving experience delay overload threshold as a function of situation complexity? In: Dorn, L., Sullman, M. (eds.) Driver Behaviour and Training, vol. VI, pp. 13–23. Ashgate Publishing, England (2013)

Injury Severity Analysis in Vehicle-Pedestrian Crashes

Younshik Chung[1(✉)], Tai-Jin Song[2], and Juyoung Kim[2]

[1] Yeungnam University, Gyeongsan 38541, South Korea
tpgist@yu.ac.kr
[2] Korea Transport Institute, Sejong 30147, South Korea
{tjsongl7, jykim}@koti.re.kr

Abstract. Since traffic accident data has been typically collected by the police who record accident information based on estimation through the engineering methods and statements of accident-involved people and witnesses in the vicinity of an accident, accident data recorded by the police may inevitably include errors. However, the technological progression of in-vehicle driving recording devices such as event data recorder, vehicle black box (VBB), and various sensors using in naturalistic driving study tackles the change of methodological paradigms in traffic safety research. The objective of this study is to analyze injury severity in taxi-pedestrian crashes using more accurate crash data by VBBs such as time-to-collision, crash speed, crash angle, and crash region of vehicle and pedestrian. To carry out this study, VBB data for two years (2010–2011) collected by taxies operating in Incheon, Korea are used. Then, the ordered probit model is applied to analyze injury severity in taxi-pedestrian crashes.

Keywords: Vehicle black box · Vehicle-pedestrian crash · Injury severity · Crash data accuracy · Pedestrian behavior · Driver maneuver

1 Introduction

There have been various efforts to identity the factors that affect pedestrian injury severity at vehicle-pedestrian crashes using collected accident dataset. In previous, traffic accident data has been typically collected by the police who record accident information based on estimation through the engineering methods of post-crash trajectories and skid marks at the accident spot and statements of accident-involved people and witnesses in the vicinity of an accident [1–3]. Thus, accident data recorded by the police may inevitably include errors. However, the technological progression of in-vehicle driving recording devices such as event data recorder (EDR), vehicle black box (VBB), and various sensors using in naturalistic driving study (NDS) tackles the change of methodological paradigms in traffic safety research.

Specifically, the VBB, one of the advanced in-vehicle driving recording devices, is equipped with a high resolution camera and wide lens that can capture clear, color images in the front of the vehicle and save them on memory, and they have a GPS device that provides location and time information anywhere on the Earth [3]. Based on

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_85

these technologies and functions, the VBB was initially introduced to be used as legal evidence regarding a traffic accident [4]. However, the VBB enables the collection of internal driving information such as GPS (global positioning system)-based location, braking start time, crash time, and speed as well as the collection of video and audio data with respect to external driving situations such as crash angle and crash region of vehicle or pedestrian. As a result, these state-of-the-art technologies may reduce the human errors during the accident data collection, they may contribute to increasing the validity of the models to assess the impact on traffic safety.

The objective of this study is to analyze injury severity in taxi-pedestrian crashes using more accurate crash data by VBBs such as time-to-collision, crash speed, crash angle, and crash region of vehicle and pedestrian. To carry out this study, VBB data for two years collected by taxies operating in Incheon, Korea are used. Then, the ordered probit model is applied to analyze injury severity in taxi-pedestrian crashes, and based on the fitted model, the discussion and conclusion are presented.

2 Prior Studies on Vehicle-Pedestrian Crash Injury Severity

As indicated by Mannering and Bhat [5], road safety research to understand factors that affect the injury severity of vehicle crashes has evolved rapidly in terms of methodological approaches. Since the injury severity data are generally represented by ordinal categories such as fatal injury, incapacitating injury, non-capacitating, possible injury, and property damage only, a family of discrete outcome models have become the prevailing analytical technique [6]. The efforts of traffic safety researchers to determine more precisely the influence of injury severity factors have enabled the development of sophisticated models such as simple binary outcome models, ordered discrete outcome models, and unordered multinomial discrete outcome models. However, the crash data to be applied these sophisticated models seem not to have a great evolution in terms of its quality and contents.

Most of prior studies on the injury severity of vehicle-pedestrian crashes have used police-recorded crash data. Although such data may include inaccurate information due to misreporting and misrecording [7, 8], there were rarely alternatives to these data sources. The injury severity studies of vehicle-pedestrian crashes in the 1990s were based on the police-recorded crashes. For example, Zeeger et al. [9] applied simple comparison methods to identify fatality likelihood difference between older pedestrian group and younger one. In this study, they used more than 26,000 police-recorded pedestrian crashes occurred in North Carolina over 11 years. A similar study was conducted by Fontaine and Gourlet [10] using police-recorded crashes in France between March 1990 and February 1991 In this study, they examined the characteristics of fatal pedestrian crash through correspondence analysis. A Danish researcher analyzed the trend of pedestrian-vehicle crash in 47 Danish cities with simple comparison method based on police-recorded crashes [11].

Since the 2000s, a wide variety of methodological techniques were applied to identify the level of pedestrian injury severity: heteroskedastic logit model [12], multinominal logistic regression [13], mixed logit model [14], binary and ordinal logistic regression [15], random parameter logit model [16], latent class with ordered

probit model and K-means with a multinomial logit model [17], logistic regression [18], ordered logit model, generalized ordered logit model, and latent segmentation based ordered logit model w [19], multivariate logistic regression [20], mixed logit model [21], statistical tests between different groups [22], ordered response model [23], linearized spatial logit model [24], and mixed logit model [25]. All of these prior studies were based on police-recorded crash data, which may include misreporting and misrecording errors.

Unlike various efforts on the injury severity analysis method, those on the improvement of crash data accuracy itself have received little attention. However, the technological progression of in-vehicle driving recording devices such as event data recorder (EDR), vehicle black box (VBB), and various sensors using in naturalistic driving study (NDS) enables the safety researchers to obtain new data contents as well as to use more accurate crash data. Specifically, there was a new technology to record various crash information by using a high resolution camera and wide lens and a global positioning system (GPS) device, namely vehicle black box (VBB). Based on VBB crash data, Song et al. [26] investigated the main factors affecting the vehicle-pedestrian injury severity through structural equation modeling (SEM). Different from the prior studies, they used more accurate speed and pedestrians' behavior at the crash moment captured from VBB.

3 Data Description

This study combines two heterogeneous dataset: police-recorded crash data and VBB-recorded crash data. First, the police collected 565 taxi-pedestrian crashes from 2010 to 2011 in the city of Incheon, Korea. This crash database includes various crash-related characteristics: pedestrian and driver characteristics, violation types of pedestrian and driver, geometric characteristics of crash location, crash time, and crash characteristics such as crash type and crash speed. Of these data, pedestrians' behavior at the crash moment and crash characteristics is basically recorded through drivers' or pedestrians' statement or other eyewitness's account. Thus, these can inevitably include human errors [8].

Table 1. Crash information extracted from VBB

Accident time	Time-to-collision	Speed (km/h)		Location (degree)		Crash characteristics						
		Running	Crash	Longitude	Latitude	Vehicle part		Pedestrian part			Other parts	
						Driver behavior	…	Pedestrian behavior	1st impact region	…	2nd impact object	3rd impact object
2010-01-27 18:02:05	0.46	28.5	26.3	126.789	37.532	Signal violation	…	Distracted	Leg	…	None	None
2010-03-19 20:48:34	0.91	38.5	25.2	126.628	37.480	Normal	…	Violated	Arm	…	Vehicle	Road
2010-05-08 05:04:47	0.00	61.1	61.3	126.729	37.489	Speeding	…	Normal	Leg	…	Road	Another moving vehicle
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

To use more accurate crash data and new crash variables, this study applies VBB-recorded crash data collected from same spatiotemporal regions. In total, 439 taxi-pedestrian crashes recorded by taxi VBBs were collected from the Incheon branch office of Korea Taxi Mutual Aid Association (KoTMA). Table 1 shows the some information collected from VBBs. In this table, variables regarding driver and pedestrian behavior at the crash moment were manually extracted from the video image recorded in each VBB. Also, the variable with respect to time-to-collision implies time difference between speed drop beginning time of taxi and crash time. That is, it implies the time from the reaction (i.e., braking) to the crash moment. These information can be reconstructed by VBB view program as seen in Fig. 1.

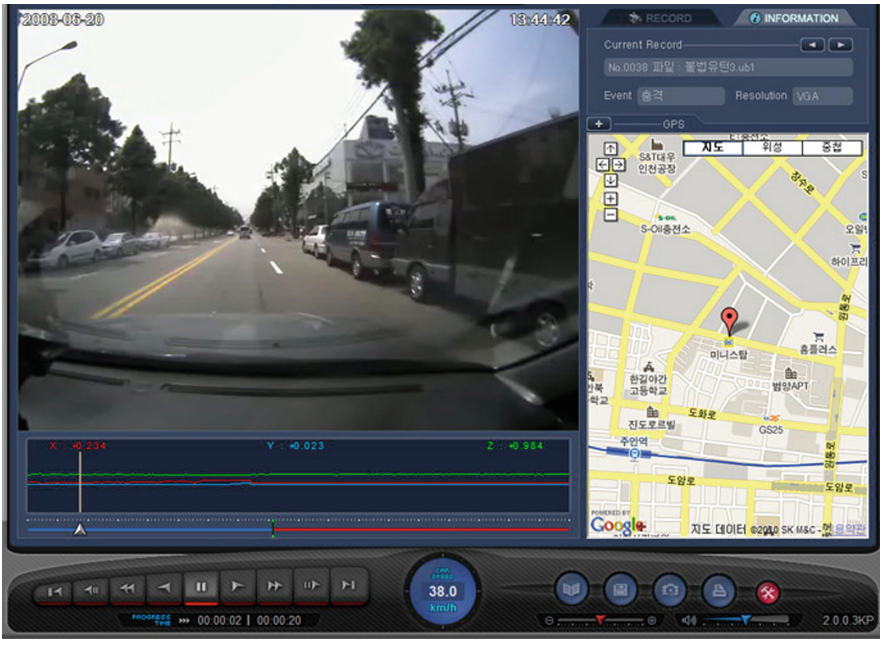


Fig. 1. VBB view program for reconstructing crash data

Some of crash records by the police would not be included in the VBB crash database, specifically in the case of minor crashes. As a result, two heterogeneous crash data sources are combined to use more plentiful crash information in the analysis of vehicle-pedestrian crash injury severity. Then, 357 detailed crash data were successfully merged. Therefore, the merged crash data were classified into five groups: vehicle characteristics, pedestrian characteristics, environmental characteristics, road characteristics, and crash characteristics. The definition of each independent variable are shown in Table 2.

Table 2. Definition of the explanatory variables

Category	Variable definition
Vehicle characteristics	**Driver behavior** Norm (1 if driver was normally driving, 0 otherwise) Distracted (1 if driver was driving with distraction, 0 otherwise) Speeding (1 if driver was speeding, 0 otherwise) Violate (1 if driver violated traffic rules, 0 otherwise)
	Vehicle maneuver U-turn (1 if vehicle was U-turning, 0 otherwise) Straight through (1 if vehicle was driving straight through, 0 otherwise) Right-turn (1 if vehicle was right-turning, 0 otherwise) Left-turn (1 if vehicle was left-turning, 0 otherwise)
Pedestrian characteristics	**Pedestrian age** Child (1 if pedestrian age was under 13, 0 otherwise) Junior_high (1 if pedestrian age was between 13 and 15, 0 otherwise) High (1 if pedestrian age was between 16 and 18, 0 otherwise) Adult (1 if pedestrian age was between 19 and 64, 0 otherwise) Elder (1 if pedestrian age was over 64, 0 otherwise)
	Pedestrian crossing characteristic Norm_cross (1 if normally crossing a crosswalk, 0 otherwise) Fault_cross (1 if faulty crossing a non-crosswalk, 0 otherwise) Violate_cross (1 if violated a crosswalk signal, 0 otherwise)
Environmental characteristics	**Time period** Daytime (1 if crash occurred during 07:00–18:00, 0 otherwise)
	Road surface condition Wet (1 if road surface was wet, 0 otherwise)
	Sight obstruction Street_park (1 if vehicle parked at street, 0 otherwise)
Road characteristics	**Segregation of vehicle and pedestrian paths** Segregate (1 if segregated, 0 otherwise)
	Number of lanes Lane
	Median type M_none (1 if median was none, 0 otherwise) M_marking (1 if median was painted-making, 0 otherwise) M_post (1 if median was flexible post, 0 otherwise) M_planting (1 if median was planting, 0 otherwise) M_barrier (1 if median was barrier, 0 otherwise)
	Land use type in the proximity of road Land_reside (1 if residential areas, 0 otherwise) Land_business (1 if business areas, 0 otherwise) Land_other (if others such as parks and school zones, 0 otherwise)
	Traffic control type Sig_normal (1 if traffic signal was normal type, 0 otherwise) Sig_flashing (1 if traffic signal was flashing type, 0 otherwise) Sig_none (1 if traffic signal was none, 0 otherwise)

(*continued*)

Table 2. (*continued*)

Category	Variable definition
Crash characteristics	**Pedestrian behavior** Ped_prep (1 if preparing the crash, 0 otherwise) Ped_walk (1 if only walking right before crashing, 0 otherwise) Ped_dist (1 if walking with distraction right before crashing, 0 otherwise)
	Pedestrian crash side Front (1 if front side, 0 otherwise) Side (1 if lateral side (i.e., left or right), 0 otherwise) Back (1 if back side, 0 otherwise)
	1st impact region of pedestrian Leg_1 (1 if leg, 0 otherwise) Other_1 (1 if others, 0 otherwise) Arm_1 (1 if arm, 0 otherwise) Head_1 (1 if head, 0 otherwise)
	2nd impact region of pedestrian None_2 (1 if none, 0 otherwise) Leg_2 (1 if leg, 0 otherwise) Other_2 (1 if others, 0 otherwise) Arm_2 (1 if arm, 0 otherwise) Head_2 (1 if head, 0 otherwise)
	3rd impact region of pedestrian None_3 (1 if none, 0 otherwise) Leg_3 (1 if leg, 0 otherwise) Other_3 (1 if others, 0 otherwise) Arm_3 (1 if arm, 0 otherwise) Head_3 (1 if head, 0 otherwise)
	2nd impact object type of pedestrian Agst_none_2 (1 if none, 0 otherwise) Agst_vehicle_2 (1 if crashed vehicle, 0 otherwise) Agst_road_2 (1 if road, 0 otherwise)
	3rd impact object type of pedestrian Agst_ None_3 (1 if none, 0 otherwise) Agst_vehicle_3 (1 if another moving vehicle, 0 otherwise) Agst_road_3 (1 if road, 0 otherwise) Agst_other_3 (1 if other road facilities, 0 otherwise)
	Vehicle crash side Left_corner (1 if left corner of the front, 0 otherwise) Front (1 if front, 0 otherwise) Right_corner (1 if right corner of the front, 0 otherwise)
	Vehicle speed Running speed (running speed prior to starting break) Crash speed (Speed at the crash moment reconstructed by VBB)
	Time to crash Time difference between braking start and crash moment

In this study, crash severity were classified into six group based on hospitalized periods: (1) none evident injury or first-aid treatment at crash scene, (2) minor injury hospitalized under two weeks, which includes bruises, lumps on the head, and minor lacerations, (3) moderate injury hospitalized between two and five weeks, (4) serious injury hospitalized between five and nine weeks, (5) severe injury hospitalized over nine weeks, and (6) fatality. The severity level is different from previous studies. Since the hospitalized period is one of the most critical factors for taxi insurance fees, it is concretely managed by the KoTMA based on VBB data.

4 Injury Severity Analysis

Table 3 summarizes the estimated model for the injury severity of taxi to pedestrian crashes. From the fitted ordered probit model, it can be concluded that nine variables affecting injury severity are statistically significant. Specifically, they include crash speed, crashes that the 1st impact region of pedestrians was their leg, crashes that the 2nd

Table 3. Result of ordered probit model on injury severity

Variables	Coefficient	p-value
Vehicle speed		
Crash speed	0.017	0.000
1st impact region of pedestrian		
Leg	−0.378	0.010
2nd impact object type		
Vehicle	0.350	0.004
3rd impact object type		
Another moving vehicle	1.428	0.001
3rd impact region of pedestrian		
Head	0.806	0.021
Driver behavior		
Speeding	0.691	0.026
Driver maneuver		
Straight through	−0.285	0.054
Pedestrian age		
Junior high school (13–15)	−0.567	0.049
Median type		
None	0.419	0.002
μ_0	−1.123	
μ_1	0.330	
μ_2	1.172	
μ_3	1.938	
μ_4	2.695	
Number of crashes	357	
McFadden's ρ^2	0.085	
LR chi2	89.290	

impact object of pedestrians was the crashed vehicle, crashes that the 3rd impact object of pedestrians was another moving vehicle, crashes that the 3rd impact regions of pedestrians was their head, crashes while the taxi was under speeding, crashes while the taxi driver was going straight through, crashes with junior high school students, and crashes that occurred at none median sections. Of these, most variables contribute to increasing injury severity, except the variables with respect to 'leg', 'straight through', and 'junior high school'. Specifically, crashes with junior high school students resulted in less severity. It can be interpreted by the perspectives of the physical flexibility and the level of crash-related perception. That is, the junior high school students would have all of these capabilities, comparing with other age groups. As for crashes while straight through, the behaviors of drivers during this maneuver could be simple, comparing with other maneuvers such as various turning movements. Lastly, the results of other variables are consistent with theoretical expectations or results from previous studies.

5 Conclusion

This study investigated injury severity in taxi-pedestrian crashes using more accurate crash data such as time-to-collision, crash speed, crash angle, and crash region of vehicle and pedestrian, which were re-constructed from a new crash recording device, namely vehicle black box (VBB). Based on ordered probit model, this study concluded that six variables affect more injury severity: crash speed, crashes that the 2nd impact object of pedestrians was the crashed vehicle, crashes that the 3rd impact object of pedestrians was another moving vehicle, crashes that the 3rd impact regions of pedestrians was their head, crashes while the taxi was under speeding, and crashes that occurred at none median sections. On the other hand, the variables with respect to 'leg', 'straight through', and 'junior high school' contribute to less severity. The findings of this study could be a basis of development of various safety-related vehicle devices as well as of development of various strategies to mitigate injury severity of vehicle-pedestrian crashes.

References

1. Chung, Y., et al.: Injury severity in delivery-motorcycle to vehicle crashes in the Seoul metropolitan area. Accid. Anal. Prev. **62**, 79–86 (2014)
2. Loo, B.P.Y.: Validating crash locations for quantitative spatial analysis: a GIS-based approach. Accid. Anal. Prev. **38**, 879–886 (2006)
3. Moes, D.J.A.R.: Optimizing immunosuppression with mTOR inhibitors in renal transplant recipients. Ph.D., Faculty of Science, Leiden University (2014)
4. Koo, J., et al.: Multithreading embedded multimedia application for vehicle blackbox. In: Gaol, F.L. (ed.) Recent Progress in Data Engineering and Internet Technology, vol. 156, pp. 379–384. Springer, Berlin (2013)
5. Michalaki, P., et al.: A time-series analysis of motorway collisions in England considering road infrastructure, socio-demographics, traffic and weather characteristics. J. Transp. Health **3**, 9–20 (2016)

6. Savolainen, P.T., et al.: The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accid. Anal. Prev. **43**, 1666–1676 (2011)
7. Imprialou, M., Quddus, M.: Crash data quality for road safety research: current state and future directions. Accid. Anal. Prev. (In Press)
8. Chung, Y., Chang, I.: How accurate is accident data in road safety research? An application of vehicle black box data regarding pedestrian-to-taxi accidents in Korea. Accid. Anal. Prev. **84**, 1–8 (2015)
9. Zegeer, C.V., et al.: Analysis of elderly pedestrian accidents and recommended countermeasures. Transp. Res. Rec. 1405, 56–63 (1993)
10. Fontaine, H., Gourlet, Y.: Fatal pedestrian accidents in France: a typological analysis. Accid. Anal. Prev. **29**, 303–312 (1997)
11. Jensen, S.: Pedestrian safety in Denmark. Transp. Res. Rec. J. Transp. Res. Board **1674**, 61–69 (1999)
12. Kim, J.-K., et al.: Age and pedestrian injury severity in motor-vehicle crashes: a heteroskedastic logit analysis. Accid. Anal. Prev. **40**, 1695–1702 (2008)
13. Damsere-Derry, J., et al.: Pedestrians' injury patterns in Ghana. Accid. Anal. Prev. **42**, 1080–1088 (2010)
14. Kim, J.-K., et al.: A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. Accid. Anal. Prev. **42**, 1751–1758 (2010)
15. Moudon, A.V., et al.: The risk of pedestrian injury and fatality in collisions with motor vehicles, a social ecological study of state routes and city streets in King County, Washington. Accid. Anal. Prev. **43**, 11–24 (2011)
16. Aziz, H.M.A., et al.: Exploring the determinants of pedestrian–vehicle crash severity in New York City. Accid. Anal. Prev. **50**, 1298–1309 (2013)
17. Mohamed, M.G., et al.: A clustering regression approach: a comprehensive injury severity analysis of pedestrian–vehicle crashes in New York, US and Montreal, Canada. Saf. Sci. **54**, 27–37 (2013)
18. Zhang, G., et al.: Analyzing fault and severity in pedestrian–motor vehicle accidents in China. Accid. Anal. Prev. **73**, 141–150 (2014)
19. Yasmin, S., et al.: Alternative ordered response frameworks for examining pedestrian injury severity in New York City. J. Transp. Saf. Secur. **6**, 275–300 (2014). (2014/10/02)
20. Koopmans, J.M., et al.: Urban crash-related child pedestrian injury incidence and characteristics associated with injury severity. Accid. Anal. Prev. **77**, 127–136 (2015)
21. Haleem, K., et al.: Analyzing pedestrian crash injury severity at signalized and non-signalized locations. Accid. Anal. Prev. **81**, 14–23 (2015)
22. Oikawa, S., et al.: Relation between vehicle travel velocity and pedestrian injury risk in different age groups for the design of a pedestrian detection system. Saf. Sci. **82**, 361–367 (2016)
23. Pour-Rouholamin, M., Zhou, H.: Investigating the risk factors associated with pedestrian injury severity in Illinois. J. Saf. Res. **57**, 9–17 (2016)
24. Prato, C.G., et al.: Considering built environment and spatial correlation in modelling injury severity for pedestrian crashes. In: 96th Annual Meeting of the Transportation Research Board, Washington D.C. (2017)
25. Tulu, G.S., et al.: Injury severity of pedestrians involved in road traffic crashes in Addis Ababa, Ethiopia. J. Transp. Saf. Secur. **9**, 47–66 (2017). (2017/03/30)
26. Song, T.-J., et al.: Exploring the vehicle-pedestrian crash severity factors based on the in-car black box recording data. In: 96th Annual Meeting of the Transportation Research Board, Washington D.C. (2017)

Road and Rail: Driving Simulation and Test Track

Motion Sickness Measurements for Young Male Adults in Vitality, Endurance, Profiles and Sensitivity

Aie Liu, Li Ding, and Yan Li(✉)

International Joint Research Center of Aerospace Biotechnology
and Medical Engineering of Ministry of Science and Technology of China,
School of Biological Science and Medical Engineering, Beihang University,
No. 37 Xueyuan Road, Haidian District,
Beijing 100191, People's Republic of China
yanli_buaa@buaa.edu.cn

Abstract. In the research field of motion sickness, one of the challenges is to measure motion sickness. In this study a revised motion sickness questionnaire (rMSQ) was proposed. 30 male subjects participated in the experiment, and they were seated in a rotating chair until the evoked motion sickness became unbearable or scheduled time was up. The rotation were repeated for two sessions, one for eyes opened and one for eyes closed. The baseline diastolic blood pressure had a significant correlation with the fear survey schedule value while baseline systolic blood pressure had a correlation with the rMSQ headache item. According to the eyes opened endurance time length, subjects were divided into three groups: the sensitive, the normal and the resistant. Statistical analysis indicated that the sensitive and resistant groups had marked differences in simplified coping style and nausea item in rMSQ. To conclude, rMSQ could describe motion sickness symptoms clearly.

Keywords: Motion sicking · Questionnaire · Rotating chair

1 Introduction

Motion sickness is defined as a series of symptoms appeared in traveling, which include pallor, yawning, headache, tiredness and nausea, etc. It could happen to people when they ride in almost all kinds of vehicles, such as bus, car, boat, plane or even in space shuttle. The most popular explanation of motion sickness generation is the conflicts among the integrating sensory inputs [1]. However, there also exist lots of influencing factors, like gender, age, sleepiness, and personal sensitivities [2, 3].

In most of the experimental studies, motion sickness is induced through three methods: the rotating drum, rotating chair, and virtual reality (mainly visual stimulus). Recently there are also a few studies that apply aural or tactile stimulus to induce the feeling of motion sickness [4, 5]. The rotating chair is a traditional and effective way to make subjects feel dizzy, nausea, or unable to concentrate. Although almost all studies ask the subjects to do pitch movements of the head during the rotation in order to produce the Coriolis effect [6], in this study, in order to evoke motion sickness more

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_86

efficiently, a rotating chair is used and the speed is set to high. Subjects are asked to sit still and do nothing during rotation, and the different endurance time lengths prove that the personal sensitivities indeed play a significant role in the rotating chair experiment. According to the eyes opened endurance time length, as usually the subjects were divided into three groups: the sensitive, the normal and the resistant.

According to the published papers, personal sensitivities weigh much in motion sickness incidence rates, which means for the same transportation conditions, people feel differently and present various symptoms. A 500 adult survey conducted in Hong Kong analyzes individual self-evaluation of the susceptibility of motion sickness. The rates of sensitive, mild to medium, and resistant types are 7.2%, 79.4%, and 14.3%, respectively [7]. Thus, to improve the endurance strength of sensitive and normal people, is of realistic importance.

Motion sickness can usually be measured either in an objective or subjective way, with the latter consisting of various kinds of questionnaires. For example, the Motion Sickness Susceptibility Questionnaire (MSSQ) investigates uncomfortable experiences of transportation in childhood and adulthood, and then predicts the motion sickness susceptibility [8]; there is also the Simulator Sickness Questionnaire (SSQ) aiming at measuring motion sickness evoked by virtual reality [9]; the Nausea Profile emphasizes measuring nausea levels [10]; and the Fast Motion Sickness Scale (FMS) [11]. The Motion Sickness Assessment Questionnaire (MSAQ) is another good example of such questionnaires in which symptoms are grouped into four subscales [12]. However, besides the abovementioned questionnaires, motion sickness can also be measured by physiological and motion parameters. Some studies use heart rate, blood pressure or photoplethysmography as indexes of motion sickness [13]; some studies measure the sway track of hip or head to indicate to what extent the subjects lose their balance [14, 15]; some analyze electrogastrogram to express the nausea levels; while some other studies observe the static stance time, the dynamic walking gait and the walking tasks in which subjects may stray from the defined route, before and after the experiment, to test the aftereffects of virtual realities [16]. However, Cobb claimed that there was no connection between subjective and objective evaluations [16]. An exploratory factor analysis has been carried out, indicating that there is much room for improving rMSQ, but more future work and careful study is needed.

To summarize, although studies are numerous, many researchers concentrate only on the countermeasures of motion sickness, and studies on personal sensitivity are still lacking. Therefore, in this study, a revised Motion Sickness Questionnaire (rMSQ) is proposed on the basis of MSAQ in the hope that more symptoms could be described concerning vestibular and visual stimulus; some emotional stress indexes are also included to measure the personal sensitivities in a comprehensive way, as some studies claim that those indexes contribute to motion sickness susceptibility indirectly [1, 3]; and the visual analog scale is changed from 0 to 10 to give a better expression of the subjects. A rotating chair experiment is also conducted with the aim of testing the validity of the proposed rMSQ and emphasizing on the differences between the motion sickness sensitive subjects and the resistant ones.

2 Methods

2.1 Subjects

30 healthy, male students participated in the experiment, aging from 18 to 28, with a mean of age 22.8 years. None of them had neurological or vestibule-related diseases and every one of them got paid. They were forbidden to drink or eat from two hours before the experiment. The study protocol was approved in advance by Capital Medical University. Each subject provided written informed consent before participating in the experiment. In this study, only male subjects were recruited considering the fact that motion sickness susceptibility could be affected by the female period.

Prior to the experiment, all the subjects finished the Edinburgh Handedness Inventory to ensure that they were right handed. Additionally, subjects were tested with respect to personality (by means of state-trait anxiety inventory (STAI) [17], simplified coping style questionnaire (SCSQ) [18], general self-efficacy scale (GSES) [19], and fear survey schedule (FSS) [8].

The motion sickness questionnaire adopted in this study was a revised version of MSAQ, consisting of four subscales, which included Behavioral Changes (e.g., eye strain, increased salvation, sweating, and shaking), Central Issues (e.g., headache, fullness of head, difficulty concentrating, disorientation, vertigo, and dizziness), Sopite and Anxiety Related Issues (e.g., fatigue, drowsy, bored, tired, anxiety and fear), and Gastrointestinal Feelings (e.g., nausea, awareness of stomach and burp). Altogether there were 19 items, rated in a form of Visual Analog Scale (VAS). The VAS was set from 0 to 10, which was in accordance with Chinese customs. Under such a VAS, 0 meant feeling nothing and 10 stood for unbearable, with 2.5, 5, 7.5 representing slight, moderate and severe, respectively.

2.2 Procedure

The subject was seated in a rotating chair and rested for three minutes. His blood pressure and heart rate were measured and these physiological parameters served as the baseline. The motion sickness questionnaire was also filled in, however, for most subjects, most items got values of 0. And the rotational session was set for 3 min unless the subject asked for a stop. During the session, the subject was required to keep his eyes closed or opened (in a randomly assigned sequence) and his BP and HR were measured again as soon as the rotating chair was stopped. The subject filled in the motion sickness questionnaire again. A break was given and the length of break was determined by the subject, until he felt he is totally recovered from motion sickness. Then his BP and HR were measured and another session of rotation (eyes opened or closed dependent on and opposite to the previous session) was started. Once again the length of rotating time was either up to the limit of endurance or for 3 min. After this session, his BP and HR were measured for the fourth time and the subject filled in the motion sickness questionnaire for the third time. The rotational sessions were video-taped by one of the experimenters.

2.3 Equipment

Since the rotation chair was operated by one of the experimenter manually, the rotation speeds were not stable. During the experiment we recorded the whole process with a video camera, and the rotation speed was calculated in a way that rotated degrees divided by rotation time. When the rotation time was up or the subject asked for a stop, the operator would try to stop the rotation chair gently, because in the pre-experiment we found that a sudden stop was quite irritating. The final analysis found that the rotation speeds ranged from 270 to 360 degrees/second, and the speeds of 30 subjects fitted for the normal distribution, both for the eyes opened and eyes closed sessions. Paired t-test was also carried out, and the results showed that there was no significant differences between the eyes opened and eyes closed rotation speeds. There were two seat belts settled on the chair, and one was fastened over the laps while the other was fastened around subject's waist. Besides, a horizontal steel bar was set before the subjects to ensure safety and the subject could use it as a handrail. In this way, the movements of subject were reduced and his safety can be assured. The apparatus for measuring BP and HR was a cuff-styled automatic electronic hemodynamometer with the model number CONTEC 08A, which could also be used for medical purposes.

2.4 Statistical Analysis

The statistical tests used in this study were the two-tailed paired t-test and Wilcoxon rank-sum test. Paired t-tests were used to compare if average values changed significantly for the same subjects under different conditions, such as eyes opened and closed conditions, or before and after rotations. The Wilcoxon rank-sum test was applied when comparing parameters of sensitive and resistant groups, which were of different lengths (n1 = 9, n2 = 12). Also Pearson linear correlation was utilized to explore whether two variables were significantly correlated or not. The statistical toolbox was with version 8.2 in MATLAB 2013a [20].

3 Results

The Pearson correlation was measured between the items in the psychometric tests and physiological parameters. It could be seen that there was a significant correlation between anxiety status and baseline heart rate ($r = 0.436$, $p = 0.042$); value of fear and baseline diastolic blood pressure had a significant correlation ($r = 0.543$, $p = 0.007$); baseline systolic blood pressure had a significant correlation with the headache item in rMSQ ($r = 0.371$, $p = 0.042$).

During the rotation sessions, the endurance time under eyes opened condition greatly shortened compared with that under eyes closed condition (mean 103.6 vs. 122.9 S, two tailed paired t-test, $p = 0.029$, df = 29), also with a reasonably much smaller number of rotations (mean 90.2 vs. 103.1 rotations, two tailed paired t-test, $p = 0.078$, df = 29). However, the rotation speeds between these two sessions were quite close (mean 285.16°/S (eyes opened) vs. 298.64°/S (eyes closed), two-tailed paired t-test $p = 0.077$,

df = 29). The endurance times and rotational cycles between two sessions were significantly correlated (Pearson correlation coefficient equaled 0.74, p = 8.5732e-07).

The paired t-test showed that both diastolic blood pressure (mean 71.20 vs. 75.33 mmHg, p = 0.020, df = 29) and systolic blood pressure (mean 122.60 vs. 126.33 mmHg, p = 0.026, df = 29) changed significantly after the eyes opened rotation; while after eyes closed rotation, only the systolic blood pressure (mean 121.80 vs. 127.90 mmHg, p = 0.005, df = 29) changed remarkably and so did the heart rate (mean 71.50 vs. 75.26 BPM, p = 0.002, df = 29).

During the whole experimental procedure, the subjects filled in the motion sickness questionnaire three times. Tables 1 to 4 show different symptoms of motion sickness under these three conditions (Rest is sitting on the chair for 3 min, ECRotation is Rotation with eye closed, EORotation is rotation with eyes opened). Paired t-tests were conducted with degree of freedom of 29, *, **, *** and **** stand for $p < 0.05$, $p < 0.01$, $p < 0.001$ and $p < 0.0001$, respectively. (Tables 2 and 3)

Table 1. Under three conditions, the average scores of Behavioral subscale

	Eye strain	Salvation	Sweat	Shaky
A(Rest)	1.033	0.017	0	0
B(ECRotation)	1.517	1.243	2.483	1.217
C (EORotation)	2.2833	1.6000	2.0333	2.1833
A-B tvalue	−1.364 n.s.	−4.879***	−4.953***	−4.067***
A-C tvalue	−3.839***	−5.42***	−7.314***	−4.903***
C-B tvalue	5.325****	3.976***	3.752***	1.039 n.s.

Table 2. Under three conditions, the average scores of Central neuro-system subscale

	Headache	Head tightness	De-concentration	Disorientation	Vertigo	Feeling dizzy
A (Rest)	0	0.133	0.367	0	0	0
B (ECRotation)	2.117	3.333	3.917	5.383	5.95	4.250
C (EORotation)	2.4333	4.2667	3.2667	5.2333	6.583	5.917
A–B tvalue	−4.456***	−5.948****	−6.653****	−8.495****	−11.444****	−8.011****
A–C tvalue	−5.151****	−7.430****	−8.815****	−10.89****	−12.559****	−9.856****
C–B tvalue	2.006 n.s.	0.902n.s.	2.796**	1.380 n.s.	1.602 n.s.	3.411**

Table 3. Under three conditions, the average scores of gastro-intestinal subscale

	Nausea	Stomch-awareness	Burp
A(Rest)	0	0	0*
B(ECRotation)	3.517	3.117	0.217
C(EORotation)	4.867	4.350	0.767
A–B tvalue	−5.991****	−5.461****	−1783 n.s.
A–C tvalue	−7.814****	−6.795****	−2.399*
C–B tvalue	2.844**	3.211**	1.553 n.s.

Table 4. Under three conditions, the average scores of Sopite-related subscale

	Fatigue	Drowsiness	Boredness	Tiredness	Anxiety	Fear
A (Rest)	1.35	1.283	0.817	0.917	0.117	0.033
B (ECRotation)	2.117	2.117	0.967	2.667	2.767	2.417
C (EORotation)	3.086	2.817	1.267	3.333	3.283	2.767
A–B tvalue	−2.269*	−1.753n.s.	−0.632 n. s.	−4.051***	−6.093****	−5.082****
A–C tvalue	−5.528****	−.3.330**	−.1.487n. s.	−5.679****	−6.618****	−5.475****
C–B tvalue	4.338***	2.242*	1.895 n.s.	3.808***	1.660 n.s.	0.975 n.s.

According to the values shown in Table 1, it can be concluded that under the eye opened rotation condition, the top five items with averaged highest scores are disorientation, feeling dizzy, vertigo, de-concentration and nausea.

According to the different lengths of rotation endurance time under eyes opened condition, the subjects were divided into three groups: the sensitive, the normal and the resistant groups, with the averaged endurance times of 27.444 ± 13.116 S (N = 9), 81.333 ± 21.059 S (N = 9) and 180 ± 0.0S (N = 12), respectively. The between-group statistical analysis of physiological, psychometric test values, and motion sickness questionnaire were conducted. Simplified coping style (resistant group: 28.333 ± 5.176, sensitive group: 22.667 ± 4.770, Wilcoxon rank-sum test, $p = 0.027$, n1 = 9, n2 = 12) and nausea item (resistant group: 3.375 ± 3.127, sensitive group: 7 ± 3.278, Wilcoxon rank-sum test, $p = 0.020$, n1 = 9, n2 = 12) in rMSQ was the only two items that possessed statistically significant differences between sensitive and resistant groups. However, concerning the averaged baseline blood pressure, the resistant group (SBP:120.5 mmHg, DBP:70.0 mmHg) was more close to the average values of the 30 people(SBP:121.8 mmHg, DBP:69.8 mmHg), while the sensitive group(SBP:127.0 mmHg, DBP:75.8 mmHg) was higher than the average value and the normal group (SBP:118.3 mmHg, DBP:63.7 mmHg) was slightly lower than the average value.

An explorative factor analysis (EFA) was applied on the eyes opened rMSQ values using SPSS 23.0, although the number of subjects in our study was only 30. However, there was study [21] which already justified the use of small samples when EFA was applied. Prior to the extraction of the factors, Kaiser-Meyer-Olkin(KMO) index and Barlett's Test of Sphericity were evaluated to assess the suitability of the data for factor analysis. The KMO index got a value of 0.746 and the Bartlett's Test of Sphericity was significant (p = 1.7479E-27, df = 171, Approx. Chi-Square = 453.801). The extraction method was principal axis factoring (PAF) with a Kaiser's criteria (eigenvalue > 1 rule). A total of 5 components (factors) were chosen with a cumulative percentage of variance of 81.167%. The rotation method was oblique olbimin with Kaiser normalization. Finally, we got an EFA result as show in Table 5.

Table 5. Results of exploratory factor analysis

	1	2	3	4	5
Eyestrain					0.930
Salvation			0.499		
Sweating			0.806		
Shaky					0.506
Headache			0.815		
Fullness of head	0.518				
Disconcentrating	0.776				
Disorientation					0.592
Vertigo					0.727
Dizzy					
Nausea		0.857			
Stomach awareness		0.880			
Burp				0.776	
Fatigue					
Drowsy			0.686		
Bored					
Tired					
Anxiety	0.527	0.690			
Fear	0.688	0.498			

Considering the relatively low value of burp, it may be desirable to keep less than 14 items in rMSQ. However, further studies should test more subjects with longer experimental time to improve revision of rMSQ.

4 Discussion

The baseline HRs were quite close among groups, and after rotation, the HR of resistant group (mean 75.083 BPM) was higher than that of the sensitive group (mean 72.889 BPM), but this could also be due to the fact that the resistant group endured for a much longer time. Moreover, it was reasonable that the anxiety status had a clear correlation

with baseline HR, which had been approved by other studies [22, 23]. Although there was relation among physiological parameters or psychometric test and physiological parameters, it was still not possible to identify the sensitive subjects depending on the physiological parameters or psychometric tests. A similar conclusion was reached by a meta-analysis which studied the visually induced motion sickness (VIMS)-sensitive and resistant subjects. It found out that there was no difference in baseline parasympathetic nervous system (PNS) tones between sensitive and resistant; however, those who were sensitive had a reduction in PNS tones [24]. In our case, we could only infer that some subjects were especially sensitive because of their relatively higher baseline blood pressure and more tests should be done before serious conclusions are drawn. Our results showed that the coping style had a significant difference between the sensitive and the resistant groups, and similar results were reached by other studies. Some of them claimed that fear or anxiety may contribute indirectly to motion sickness susceptibility [1, 3] while some others believe that learned control could be a contributing factor [25]. All these studies suggest that the individual sensitivities are correlated with psychometric characteristics, which is also the object of our future study.

In this study, two means of rotation were conducted with similar speeds and in a random order. However, the results showed that the endurance time of rotation with eyes opened was much shorter than that of the rotation with eyes closed, and this meant that it would be much more effective to use rotation with eyes open to find out the sensitive subjects. And it is not difficult to understand why open-eyed rotation is much more unbearable, which is attributable to the sensory conflicts, just as one of the classical explanations of motion sickness considered. The most common daily strategy dealing with motion sickness is to close eyes or even get to sleep, which is in good accordance with our results.

Another limitation about our study was the relatively short endurance time because of high rotation speeds. Although by this way we were able to find out the sensitive subjects efficiently, the extremely short exposures times which could be tolerated (with a mean value of 113 s) did not correspond to many motion sickness environments in the real world, apart from airsick trainee pilots in agile military aircraft. Consequently some symptoms may simply not have time to develop (e.g. the scores for ‘salivation’ were very low) and the applicability of these results to longer exposures lasting many minutes or even hours, in cars, coaches or ships, may be questioned. In the future studies, it would be interesting to re-test the rotation chair endurance time with variable speeds.

The average endurance time of the sensitive group for rotation with eyes opened was 27.4 s, while the resistant felt nothing for 180 s, which was the upper limit of the rotation time. These results roughly proved that the motion sickness resistant people had endurance at least more than six times of the motion sickness sensitive people, which can be considered an initial quantified index. Although some studies [7] have investigated the ratio of sensitive and resistant people, specific data is lacking. Thus, to quantify motion sickness should be one of our future goals. We assume that our ratio of sensitive group is much higher than the previous investigation, which could be attributed to the fact that the experiment is much more irritating. And it would be of great importance to train or modulate the sensitivities, which may shed light on future studies.

Although the studies about blood pressure and motion sickness haven't reached a uniform conclusion, most of them confirm that the blood pressure would increase along with the aggravation of motion sickness [26]. Some of the studies took the BP parameters into building a model to predict the symptoms of motion sickness. Another study found that sensitive subjects would have their blood pressure increased before the appearance of motion sickness, while the resistant subjects had no change in blood pressure [27]. In our study, all the subjects presented an increase of blood pressure during the experiment and the sensitive subjects had a relatively high baseline blood pressure. Moreover, there was a significant correlation between baseline systolic pressure and headache. Therefore, in the future studies, it is desirable to measure real-time blood pressure. Besides, some research regarding space motion sickness thought there could be a relationship between motion sickness and intracranial blood velocity, which affected the intracranial blood pressure and consequently increased the blood pressure [28]. This could be regarded as one of the explanations for the increase in blood pressure.

There was no significant difference between the sensitive and the normal groups, or the normal and the resistant groups in rMSQ values by Wicoxon rank-sum test while the sensitive and the resistant groups presented differences on the nausea item. From this result, it can be concluded that different subjects have almost no difference in subjective values of motion sickness except for nausea, thus it would be inappropriate to identify sensitive subjects depending only on the questionnaire.

From the abovementioned facts, it can be concluded that physiological and psychological parameters are strongly connected, and the sensitive and resistant groups have significant differences in some of these parameters. In this study, these differences have been discerned, but the sensitive subjects can still hardly be identified if relying only on these parameters. In the future studies, it would be worthwhile to lay more emphasis on the motion sickness sensitive people and give them a better evaluation. To obtain this goal, physiological and psychological parameters need more scrutinized examination and more subjects of different ages and genders must be involved.

Acknowledgements. This study was supported by Science and Technology Support Program 2014BAK01B01.

References

1. John, F.G.: Motion sickness susceptibility. Auton. Neurosci. **129**(1), 67–76 (2006)
2. Lackner, J.R.: Motion sickness: more than nausea and vomiting. Exp. Brain Res. **232**(8), 2493–2510 (2014)
3. Paillard, A.C., Quarck, G., Paolino, F., Denise, P., Paolino, M., Golding, J.F., et al.: Motion sickness susceptibility in healthy subjects and vestibular patients: effects of gender, age and trait-anxiety. J. Vestib. Res. **23**(4–5), 203–209 (2013)
4. Keshavarz, B., Hettinger, L.J., Kennedy, R.S., Campos, J.L.: Demonstrating the potential for dynamic auditory stimulation to contribute to motion sickness. PLoS ONE **9**(7), e101016 (2014)

5. Meyer, G.F., Shao, F., White, M.D., Hopkins, C., Robotham, A.J.: Modulation of visually evoked postural responses by contextual visual, haptic and auditory information: a 'virtual reality check'. PLoS ONE **8**(6), e67651 (2013)
6. Weimer, K., Horing, B., Muth, E.R., Enck, P.: How to study placebo responses in motion sickness with a rotation chair paradigm in healthy participants. J. Vis. Exp. **94**, 52471 (2014)
7. So, R.H.Y., Finney, C.M., Goonetilleke, R.S.: Motion sickness susceptibility and occurrence in Hong Kong Chinese. Contemporary Ergonomics, pp. 88–92 (1999)
8. Geer, J.H.: Fear Survey Schedule-II. Measure for Clinical Practice: A Sourcebook, p. 160. Free Press (1987)
9. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol. **3**(3), 203–220 (1999)
10. Muth, E.R., Stern, R.M., Thayer, J.F., Koch, K.L.: Assessment of multiple dimensions of nausea: the Nausea Profile (NP). J. Psychosom. Res. **40**, 511–520 (1996)
11. Keshavarz, B., Hecht, H.: Validating an efficient method to quantify motion sickness. Hum. Factors **53**(4), 415–426 (2011)
12. Gianaros, P.J., Muth, E.R., Mordkoff, J.T., Levine, M.E., Stern, R.M.: A questionnaire for the assessment of the multiple dimensions of motion sickness. Aviat. Space Environ. Med. **72**, 115–119 (2001)
13. Sugita, N., Yoshizawa, M., Tanaka, A., Abe, K.I., Chiba, S., Yambe, T., et al.: Quantitative evaluation of the effect of visually-induced motion sickness using causal coherence function between blood pressure and heart rate. In: 26th Annual International Conference of the IEEE, 2004 Sept 1–5, pp. 2407–2410. Engineering in Medicine and Biology Society, San Francisco (CA) (2004)
14. Apthorp, D., Nagle, F., Palmisano, S.: Chaos in balance: non-linear measures of postural control predict individual variations in visual illusions of motion. PLoS ONE **9**(12), e113897 (2014)
15. Cobb, S.V., Nichols, S.C.: Static posture tests for the assessment of postural instability after virtual environment use. Brain Res. Bull. **47**(5), 459–464 (1998)
16. Cobb, S.V.: Measurement of postural stability before and after immersion in a virtual environment. Appl Ergon. **30**(1), 47–57 (1999)
17. Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., Jacobs, G.A.: Manual for the State-Trait Anxiety Inventory. Palo Alto, CA (1983)
18. Xie, Y.N.: Simplified coping style questionnaire. Chin. Ment. Health J. **13**, 122–124 (1999). (Chinese)
19. Sherer, M., Maddux, J.E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., Rogers, R.W.: The self-efficacy scale: construction and validation. Psychol. Rep. **51**(2), 663–671 (1982)
20. MATLAB and Statistics Toolbox Release: The MathWorks Inc. Natick (2013)
21. de Winter, J.C.F., Dodou, D., Wieringa, P.A.: Exploratory factor analysis with small sample sizes. Multivar. Behav. Res. **44**(2), 147–181 (2009)
22. Knight, W.E.J., Rickard, N.S.: Relaxing music prevents stress-induced increases in subjective anxiety, systolic blood pressure, and heart rate in healthy males and females. J. Music Ther. **38**(4), 254–272 (2001)
23. Watkins, L.L., Grossman, P., Krishnan, R., Sherwood, A.: Anxiety and vagal control of heart rate. Psychosom. Med. **60**(4), 498–502 (1998)
24. Farmer, A.D., Al Omran, Y., Aziz, Q., Andrews, P.L.: The role of the parasympathetic nervous system in visually induced motion sickness: systematic review and meta-analysis. Exp. Brain Res. **232**, 2665–2673 (2014)

25. Cowings, P.S., Toscano, W.B.: The relationship of motion sickness susceptibility to learned autonomic control for symptom suppression. Aviat. Space Environ. Med. **53**(6), 570–575 (1982)
26. Blue, R.S., Pattarini, J.M., Reyes, D.P., Mulcahy, R.A., Garbino, A., Mathers, C.H., et al.: Tolerance of centrifuge-simulated suborbital spaceflight by medical condition. Aviat. Space Environ. Med. **85**(7), 721–729 (2014)
27. Serrador, J.M., Schlegel, T.T., Black, F.O., Wood, S.J.: Cerebral hypoperfusion precedes nausea during centrifugation. Aviat. Space Environ. Med. **76**(2), 91 (2005)
28. Karunaharan, C.P., Atkov, O.: Measurements of raised intra-cranial pressure, a cause of space motion sickness. In: International Space Station, pp. 251–253. Springer Netherlands (2000)

Modeling the Real World Using STISIM Drive® Simulation Software: A Study Contrasting High and Low Locality Simulations

Craig K. Allison(✉), Katie J. Parnell, James W.H. Brown, and Neville A. Stanton

Faculty of Engineering and the Environment, University of Southampton,
Boldrewood Innovation Campus, Southampton, SO16 7QF, UK
{Craig.Allison,kp4g13,J.W.Brown,N.Stanton}@soton.ac.uk

Abstract. This paper presents work undertaken to develop a high locality simulation of a real world driving environment within the STISIM Drive® v3 software. The paper considers the tools and techniques used to develop the simulation model and then examines the impact that the addition of realistic models has on driver performance metrics. In order to test the developed simulation model, participants drove both a high and low locality version of the same route in a driving simulator as well as the real world route. Drivers' levels of immersion, perception of presence, performance metrics and subjective workload ratings were considered. Results from an initial pilot study suggest that individual differences play a role in participants' levels of presence between simulations. Results are discussed in terms of potential future research.

Keywords: Human factors · Simulation · Tool development · STISIM

1 Introduction

Driving simulators have facilitated research into a range of driving behaviors within Human Factors research, including responses to automated driving [1], the potential distractive effect of in-vehicle technologies [2] and the environmental impact of different driving behaviors on fuel economy [3]. Simulators allow naturalistic behaviors to be studied under high levels of control as well as enabling the manipulation of variables found in the real world, within a low risk and safe environment. This is useful within the driving domain where on-road studies are associated with large ethical concerns regarding the drivers, as well as other road users, safety. Yet, the utility of driving simulators is dependent on their ability to accurately replicate real-world conditions. Often, the simplicity of simulation models used within a simulator setting mean results cannot be reliably applied to the real world without further testing. Shechtman et al. [4] vocalized the concern that work is needed to support the use of driving simulators in order to accurately reflect real world environments and situations by suggesting that "driving simulators must be validated before they are used for driving assessment and/or training."

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_87

One popular piece of driving simulation software used for research purposes is STISIM Drive® v3 (STISIM). STISIM facilitates the recording of a variety of driving metrics as well as allowing freedom in scenario development. This program has been used to explore a variety of different driving related influences and tasks including the effect of alcohol on driving performance [5], the distractive effects of in-vehicle technology [6], and even the effects of chewing gum whilst engaged in the driving task [7].

As mobile beings, humans possess a spatial reference system, allowing them to place objects within a spatial array [8] and understand their surroundings. One influential theory accounting for how individual understand their surroundings is the Landmark-Route-Survey model [9]. This suggests that individuals learn about salient visual items within their environment, the routes that link these salient landmarks, before finally developing an ornithological or "birds eye view" understanding of the environment. As landmark cues are salient markers within a route, presenting these cues within a simulated environment should allow an individual to place their current position within a known spatial context. By placing an individual within a simulation of a known environment, populated by known salient landmarks, it would be expected that participants driving would more accurately match the real world, and that participants would experience a greater level of involvement within the environment, referred to in the literature as higher levels of presence [10].

Salient localized features do not come as standard within typical simulation tools such as STISIM. It would be incredibly difficult to create a package that could be customizable to possess local route features to suit all users needs. Therefore, within most simulator research that has utilized this tool, local features are absent. Yet, it is suggested here that the inclusion of local landmark cues within simulation models may influence the driver's levels of presence. To explore the effect of local landmark cues, a simulation model of a real roadway environment was developed using STISIM. Central to the development of the simulation environment was the creation of a custom library of models, used to supplement the default set available within STISIM. Three open source programs were used to create the models, Inkscape, GIMP and Blender, with the completed models subsequently imported into STISIM. The development of the custom objects included capturing adequate textures and images of the real environment and applying appropriate scaling to ensure their accurate representation within the simulation model.

To assess whether the inclusion of custom objects influenced participants driving metrics and subjective experience during the simulation, two simulations were developed, a high locality simulation, which included the developed custom landmark objects, and a low locality simulation, which only used objects available in the standard STISIM library. In addition, participants were invited to drive the real world route in order to examine whether presence within the simulation differed from their presence on road. To summarize, this paper aims to explore the effects of local landmark cues on drivers subjective level of presence and speed, both with simulations and compared to on road driving.

2 Experimental Method

2.1 Participants

Seven participants, (4 Male, Mean age = 28.5 years and 3 Female, Mean age = 32.3 years) were recruited. Participants were required to be over the age of 25 and hold a valid UK driving license with no more than 3 points on their license. Participation was voluntary and the study lasted approximately 1.5 h. As the study was looking to explore the effect of local landmark cues it was important that all participants were familiar with the area that was being constructed within the simulation (Southampton and the surrounding area). Therefore, participants were recruited that lived or worked in Southampton and frequently (2 times a week or more) drove the roads included in the route. Ethical approval was granted from the University of Southampton research ethics committee (Ergo ID:25422). Informed consent was obtained from each participant prior to the start of the study.

2.2 Equipment

University Vehicle. For the on-road condition participants drove the University of Southampton TRG SRIF2 instrumented vehicle. This was a right-hand drive Fiat Stilo with an automatic transmission.

Driving Simulator. In the two conditions that required the participants to drive simulated routes, the Southampton University Driving Simulator (SUDS) was used. This is comprised of a fixed base right-hand drive Land Rover Discovery, with an automatic transmission. STISIM M500W wide-field-of-view system was used to create the simulated environment. Three screens positioned in front of the vehicle captured the forward projection of the road scene with 135-degree field of view.A single screen at the rear of the vehicle shows the projected rear-view image that can be seen in the vehicle's rear-view mirror (Fig. 1).

Fig. 1. The university vehicle used for real world driving and the SUDS driving simulator, University of Southampton.

2.3 Questionnaires

In order to test the level of presence felt by participants the "Presence Questionnaire" [11] was presented following each trial. In addition, participants were required to complete the NASA-TLX to measure subjective workload. In order to gain an understanding of the underlying characteristics of the participant to feel immersed in virtual environments, participants' completed the Immersive Tendencies Questionnaire [10].

2.4 Procedure

All participants completed a demographics questionnaire at the start of the experiment. A map of the route participants were asked to drive was then presented (Fig. 2). The route was 13.2 miles long, taking approximately 30 min to drive. The route contained a variety of different road types including motorways (Highways), residential roads, rural roads and urban roads The order participants completed the three experimental conditions (On-road (OR), low-Locality simulation (LL), and high-locality simulation (HL))

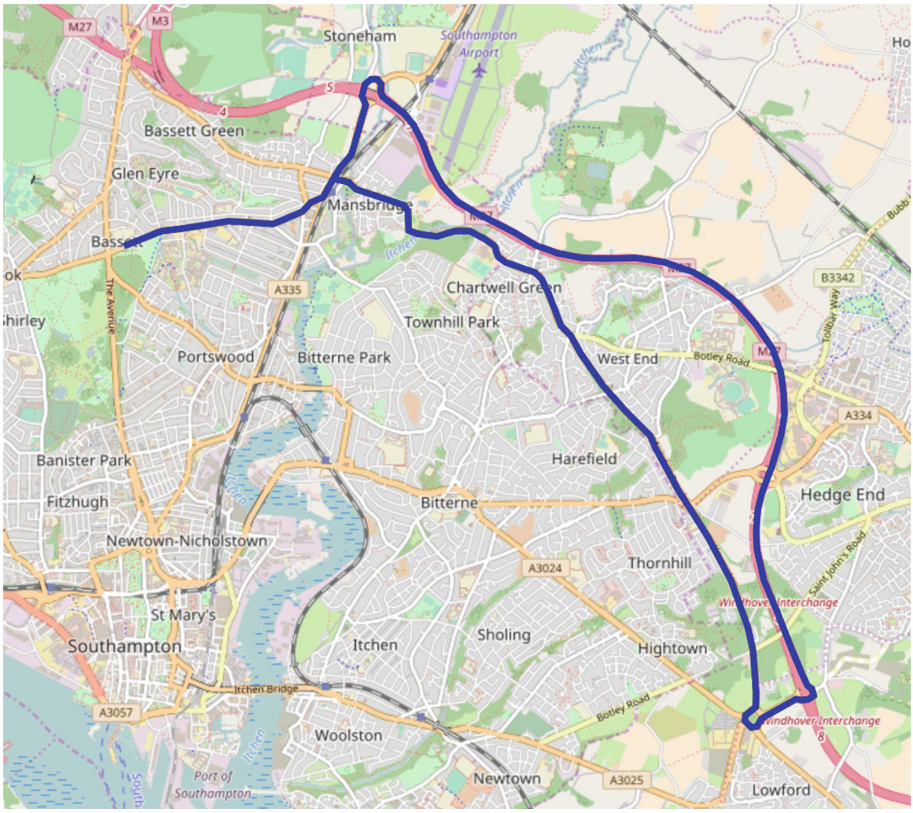


Fig. 2. The route participants were asked to drive within the study. Map produced using OpenStreetMap.

was counterbalanced to limit potential order effects. Participants completed a 5-min training drive in the driving simulator before completing the simulator trials to familiarize themselves with driving in a simulator. They were also encouraged to drive the instrumented vehicle around the car park before they began the on-road trial to acquaint themselves with the vehicle and its capabilities. For all conditions, a member of the research team sat in the passenger seat and asked the participant to drive the constructed route, providing directions where necessary. At the end of each trial, participants completed both the presence questionnaire and NASA-TLX before progressing to the next trial. Once participants had completed all three trials, they completed the immersion questionnaire. Figure 3 presents the same scene from the real world, and both low and high locality simulation environments.

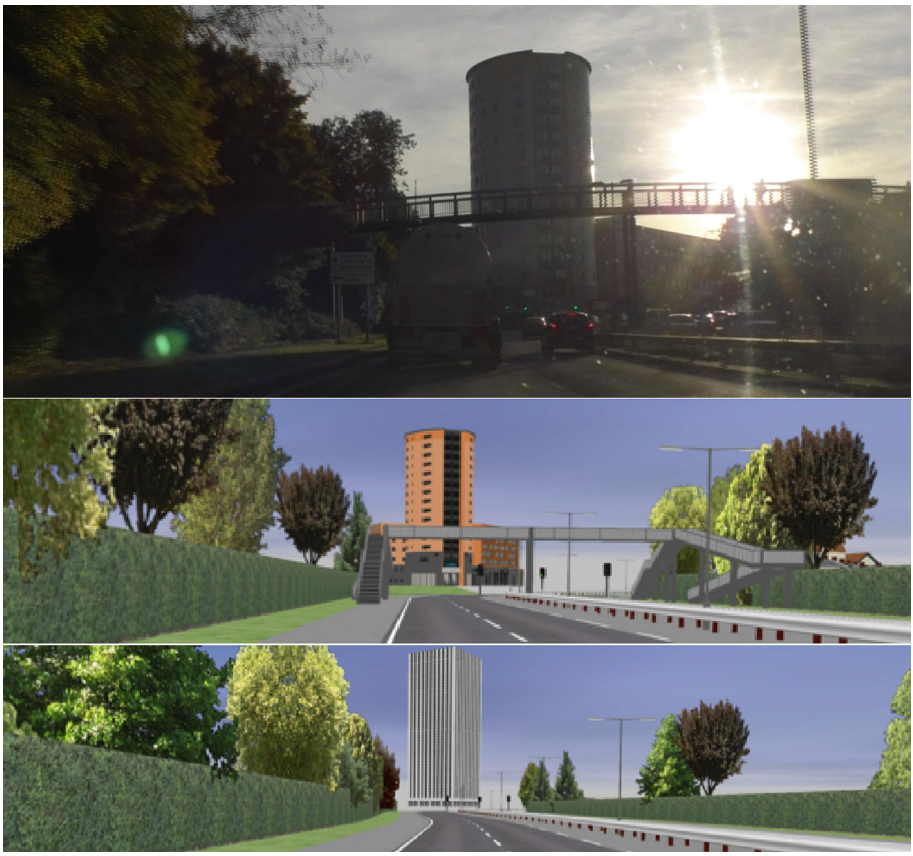


Fig. 3. A scene from the real world, and both high locality and low locality simulations.

3 Results and Discussion

As the data reported within this paper examines responses from an initial pilot study of seven participants, no inferential statistics tests have been calculated. Rather, the results are discussed in terms of current trends within the data, highlighting areas which require additional attention and data to be collected. The use of pilot studies to trial simulator studies is well known [4], and has been used within this simulation software to good effect before [12].

3.1 Immersion and Presence

Figure 4 presents participants' level of immersion as percentages. It can be seen from the graphs that large individual differences are apparent in immersion scores. This is apparent when examining participants "Play/Gaming" scores, although no participants reported themselves as particularly high in "Play/Gaming", participant 5 did score very low in this metric. Participants also reported themselves as relatively high in "Focus" with the exception of Participant 6.

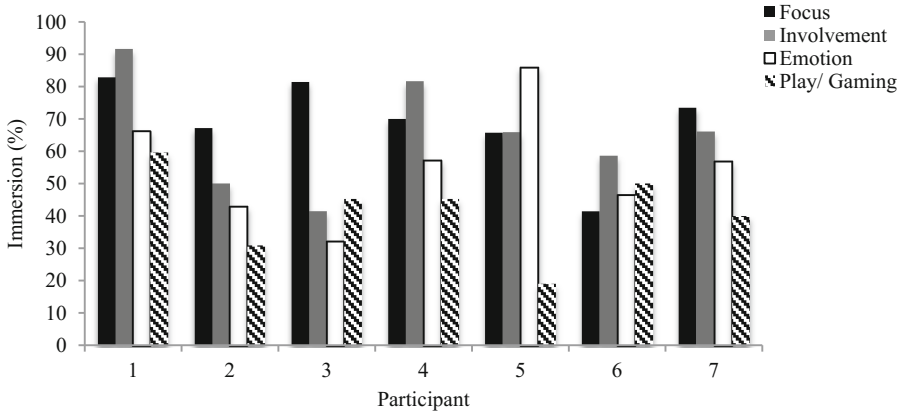


Fig. 4. Participants level of self reported immersion.

Figure 5 presents recorded presence scores for each participant, by trial type. For all participants it can be seen that presence was rated highest in the on road trial, an expected outcome, however mixed results can be observed when examining the simulation environments. This initial finding suggests that increased detail and local features does not guarantee an increase in recorded presence for all participants. However, there appears to be a link between presence within the simulations and reported levels of "Play/Gaming". Participants who reported low "Play/Gaming", for example participant 5, reported greater presence within the low locality simulation. In contrast, participant 1, who recorded the highest level of "Play/Gaming" expressed significantly greater presence within the high locality simulation. Further work is warranted to examine the interplay of these two variables.

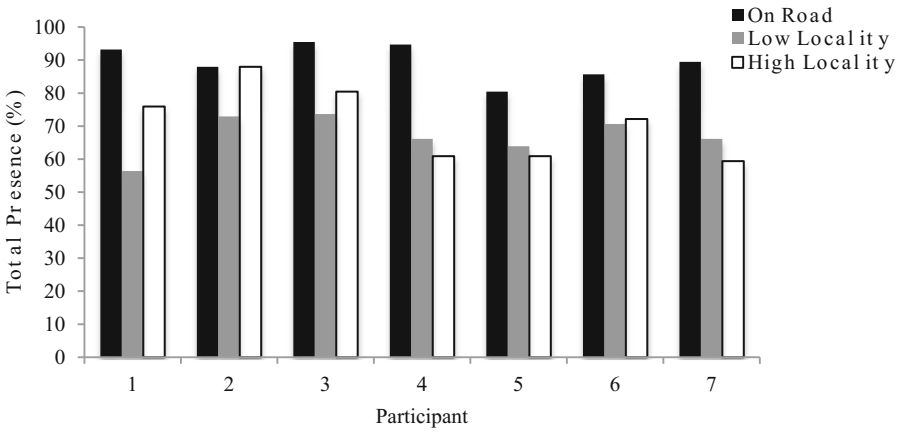


Fig. 5. Participants self-reported presence across the three conditions.

3.2 Workload

Figure 6 presents the mean self reported workload scores for each NASA-TLX sub-scale. It can be seen that participants recorded lowest workload when driving on road. Whilst workload is typically higher in the low locality version of the simulation, participants experienced considerably greater mental demand in the high locality simulation. The root cause of the increase in this sub-component is however currently unknown, and may be a consequence of the high salience of the custom objects present within this trial. Due to the increased graphical complexity of these objects, participants attention may have been artificially drawn to these objects, increasing the relative demand of the driving task. Future research utilizing eye-tracking technology is required to confirm this explanation.

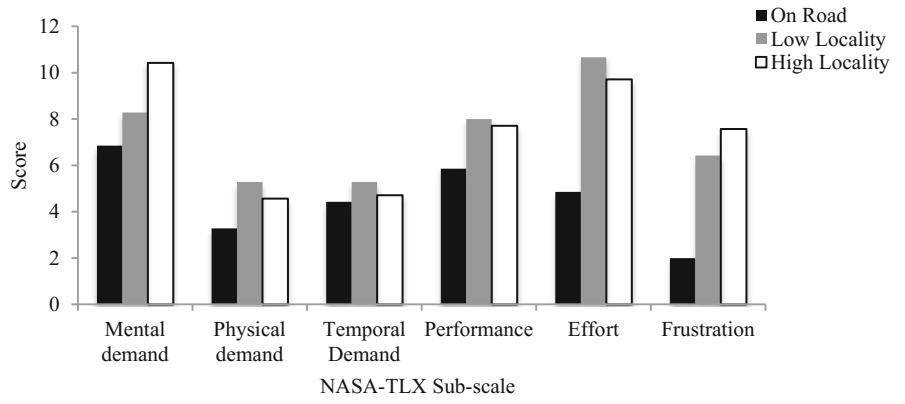


Fig. 6. Participants self reported workload, based upon NASA-TLX scores.

3.3 Speed Profiles

Figure 7 presents the mean speed profiles recorded within the simulation environment. It appears from Fig. 7 that participants drove at generally lower speeds within the high locality simulation than the low locality simulation.

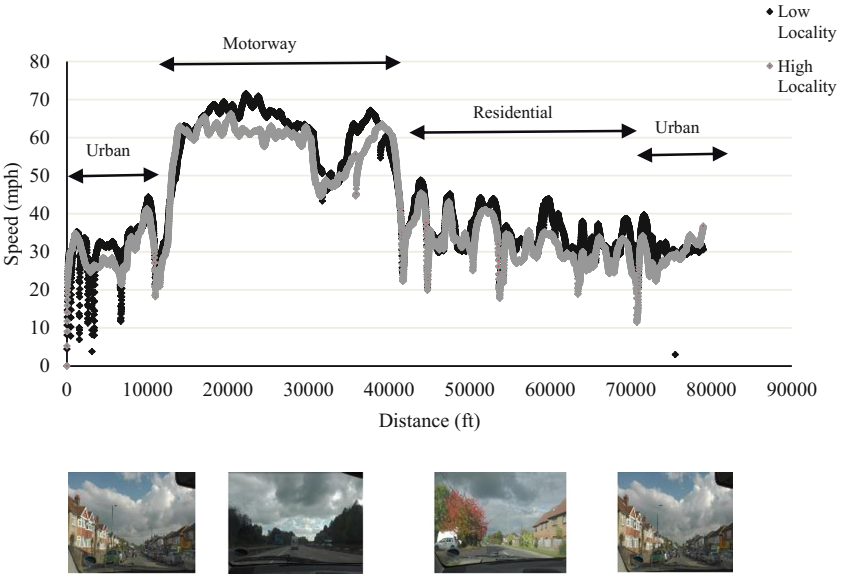


Fig. 7. Speed profiles recorded with both low and high locality simulation trials.

Several explanations can be posited to suggest why participants speed varied as a consequence of the simulation locality. It is possible that there was a link between participants subjective workload and their speed, with higher workload being associated with lower speed. Alternatively, it can be posited that as the high locality trial offered participants a greater number of local landmark cues with known dimensions, which could be used to judge relative speed of travel in addition to the vehicles' speedometer. Further research is needed to explore this finding in greater depth.

3.4 Lessons Learnt and Future Work

This study provided initial pilot data regarding the use of a simulation environment based on a real-world environment that was familiar to participants, using specially created land cues. The study highlighted future opportunities for further studies and began to consider the importance of individual differences which should be examined when conducting simulator based research. Further work, recruiting a greater number of participants is needed to support the current findings and allow for the meaningful use of statistical testing. Participants from the current study are all familiar with the area of Southampton that was modeled and driven within the current study. Engaging with a

wider demographic, and employing a scale of familiarity is needed in order to establish how locality within a simulation is affected by previous knowledge of the local area. In addition, exploring the impact of variables such as age and driving experience may offer future avenues for research.

A methodological limitation of the study relates to the use of the Fiat Stilo for the on-road component of the study. Participants were not able to drive their own cars due to concerns over insurance, and as such all participants drove the university vehicle. Although the use of a consistent vehicle enables the direct comparison of driving behaviors, participants may have handled this car differently than their own vehicle due to established knowledge of the vehicle's capabilities. In addition, participants may have been less experienced driving with automatic transmission, and as such were required to adapt their behavior. The use of the university vehicle may have influenced subjective ratings recorded in the study and the role of vehicle should be explored further.

Future work should seek to collect objective data during the on-road driving trial including speed and headway to enable direct comparison between these variables for both on road and simulator.

3.5 Conclusions

This paper has presented an initial pilot study data exploring the impact of including custom local landmark objects within a simulated driving route. The paper considered the impact of these objects on participants ratings of presence, awareness of speed and self-rated workload. Although, mixed findings were apparent, data suggests that individual differences play a role in the impact of custom objects. Participants with greater reported levels of immersion, particularly "Play/Gaming", reported greater levels of presence within the high locality route. Furthermore, it appears that the inclusion of the custom objects impacted participants awareness of speed, with participants driving faster in the low locality simulation, which included no landmark custom objects. This can be directly contrasted with data suggesting that participants level of self perceived mental workload was highest in the high locality simulation, suggesting the task was more mentally taxing. Further research is required to fully understand the impact of local landmark cues and high locality simulations on participants performance in the driving laboratory, however it may be the case that the availability of local landmark cues does not always lead to an increased feeling of presence within simulator studies.

Acknowledgments. This work was partially supported by supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/N022262/1 "Green Adaptive Control for Interconnected Vehicles", EPSRC grant EP/G036896/1 and industrial sponsor Jaguar Land Rover, under the Industry Doctoral Training Centre in Transport and the Environment. The researchers would like to thank Karen Ghali for assistance in accessing the university instrumented vehicle. The researchers also extend their thanks participants who donated their time to be part of the study.

References

1. Young, M.S., Stanton, N.A.: What's skill got to do with it? Vehicle automation and driver mental workload. Ergonomics **50**(8), 1324–1339 (2007)
2. Lee, J.D., Caven, B., Haake, S., Brown, T.L.: Speech-based interaction with in-vehicle computers: the effect of speech-based e-mail on drivers' attention to the roadway. Hum. Factors **43**(4), 631–640 (2001)
3. McIlroy, R.C., Stanton, N.A., Godwin, L., Wood, A.P.: Encouraging eco-driving with visual, auditory and vibrotactile stimuli. IEEE Trans. Hum. Mach. Syst. (2017, in press)
4. Shechtman, O., Classen, S., Awadzi, K., Mann, W.: Comparison of driving errors between on-the-road and simulated driving assessment: a validation study. Traffic Inj. Prev. **10**(4), 379–385 (2009)
5. Mets, M.A., Kuipers, E., Senerpont Domis, L.M., Leenders, M., Olivier, B., Verster, J.C.: Effects of alcohol on highway driving in the STISIM driving simulator. Hum. Psychopharmacol. Clin. Exp. **26**(6), 434–439 (2011)
6. Ranney, T.A., Baldwin, G.H., Vasko, S.M., Mazzae, E.N.: Measuring distraction potential of operating in-vehicle devices (No. HS-811 231) (2009)
7. Yoo, I., Kim, E.J., Lee, J.H.: Effects of chewing gum on driving performance as evaluated by the STISIM driving simulator. J. Phys. Ther. Sci. **27**(6), 1823–1825 (2015)
8. Shelton, A.L., McNamara, T.P.: Systems of spatial reference in human memory. Cogn. Psychol. **43**, 274–310 (2001)
9. Siegel, A.W., White, S.H.: The development of spatial representations of large-scale environments. In: Reese, H.W. (ed.) Advances in Child Development and Behaviour, vol. 10, pp. 9–55. Academic Press, New York (1975)
10. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence Teleoper. Virtual Environ. **7**(3), 225–240 (1998)
11. Witmer, B.G., Singer, M.J.: Measuring Immersion in Virtual Environments (ARI Technical Report 1014). U. S. Army Research Institute for the Behavioural and Social Sciences, Alexandria (1994)
12. Banks, V.A., Stanton, N.A., Harvey, C.: What the drivers do and do not tell you: using verbal protocol analysis to investigate driver behaviour in emergency situations. Ergonomics **57**(3), 332–342 (2014)

What Driving Abilities Do Racing Video Games Stimulate? Rating the Levels of Realism Experienced in Commercial Racing Video Games

Daniele Ruscio[1,2(✉)]

[1] Department of Psychology, Università Cattolica del Sacro Cuore, Milan, Italy
daniele.ruscio@unicatt.it
[2] Department of Mechanical Engineering, Politecnico di Milano,
via G. La Masa 1, 20156 Milan, Italy
daniele.ruscio@polimi.it

Abstract. The relationship between racing video games and road accidents is producing relevant research, but there are no studies (or game-content rating systems) that analyze the quality of the driving experience in video games. The aim of the present research is to propose a taxonomy of the driving experience content, to evaluate and cluster the realism of the driving experience created by different racing video game. A selection of 11 commercial video games, previously considered in safety researches, was used to build and test a Driving Experience Analysis – Rating Grid (DEA-RG). Levels of "environmental realism", "driving realism" and "accident realism" were assessed. DEA-RG scores showed that different categories of racing games generate very different user's experiences in driving simulations, with mixed levels of realism even within game-categories previously considered as homogenous. Implications for video game developers and road-safety researchers are discussed, for more systematic research hypothesis.

Keywords: Video games · Driving experience · Assessment · Human factors · Accident prevention

1 Introduction

Young drivers are more exposed to road accident compared to other road users [1]. One of the main reason behind this predisposition is that young driver tend to overestimate their driving abilities while underestimating the danger level of critical driving situations [2]. In addition, young drivers represent the main users of racing video games [3]. The relationship between racing video games and road accident has begun to produce relevant research in the recent year, suggesting a possible mediating factor in the road accidents exposure [4]. Historically, the effect of video games has been studied in relation to violence, aggressive attitudes and social behaviors [5]. According to General Aggression Model, aggressive video games can trigger aggressive ideas that could generate aggressive behaviors in real-life [6], depending on mediating factors like

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_88

dispositional traits and self-perception of acting aggressively in the video game [7]. Fisher and colleagues have demonstrated that a similar paradigm can be found also for racing video games in relation to risk-taking attitudes while driving. The experience of playing racing video games can trigger careless attitudes toward driving behaviors, which are mediated by the self-concept of being a risky driver in racing video games [8]. These effects were more evident when young gamers played racing games that reward and encourage traffic violations [4].

However, playing video games can also have several positive effects, and several serious games are being developed to train driving skills [9]. Playing video games can improve learning efficacy [10], enhance spatial resolution of vision [11], improve speed of visual processing [12] and target localization [13]. Video games can also can also quicken decision-making process and improve spatial competence, increasing the direct experience with routes, nodes, landmarks [14] that facilitate the ability to use cognitive mapping to explore an environment. All skills and cognitive abilities that could be crucial for safe driving [15].

Still, it is not clear what specific features of racing video games can directly affect the driving abilities. In addition, current research on video games has used racing video games as "experimental factor", without a systematic analysis of the content of the driving-experience, ignoring the way the potential user will experience in the virtual driving simulations. Video game content rating systems (e.g. PEGI and ESRB) are not evaluating the potential negative driving-related contents in commercial video games and there are currently no researches that analyze the quality of the driving experience in video games. Experimental studies usually use prepackaged labels to describe racing video games, comparing for instance "*arcade-style racers*" vs. "*racing simulators*", "*drive'em up games*" vs. "*circuit racing games*" [8], or "*typical street-racing games*" vs. "*F1 games*" [4]. This labelling process is usually based on the type of reward system (i.e. games that reward traffic violations are described as "*drive'em up games*", while games that rewards accuracy of driving are described as "*circuit racing games*"), or it is based on a list of a series of game titles that are supposed to implicitly define the category: "*drive em' up' style racing game such as VG1, VG2, etc. and realistic style racing games such as VG3, VG4, etc.*" [16]. Thus, the racing video game effect has been studied, so far, in relation to the reward system and not on the analysis of the actual content of driving experience or on the level of realism that are created by a video game. Does it mean that all "*drive'em up*" games present similar driving experiences? Do all "*circuit racing game*" propose a realistic driving experience? Do all the games that reward traffic violations entail negative effects on the driving skills?

Generate an analysis tool to rate the driving experience could help to understand the way video game content could eventually generate positive or negative effects on real-driving experience.

1.1 Rating the Driving Experience

To understand the content of driving experience in video game, it is required a definition of the real-life elements that define the driving experience. Following the driving task analyses tool for traffic safety research [17] the overall driving task can be divided

into different subtasks: e.g. controlling the vehicle, monitoring the changes in the road intersection and managing the interaction with other road users. Each subtask requires that the driver will use different cognitive abilities, like risk perception, working memory actions and decision-making processes [18]. Therefore, understanding the way each subtask is represented in a racing video game, can provide a measure of the complexity of the driving experience provided by video games. Realism score can be attributed to the driving-subtasks, creating a labelling system based not only the reward system, but also on the way video games represent the overall driving experience. Using the revised approach to Grounded Theory [19], it is possible to analyze what elements of the driving task can be categorized in subtask, to understand the complexity-level that is recreated by racing video games.

The aim of the present research is to propose a taxonomy of the content of the driving experience of racing video games to evaluate realism of the driving experience created by different racing video game. More precisely the objectives are: (1) built a rating grid to evaluate different aspects of the driving experience of racing video games; (2) implement the analysis grid to evaluate the realism of different video games; (3) test if the commercial labels *drive'em up game*" vs. "*circuit racing game*" can describes coherent clusters within the driving experience assessed by the rating grid.

Fig. 1. The eleven racing video games selected to test and build the driving-experience rating grid.

2 Methods

2.1 Sample

A selection of 11 commercial video games was used to build and test the driving experience rating-grid. The sample was build selecting racing video games that were already used in previous safety researches and it was completed trying to match for any video game categorized as "*drive'em up*" (games that reward traffic violations), one "*circuit racing game*" (that rewards accuracy of driving) published in the same time period. The resulting sample (Fig. 1) consisted of 5 "drive'em up" titles: *Burnout Paradise Ultimate Box [EA Swiss Sarl]; Need for Speed Undercover, Need for Speed Pro Street [EA Swiss Sarl]; FlatOut Ultimate Carnage [Bugbear Entertainement & Empire Interactive]; Juiced 2 HIN [THQ International Operations]*, 5 "circuit games" games: *RaceDriver GRID [Codemaster]; Test Drive Unlimited [Atari Europe]; Race Official WTCC Game [Atari United Kingdom Limited]; Overspeed High Performance Street Racing [City Interactive SA]; rFactor [Image Space Incorporated];* and 1 video game considered as "drive'em up", but that penalize traffic violations since the players are traffic police officers: *CrashTime 2 [RTL Games]*.

2.2 Driving Experience Analysis – Rating Grid (DEA-RG)

To assess the driving experience, the driving task was considered composed by three subtasks: perception of the road environment, vehicle control, and interaction with other road users. Grounded Theory approach was adapted to construct the actual evaluation grid tool, starting from the systematic categorization of the driving subtask inside the video games gameplay. Initial open coding consisted of assignment of conceptual labels that described for each game: (1) the "environmental realism"; (2) the "driving experience realism"; (3) and the "accident realism". The labeling system was refined through repeated inspection of the gameplay experience of different video games and it lead to three sub-categories for each conceptual label. The "environmental realism" categories that emerged were: 1.1 structural elements of the background environment; 1.2 structural elements of the roadway environment; 1.3 nodes, landmarks, and edges that define the user's movement inside the environment. The "driving realism" categories that emerged were: 2.1 driver's aids for driving experience; 2.2 factors that affect the driving performance; 2.3 accuracy of vehicle responses. The "accident realism" considered were: 3.1 physical components of the car that could be damaged; 3.2 external and infrastructural elements that could be involved in road accidents; 3.3 rewards, fines and penalties in consequence of road accidents.

Secondary axial coding involved the definition of concrete series of items that were able to signal the presence/absence of the conceptual labels in the video game. For instance: the structural elements that compose the background (1.1) were coded as the presence/absence of buildings variations, presence/absence of different road pavements, day-night cycle and changes in weather conditions. The driver's aids for driving experience (2.1) were coded as presence/absence of rearview mirror, side view mirrors, gauges and blinkers. The physical components of the car that could be damaged (3.1)

were coded as presence/absence of: crumpling zones, external car body damage, internal working parts critical for driving. All the secondary coding are shown in Fig. 2. The presence of each item was scored assigning one point every time the element was tracked in the video game.

Final selective coding was used to create the Driving Experience Analysis - Rating Grid (DEA-RG), that generated an overall score for each conceptual label, for each racing video game considered (Fig. 2).

Fig. 2. Driving Experience Analysis - Rating Grid (DEA-RG) was build using a Grounded Theory approach, to understand the way racing video games represent the driving environment, the driving experience and the road accidents. The driving content in racing video game was generated considering a grade for presence/absence of elements from three five-item subscales for each category. The total scores were used to create a rating score of the driving experience present in the racing video games.

2.3 Calculation

First: the DEA-RG scores for all the realism categories were calculated assigning one point every time one element of the subscale was tracked in the game. For instance, if one video game presented rearview mirrors, side view mirrors, gauges, blinkers and the option to drive using driver's view prospective, then 5 points were assigned to the "driver's aids factor" in realism of driving experience. If another video game presented only rearview mirror but no side view mirrors nor other driver's aids, then only 1 point was assigned for that particular subscale. The scores for each category were standardized into t-scores and simplified to a discrete scale from 0 to 10, with 5 as a mean value and 1 as standard deviation, to allow between games comparison.

Second: starting from the DEA-RG scores for each video game, a Hierarchical Cluster Analysis was performed using Ward's methods to define dendrogram linkage with the minimum increase in the sum of squared distance from the scores of all video games in all the categories. Additional *k*-means clustering procedure was used to calculate the values of the cluster centers for each realism categories.

3 Results

The DEA-RG scores showed that commercial racing games engage the user into a wide range of driving experiences. Every video game presented a different combination of realism levels (Fig. 3) and not always high scores in one category corresponded to high scores in the other two categories. Some racing video games presented high scores in the way the driving environment is graphically represented, but at the same time very low scores in the driving realism or in accident realism.

No game presented simultaneously high scores for environmental and driving experience realism and accurate representation of the consequence of road accidents.

It was also possible to define some similarities between different video games. Using repeated hierarchical cluster analysis, it was possible to provide a numerical quantification of the similarities for each category of realism.

3.1 Environmental Realism

Applying cluster analysis method to the "environmental realism" scores, three clusters with minimal-increase of the sum of squared distance emerged. The first cluster presented a *k*-mean center value of 49.00 and it included the three games with the higher scores for the "environmental realism": *FlatOut UC.*; *CrashTime2*; *BurnoutParadise*. The second cluster presented a *k*-mean center value of 34.80 and it included five games with medium scores for the "environmental realism": *Need for Speed Undercover*; *Juiced 2*; *RaceDriver GRID*; *Test Drive Unlimited*; *rFactor*. The last cluster presented a *k*-mean center value of 18.33 and it included three games with the lower scores for the complexity of the driving environment: *Race Official WTCC Game*; *Need for Speed Pro Street*; *Overspeed.*

3.2 Driving Realism

For the "driving realism" scores, three clusters emerged from the minimal-increase analysis of the sum of squared distance within the video games scores. The first cluster presented a *k*-mean center value of 47.67 and it included the three games with best scores for the realism of the driving experience: *Race Official WTCC Game*; *RaceDriver GRID*; *Test Drive Unlimited*. The second cluster presented a *k*-mean center value of 31.62 and it included four games with medium scores for the "driving realism": *Overspeed*; *CrashTime2*; *Juiced 2*; *rFactor*. The last cluster presented a *k*-mean center value of 16.00 and it included four games that do not represent a driving experience that can be considered realistic: *FlatOut UC.*; *BurnoutParadise*; *Need for Speed Pro Street*; *Need for Speed Undercover*.

3.3 Accident Realism

Considering the "accident realism" scores, only two clusters emerged from the minimal-increase analysis of the sum of squared distance. The first cluster presented a *k*-mean center value of 30.98 and it included four games with very high scores in the way road accident are represented: *FlatOut UC.*; *CrashTime2*; *BurnoutParadise; RaceDriver GRID; Race Official WTCC Game*. The second cluster presented a *k*-mean center value of 12.66 and it included seven games where accidents and collision almost do not produced any damage to the vehicle: *Need For Speed Undercover*; *Juiced 2*; *Test Drive Unlimited*; *rFactor*; *Need For Speed Pro Street*; *Overspeed*.

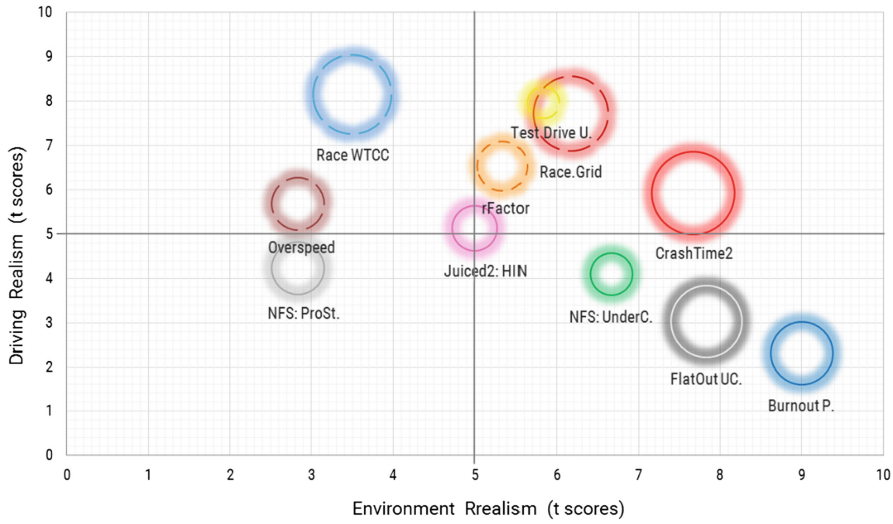


Fig. 3. DEA-RG scores plotted using "environmental realism" on x-axe, "driving realism" on y-axe and using "road accident realism" for the dot width. "*Drive'em up*" games are represented by solid lines, "*circuit racing*" games by dashed lines.

3.4 Overall Driving Experience Realism

Considering the overall DEA-RG scores for all the three realism categories, it was possible to obtain a synthetic index to describe the driving experience of each video game. Repeating the hierarchical cluster analysis for the overall DEA-RG scores it was possible to get four groups of racing video games (Fig. 4).

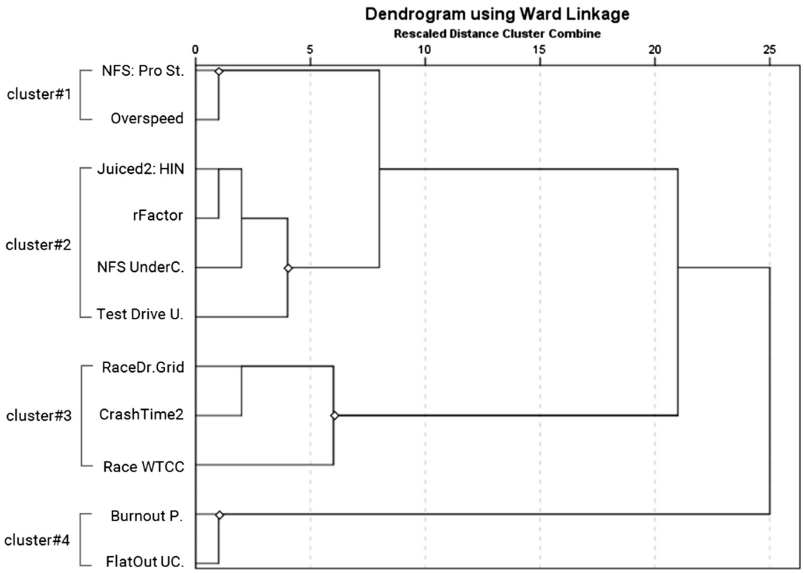


Fig. 4. Hierarchical Cluster Analysis using Ward's methods to combine the DEA-RG scores for all the three categories of realism. A four-cluster solution was able to group the overall-scores at minimal distance for the eleven video games. The emerging clusters describe a more complex classification of racing video games compared to "*drive'em up*" and "*circuit racing*".

The first cluster grouped together two games (*Need for Speed Pro Street*; *Overspeed*) that presented a driving experience characterized by medium levels of "driving realism" and low scores of "environmental realism" and "accident realism".

The second cluster grouped together a series of "*drive'em up*" and "*circuit racing*" games that presented medium levels of both "driving realism" and "environmental realism", but with very low scores of "accident realism" (*Juiced 2*; *rFactor*; *Need For Speed Undercover*; *Test Drive Unlimited*).

The third cluster included racing video games that present overall scores similar to the second cluster, but with a better description of the consequence of a "road accident" (*Race Official WTCC Game*; *RaceDriver GRID; CrashTime2*).

Finally, the last cluster grouped together classic arcade-style games, with very high scores of driving environment articulation and "road accidents" representation, but with very poor scores of "driving realism" (*FlatOut UC.*; *BurnoutParadise*).

4 Discussion

The aim of the present research was to propose an analysis tool (DEA-RG) to rate the content of racing video games in terms of driving experience. Different aspects of the driving task were evaluated in racing video games, in order to understand: the complexity of the driving environment in which the user can drive; the driving task accuracy and the driving skills required to operate vehicle during gameplay; the realism and reward system of eventual road accidents. Cluster analysis was performed to group racing video games based on overall ratings of the driving content experience.

Results showed that is possible to analytically describe the content of the driving simulation generated by video games; in terms of accuracy and distance to the multi-factory driving experience complexity. The DEA-RG scores showed that the common labels used to categorized racing video games do not completely overlap with the driving content of the different video games. Within same type of category of racing video game, different driving experience can be assessed. For instance, some "*drive'em up*" games present medium scores for driving realism (e.g. *Juiced 2*) while other games present low scores in driving realism (e.g. *Burnout Paradise).*

At the same time, mixed levels of realism are present within each video game, with accurate descriptions of the driving realism but low level of environment complexity (e.g. Race WTCC). If we consider that video games are used by novice and young drivers that lack of direct driving experience, it is possible to see that these mixed levels of realisms could be potentially confounded, generating some interesting implications for road safety research. For instance, video games described in *cluster #4*, present realistic environment complexity and road accident dynamics therefore they could suggest to unexperienced adolescents the idea of presenting also accurate driving realism. With no or limited actual driving experience, it could be possible that novice driver could suppose that being a good driver in *Burnout Paradise* or *FlatOut,* could mean being a good driver also in real life, because those video games *seem* to present an overall realistic driving experience. Several research have shown that self-assessment of being a skilled driver can affect the risk-taking behaviors in critical driving situations, leading to more accidents [20]. Thus for research purposes, select racing video games with mixed realism rating, could help to promote a more focused findings on potential negative effects of racing video games.

On the other hand, some the video games that present a sufficient-enough combination of levels of realism, could potentially be used to develop general driving skills useful also in real-life driving. For instance, games in *cluster #3* present medium to high scores for the driving realism accuracy, requiring the user to engage in a complex driving task to successfully play the game (e.g. pay attention to rearview mirror or interact with advanced assistance system). These games require the driver to interact with several elements of the roadway, they present a sufficient environmental complexity, they include accurate consequences of the driver's action in case of accidents and they present complex behavioral response from other road users, too. Playing these racing games, may require the driver to master specific driving skills related to visual attention and target localization that are more close to the visual demands of real-life driving, compared to *cluster #4* video games. However, the *cluster #3* games present

still some severe limitations, like the absence of cyclists, pedestrians and motor bikes, which restrict the potential enhancement of visual abilities while driving and they require a more general consideration on the external validity before considered as potential serious game training. External validity could also be the problem with games in *cluster #2 and cluster#1*, as they present a driving environment that are over simplistic and they represent the consequence of road accidents in an inaccurate way, compared to other games or compared to real-life conditions. They do not offer a driving experience complex or realistic enough to entail the specific cognitive abilities related to the driving task, considered in its complexity.

Certainly, racing video games are not meant or designed to represent the driving task in its complexity. The present research did not took in consideration the gameplay strategies implemented to achieving user engagement. For instance, the simplified driving task or the exaggeration of the consequence of the accidents without penalties, are intrinsic part of the arcade-style or kart racing gameplay. The goal of this type of video games is entertain the player allowing edgy racing experience with adventurous drifting through turns, thrilling overspending and exaggerated collisions. To achieve those goals it is essential to remove the driving precision that could become an obstacle for the engagement purposes of the game. On the other hand "racing simulations" will require the player to focus more strictly on vehicle handling process in order to achieve game engagement.

However, the DEAR-RG can represent a preliminary tool to rate the driving content of individual video games, which could provide an opportunity to inform the users and researchers on the driving complexity created by each game. Future researches could test the effects of racing video games with different DEA-RG scores, to determine the different effects of racing video games on real driving abilities [21]. Understanding the way different video games impact user's driving ability could indeed provide some insight on the positive and negative effects of racing video games and it could help the design of serious games for training specific driving abilities using the efficient user-engagement that video games are able to generate.

References

1. White, M., Cunningham, L., Titchener, K.: Young drivers' optimism bias for accident risk and driving skill: accountability and insight experience manipulations. Accid. Anal. Prev. **43** (4), 1309–1315 (2011)
2. Scott-Parker, B., Watson, B., King, M.J.: Understanding the psychosocial factors influencing the risky behaviour of young drivers. Transp. Res. Part F Traffic Psychol. Behav. **12**(6), 470–482 (2009)
3. ESA: Essential Facts About the Computer and Video Game Industry. http://essentialfacts.theesa.com/Essential-Facts-2016.pdf. Accessed 2016
4. Fisher, P., Greitemeyer, T., Morton, T., Kastenm, A., Postmes, T., Frey, D., Kubitzki, J., Odenwälder, J.: The racing-game effect: why do video racing games increase risk-taking inclinations? Pers. Soc. Psychol. Rev. **35**, 1395–1409 (2009)

5. Prot, S., Anderson, C., Gentile, D., Brown, S., Swing, E.: The positive and negative effects of video game play. In: Media and the Well-Being of Children and Adolescents. Oxford University Press, New York, pp. 109–128 (2014)
6. Triberti, S., Villani, D., Riva, G.: Moral positioning in video games and its relation with dispositional traits: the emergence of a social dimension. Comput. Hum. Behav. **50**, 1–8 (2015)
7. Anderson, C., Dill, K.: Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. J. Pers. Soc. Psychol. **78**, 772–790 (2000)
8. Vingilis, E., Yıldırım-Yenier, Z., Fischer, P., Wiesenthal, D., Wickens, C., Mann, R., Seeley, J.: Self-concept as a risky driver: mediating the relationship between racing video games and on-road driving violations in a community-based sample. Transp. Res. Part F Traffic Psychol. Behav. **43**, 15–23 (2016)
9. Breuer, J., Bente, G.: Why so serious? On the relation of serious games and learning. Eludamos J. Comput. Game Cult. **4**(1), 7–24 (2010)
10. Cardoso-Leite, P., Bavelier, D.: Video game play, attention, and learning: how to shape the development of attention and influence learning? Curr. Opin. Neurol. **27**(2), 185–191 (2014)
11. Green, C., Bavelier, D.: Action-video-game experience alters the spatial resolution of vision. Psychol. Sci. **18**(1), 88–94 (2007)
12. Dye, M., Green, C., Bavelier, D.: The development of attention skills in action video game players. Neuropsychologia **47**, 1780–1789 (2009)
13. Achtman, R., Green, C., Bavelier, D.: Video games as a tool to train visual skills. Restor. Neurol. Neurosci. **26**(4–5), 435–446 (2008)
14. Evans, G., Marrero, D., Butler, P.: Environmental learning and cognitive mapping. Environ. Behav. **13**(1), 83–104 (1981)
15. Ciceri, M., Ruscio, D.: Does driving experience in video games count? Hazard anticipation and visual exploration of male gamers as function of driving experience. Transp. Res. Part F Traffic Psychol. Behav. **22**, 76–85 (2014)
16. Stinchcombe, A., Kadulina, Y., Lemieux, C., Aljied, R., Gagnon, S.: Driving is not a game: video game experience is associated with risk-taking behaviours in the driving simulator. Comput. Hum. Behav. **69**, 415–420 (2017)
17. Fastenmeier, W., Gstalter, H.: Driving task analysis as a tool in traffic safety research and practice. Saf. Sci. **45**(9), 952–979 (2007)
18. Groeger, J.: Understanding Driving: Applying Cognitive Psychology to a Complex Everyday Task. Psychology Press, Philadelphia (2000)
19. Charmaz, K.: Constructionism and the grounded theory method. In: Holstein, J.A., Gubrium, J.F. (eds.) Handbook of Constructionist Research, pp. 397–412. Guilford Press, New York (2008)
20. Lajunen, T., Summala, H.: Driving experience, personality, and skill and safety-motive dimensions in drivers' self-assessments. Pers. Individ. Differ. **19**(3), 307–318 (1995)
21. Ruscio, D.: Simulating real danger? Validation of driving simulator test and psychological factors in brake response time to danger. In: ICMI 2012 Proceedings of the 14th ACM International Conference on Multimodal Interaction, New York, NY, USA, pp. 345–348 (2012)

Multivariate Differences in Driver Workload: Test Track Versus On-Road Driving

Jack L. Auflick(✉)

Engineering Systems Incorporated, 1174 Oak Valley Drive, Ann Arbor, MI 48108, USA
jlauflick@engsys.com

Abstract. In a continuing effort to examine data from the Crash Avoidance Metrics Partnership (CAMP), Driver Workload Metrics Project (DWM) (The Driver Workload Metrics project, a co-operative agreement between the NHTSA, Ford, GM, Nissan, and Toyota, was conducted under the Crash Avoidance Metrics Partnership (CAMP), established by Ford and GM to undertake joint precompetitive work in advanced collision avoidance systems.) this paper identified correlates of driver workload, a construct defined as the competition in driver resources (perceptual, cognitive, or physical) between the driving task and a concurrent secondary task occurring over that task's duration. Data from 24 on-road and 24 test-track subjects who performed visual-manual, auditory vocal, a combination of both, or just drive tasks, were analyzed using Maximum Likelihood Factor Analysis (MLFA). Results found that there were seven hidden factors that still explained about 62% of the original variance from the original 42 variables. Factors scores for each subject were analyzed with Multivariate Analysis of Variance (MANOVA). Results found highly statistically significant workload differences in venue, age groups, task type, and gender. Results from radar plots visually define the subjective concept of driver workload.

Keywords: Driver Workload Metrics · Driver workload · Driver distraction · Human factors

1 Introduction

As noted in several major research publications over the past decade automotive manufacturers are introducing new technologies and automated systems into their vehicles touting how the new tools will provide enhanced safety for drivers. Examples of some of these new tools range from passive to active, including advanced interactive navigation systems, active lane departure and blind spot warnings, collision avoidance and active braking systems. These past studies tried to examine how semi- or fully autonomous vehicles may emulate driver behavior and reduce the number of traffic accidents and fatalities. While fully autonomous driving is still in its infancy, much research has been done that examines driver behavior interacting with secondary, hand-held technologies, such as cell phones, tablets, and navigation systems.

In the United States, the National Highway Transportation Safety Administration had voiced concerns that use of these new technologies and devices within the automotive

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_89

cockpit will at times overload and distract the driver. One such project intended to investigate driver workload was the Crash Avoidance Driver Workload Metrics (DWM) [1]. In its experimental design, drivers interacted with secondary tasks while driving on a four-lane divided highway or alternately while driving on an oval test track. Workload was empirically defined to be the competition in driver resources (perceptual, cognitive, or physical) between the primary driving task and the concurrent secondary task occurring over that task's duration. DWM hypothesized that, depending on the type of secondary task performed, measured workload and the correlated quality of driving would manifest in degraded measures of lane keeping, longitudinal control, object-and-event detection, or eye glance behavior.

DWM's experimental tasks, based on published literature on driver workload and Multiple Resource Theory (MRT) [2] involved visual input and manual output, auditory input and vocal output, a combination of these two, or just driving. The ensemble of tasks was like tasks done by drivers today and included seven auditory vocal and seven visual manual tasks, and the base line just-drive task plus a combination cell phone task. Driver performance data was collected from on-road (highway) and a test track. The test track venue included several additional visual manual tasks that an institutional review board believed to be too dangerous to do while driving on the highway. Details of the experimental tasks can be found in the CAMP DWM Appendices [3].

As explained in the DWM final report [1], driver performance data collected driver responses for 69 test track drivers and 108 on-road drivers using a canonical repeated measures experimental design where each test subject repeatedly performed the secondary tasks during real driving. Test subjects drove an instrumented car that was the center car of an instrumented three-vehicle platoon, providing a realistic car-following driving experience. Data was captured through an extensive array of sensors, cameras, and on-board instrumentation. For example, vehicle kinematic data was recorded during task performance for longitudinal and lateral vehicle control, task duration, speed, range, range rate, time headway, time to contact, (mean, median, standard deviation), minimum and maximum distances, and time durations (at the minimum or maximum). In addition, driver steering wheel behavior was measured with respect to how much time, during a task, that the steering wheel was held at zero degrees, with respect to the midpoint of the lower steering wheel, or when it had been rotated to the right or left and was held at a 15° or 20° offset from the center line, zero-degree position. Also included were the number of lane exceedances and durations, plus the number of center stripe touches and durations. In addition to the kinematic variables, driver eye glance patterns were also recorded for glances to the road, mirror, situation awareness (outside left or right), task related (glances to in-cockpit locations during visual manual tasks), not road, and to not assigned (other) locations. These eye glance metrics included median glance duration, standard deviations, and the percent of a task's duration attributed to that specific location.

The DWM project analyses developed performance metrics and test procedures that reliably assessed how driving performance may be negatively affected due to auditory, visual, manual, and cognitive aspects of driver workload associated with using secondary tasks and in-vehicle systems. While the results from the DWM study created an extensive array of bi-variate correlations and regression analyses, the results allowed

the researchers to infer the existence of driver workload related to task type. These correlations were empirically observed across multiple measures of performance.

2 Factor Analysis

However, beyond the DWM results, workload, the construct of interest, had no direct empirical measure within the research literature that was available at the time. While the individual DWM analyses provided brief glimpses into the complex, hidden inter-relationships within the collected data, they fell short of fully defining workload mainly because workload was believed to be a multidimensional, transient mental state that changed from task to task. The DWM researchers believed that workload, if it existed, could be observed through the analysis of the hidden, multivariate correlations by using exploratory factors analysis (EFA). When the original DWM project ran out of time and funding, one member of the research team continued to experimentally define workload through the use of Maximum Likelihood Factor Analysis (MLFA). The purpose of these analyses was to extract the underlying latent, highly inter-correlated structure within these data, believed to be workload, caused by the variation in driving performance during multitasking. MLFA [4, 5] is one of several commonly used EFA techniques that in general seek to: (1) reduce the number of variables and (2) to detect unobserved structure (latent variables) in the relationships between variables. MLFA approaches this with general methods that estimate the loadings and communalities in a data set and then maximizes the likelihood function to determine the parameters that are most likely to produce the observed data.

2.1 Hidden Structure – On Road Data

The first of these two past studies used MLFA to analyze the data from test subjects who drove on the highway [6]. This analysis used 42 vehicle kinematic variables (like speed, range, range rate, longitudinal and lateral lane positions, etc.), time variables (like time for the vehicle at minimum or maximum lane position right and left, plus minimum and maximum times for the kinematic variables), and the steering wheel position and durations. Additionally, the analysis included the data from 24 test subjects with complete eye glance data. These data included eye glance variables (such as median, standard deviation, and percent task duration, etc. from on-road, mirror, situation awareness, task-related, and not-road or not-assigned locations). The sample size had 12 males and 12 females and was also equally parsed into older, middle, or younger groups.

The first step of MFLA identifies how many factors should be retained in the analysis. In factor analysis, a factor is a latent (unmeasured) variable that expresses itself through its relationship with other measured, observable variables. In this analysis, MFLA was used in an iterative fashion starting with all 42 independent variables, meaning that there could have been up to 42 orthogonal or independent factors. Then, using the Kaiser criterion [7] and Cattell's scree plots [8], results showed that from the original 42 variables, seven new orthogonal factors described the hidden structure within the data while still explaining 66% of the original variance.

MLFA not only generates individual orthogonal factors, but provides factor loadings (i.e., the correlations of an observed variable with the underlying factor conceptually like Pearson's *r)* from those independent variables that contributed to that factor. Based on their individual directionality and magnitude, factor loadings provide important insights when interpreting the MFLA solutions. When the loadings are squared, the resulting values describe the percentage of variance within a specific independent variable explained by a given factor. MLFA also generates individual factors and related factor scores for each test subject on each variable and task. These scores were also normalized, and rotated using the varimax orthogonal transformation [8–10] creating a matrix of scores representing a person's relative spacing or standing on each of the seven latent factors. Factor scores were subsequently analyzed using Multivariate Analysis of Variance. Results from the MANOVA found there were highly statistically significant main effects for task type, gender and, age, with highly significant interactions between task type and the age group as well as between gender and age group. Overall results from this on-road data analysis clearly showed that MFLA and subsequent MANOVA analyses could define and portray the hidden structure related to driver workload.

2.2 Hidden Structure – Test Track Data

The second of the two studies also used the same MLFA process to analyze the data from test subjects who drove on the test track [11]. This analysis used the same set of 42 independent variables but also included the set of 8 new visual manual experimental tasks that were not allowed to be done on the highway. As above, there were 24 test subjects with complete eye glance data, grouped by gender, 12 males and 12 females, and by age group: older, middle, or younger groups. It was hypothesized that the new tasks would demonstrate higher levels of visual-manual workload when compared to less demanding tasks.

As in the prior on road DWM study, test subject responses from the kinematic and eye glance behavior from the test track data were first analyzed using MLFA to uncover the underlying unobserved structure within the large set of variables. It is this hidden multi-dimensional structure that must be examined to empirically identify the concept of driver workload. As in the DWM on-road analyses, these analyses found that task-induced workload affected driving performance and was multi-dimensional in nature. Visual-manual tasks exhibited fundamentally different performance profiles than auditory-vocal tasks or just driving. After rotating and naming factors, results also found seven factors that explained approximately 63% of the original variance in the data These factors had different loadings and interpretations that appeared to come from the ambient environment while driving on an automotive test track. Furthermore, when secondary statistical analyses of the normalized factor scores were done using MANOVA, the results found highly statistically significant workload differences in age groups and task type. Interestingly, there were no gender differences. There were also workload differences between the more egregious visual manual tasks and the less hazardous ones. Results from this test track data analysis also clearly showed that MFLA and subsequent MANOVA analyses could define and visually depict the hidden structure related to driver workload.

3 Driver Workload: Test Track Versus On-Road Driving

As can be seen in the discussion above, the two preliminary exploratory analyses provide strong results indicating that the theoretical concept of driver workload can be empirically seen within a complex set of kinematic, visual, and behavioral test subject variables. However, there were still questions about the overall sensitivity of the method, e.g. using MFLA, followed by MANOVA of subjects' factor scores. So a third study was done where the data from the test track and the on-road driving were combined into one data set. This analysis continued to use the same 42 independent variables discussed above, but did not use the set of more hazardous visual manual experimental tasks from the test track. The resulting data set had 48 subjects. There were 24 males and 24 females with 16 subjects in each of the three age groups.

There could have been 42 orthogonal or independent factors from the MLFA, but this analysis identified that there were only seven factors that defined the hidden relationships believed to be workload. These seven factors had Eigenvalues greater than 1.0, using both the Kaiser criterion and Cattell's scree plots. Subjects' factor scores and the factors were all normalized and rotated using the Varimax procedure. A varimax rotation makes the output more understandable by maintaining the orthogonality of the factor axes but maximizing the variance of the squared loadings in the factor matrix (i.e. correlation coefficients between the cases [rows] and factors [columns]). These rotated factors identify the simple structure or hidden relationships within the unobserved configuration in the data, if such structure exists.

Based on their individual directionality and magnitude, factor loadings were again interpreted through a highly subjective activity that generated "names". While there are several approaches on how to do this, a generic rule of thumb has been developed suggesting that interpretations should be done only on those loadings exceeding |0.7|. The rationale for this is that when a loading of 0.7 or higher is squared, about half of the variance in that variable is being explained in the factor. Table 1 contains the names applied to the seven factors that were derived through analysis of loadings on each of the seven factors. Test subjects' factor scores on each factor were analyzed using (MANOVA). The MANOVA tested hypotheses that there were statistically significant differences between venues, gender, age groups and task type (Visual-Manual, Auditory-Vocal, Just Drive, Combination).

Table 2 presents the MANOVA analyses and statistical significance. As can be seen in Table 2, there were highly statistically significant main effects for Task Type, Venue, Gender, as well as for Age Group. In addition, there were highly significant two-way interactions for Task Type by Venue, Task Type by Age Group, and Gender by Age Group. Finally, the results had one significant three-way interaction for Venue by Gender by Age Group.

Conceptually it is very hard to envision a multi-dimensional space where there are seven orthogonal axes. The prior two analyses discussed earlier have shown that one way to effectively do this is to use radar plots where test subjects' factor scores on each factor were averaged by Task Type, Venue, Gender, or Age Group. Means are then plotted for each factor on the radar plots that present a two-dimensional cross- section of a multi-dimensional driving workload profile for groups of subjects.

Table 1. Eigenvalues, named factors, and explained variance.

Factor	Eigenvalue	% Total variance	Cumulative eigenvalue	Cumulative %
Eyes up with kinematic variability	11.8	28.0	11.8	28.0
Road glances/Not SA or MR	5.0	11.8	16.7	39.8
Task glances	2.8	6.8	19.5	46.5
Cross centerline & duration	2.2	5.2	21.7	51.8
Touch centerline & duration	1.2	2.9	23.0	54.7
Glances to MR, SA, Avg. range rate	1.2	2.7	24.1	57.4
Glances to MR & Range, Rate, Speed Variability	1.9	4.4	26.0	61.9

Table 2. MANOVA tests of significance (Wilks Lambda)

	Value	F	Effect	Error	p
Intercept	**0.745**	**56.854**	**7**	**1164.0**	**0.000000**
{1}TaskType	**0.076**	**231.279**	**21**	**3342.9**	**0.000000**
{2}VenueName	**0.947**	**9.231**	**7**	**1164.0**	**0.000000**
{3}GenderName	**0.980**	**3.344**	**7**	**1164.0**	**0.001557**
{4}AgeGroup	**0.914**	**7.612**	**14**	**2328.0**	**0.000000**
TaskType * VenueName	**0.921**	**4.643**	**21**	**3342.9**	**0.000000**
TaskType * GenderName	0.989	0.642	21	3342.9	0.890611
VenueName * GenderName	0.994	0.920	7	1164.0	0.489946
TaskType * AgeGroup	**0.935**	**1.881**	**42**	**5463.1**	**0.000513**
VenueName * AgeGroup	0.980	1.687	14	2328.0	0.051781
GenderName * AgeGroup	**0.979**	**1.808**	**14**	**2328.0**	**0.032252**
TaskType * VenueName * GenderName	0.982	1.006	21	3342.9	0.452026
TaskType * VenueName * AgeGroup	0.973	0.767	42	5463.1	0.861418
TaskType * GenderName * AgeGroup	0.968	0.904	42	5463.1	0.648634
VenueName * GenderName * AgeGroup	**0.955**	**3.836**	**14**	**2328.0**	**0.000002**
1 * 2 * 3 * 4	0.968	0.912	42	5463.1	0.633452

The purpose of the MFLA was to identify hidden structure that was indicative of driver workload. This MLFA found that these data contained extensive hidden relationships, i.e. the latent structure that can be defined as driver workload. In the present case, the multidimensional workload could be described by seven factors while explaining ~62% of the original variation in the data. As postulated by the original DAM research team, the seven underlying factors reveal key effects of multitasking on driver workload caused by interference from secondary tasks. The secondary MANOVA affirmed the qualitative differences in driving workload as defined by the dependent grouping variables. While there are very strong main effects for these groups, the data

also showed statistically significant interactions. Figures 1, 2, 3 and 4 below show examples of the radar plots from the four main effects.

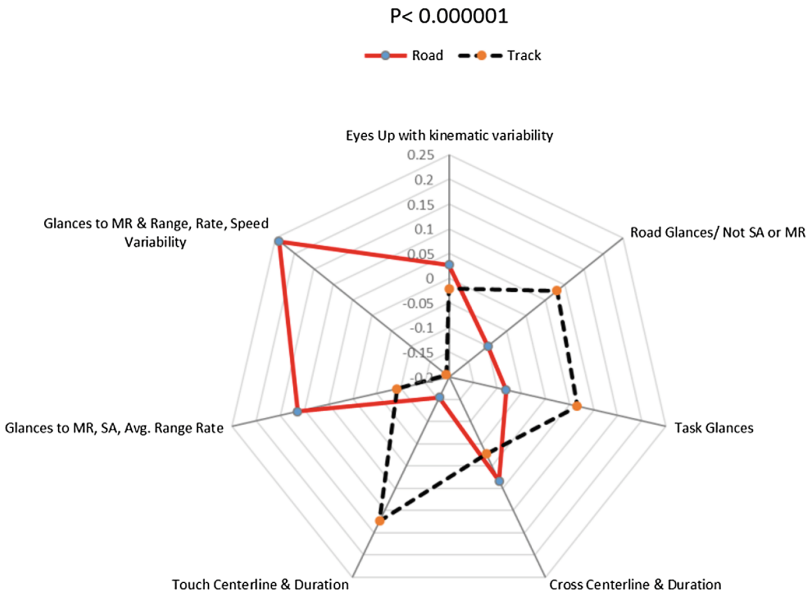


Fig. 1. Radar plot for venue

Figure 1 depicts difference in driver workload related to the venue. One would expect that driving on a very active four-lane highway should have a different workload profile compared to a tightly structured driving environment on a test track. This is what was found in the MFLA and MANOVA.

Of more interest is the radar plot in Fig. 2 showing the highly significant statistical differences based on Gender. When data from both venues is combined in one MFLA, there are clear, significant differences in driver workload due to Gender. More research will be required to further understand how and why these Gender differences occur while driving.

Figure 3 is a radar plot based on the Age Group main effect that was again highly statistically significant. Numerous past studies and insurance data have stated that driver abilities vary with age. This study confirms that there are age-related differences, both qualitative and quantitative, with respect to driver workload.

Finally, Fig. 4 shows the highly significant differences in workload based on the type of task the test subjects performed. This was one of the original DWM conclusions that was based on inferences drawn from numerous bi-variate correlation and regression analyses. The present MFLA confirms this finding and can qualitatively and quantitatively show that driver workload does vary in relation to the type of task performed while driving.

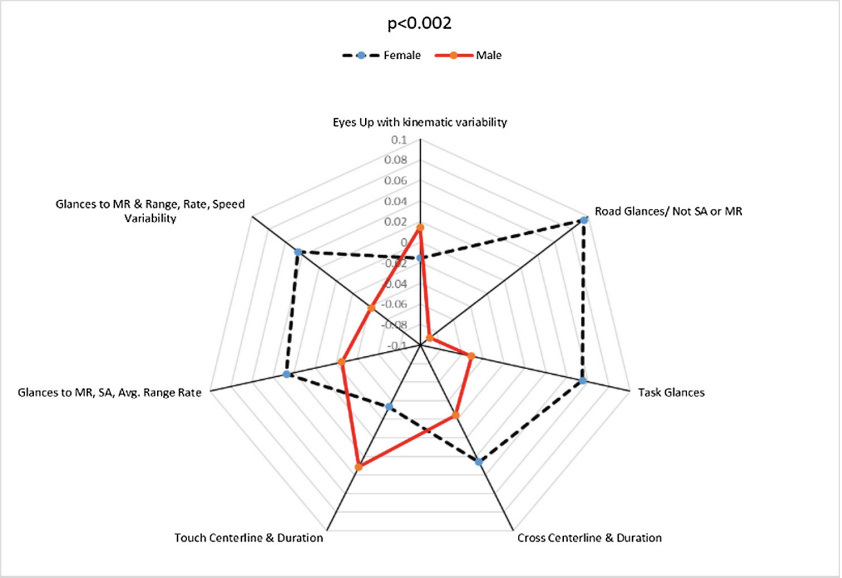


Fig. 2. Workload differences due to gender

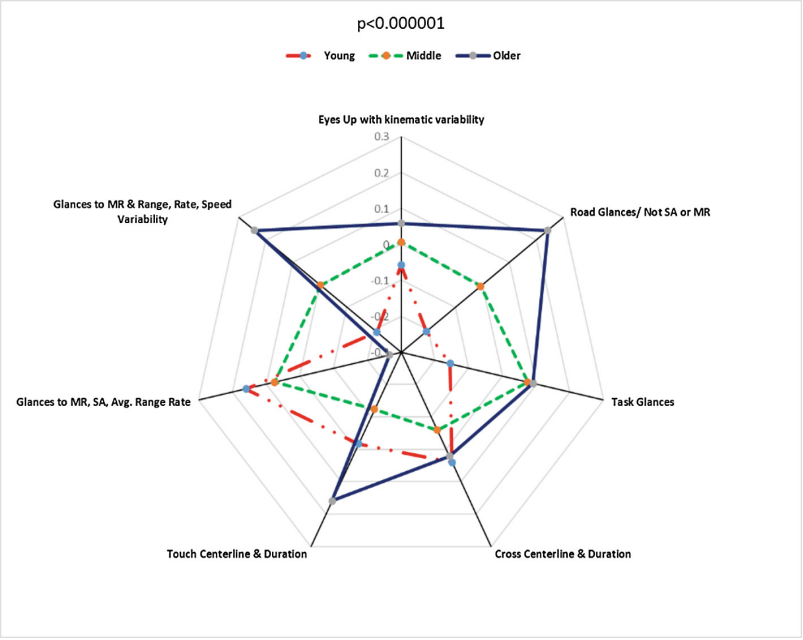


Fig. 3. Workload differences by age group

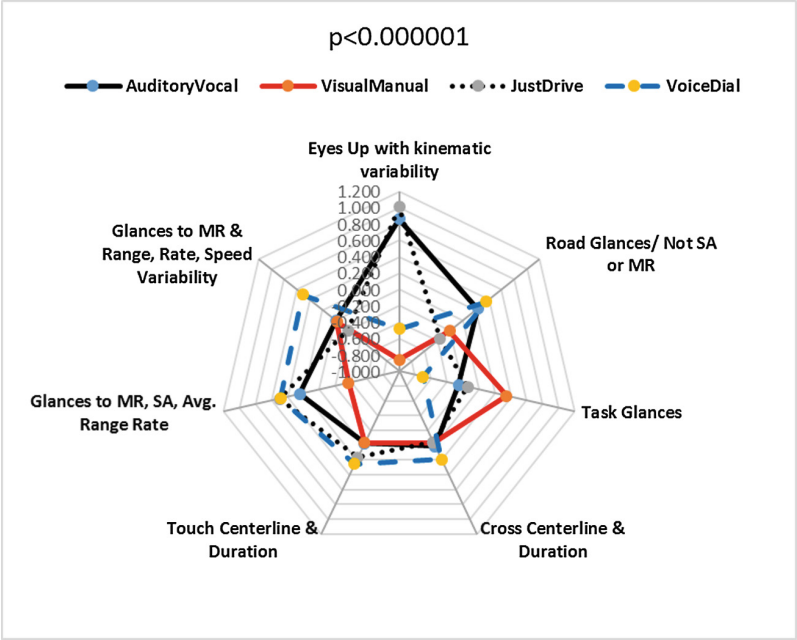


Fig. 4. Differences in workload due to task type

4 Conclusions

This new MLFA, combining on-road and test track data, found that there was extensive multicollinearity, e.g. hidden relationships tied to driver workload. As in the previous analyses of DWM data, the latent structure which defined workload, could be described by seven factors, while still explaining 62% of the original variation in the data. When interpreted by names, these seven factors revealed key effects of multitasking on driver workload caused by interference from secondary tasks. The actual driving task while performing secondary tasks, was shown to be multidimensional in nature, meaning that it was represented in the data by simultaneous effects on multiple factors. Driver workload appears to be affected by different allocations of driver resources across input modalities as well as being affected by task type, gender, venues, or age group. However, it must be noted when applying MFLA, there was an explicit recognition that it was exploratory in nature where the underlying dimensions would need to be attributed with meaning and interpreted through subjective analysis. As exploratory work, the nature of the dimensions could change if the input to the analysis were different.

Acknowledgements. Much praise goes to the Driver Workload Metrics project team: L. Angell, P.A. Austria, D. Kochhar, L. Tijerina, W. Biever, T. Diptiman, J. Hogsett, and S. Kiger. They endured much but made a significant contribution to the understanding of driver workload.

Thanks go to Manuel Meza-Arroyo, Ph. D. who worked to help generate meaningful factor names based on the data.

References

1. Angell, L., Auflick, J.L., Hogsett, J., Austria, P.A., Kochhar, D., Biever, W., Diptiman, T., Tijerina, L., Kiger, S.: Driver Workload Metrics Task 2 Final Report, National Highway Traffic Safety Administration, Report Number DOT HS 810 635 (2006)
2. Wickens, C.D.: Processing resources in attention. In: Parasuraman, R., Davies, D.R. (eds.) Varieties of Attention, pp. 63–102. Academic Press, New York (1984)
3. Angell, L., Auflick, J.L., Hogsett, J., Austria, P.A., Kochhar, D., Biever, W., Diptiman, T., Kiger, S.: Driver Workload Metrics Project Task 2 Final Report Appendices (2006)
4. Thurstone, L.L.: The vectors of the mind. Psychol. Rev. **41**, 1–32 (1934)
5. Lawley, D.N.: The estimation of factor loadings by the method of maximum likelihood. Proc. R. Soc. Edinb. (A) **60**, 64–82 (1940)
6. Auflick, J.L.: Resurrecting driver workload metrics: a multivariate approach. In: Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, Procedia Manufacturing, vol. 3, pp. 3160–3167 (2015)
7. Kaiser, H.F.: The application of electronic computers to factor analysis. Educ. Psychol. Meas. **20**, 141–151 (1960)
8. Cattell, R.B.: The scree test for the number of factors. Multivar. Behav. Res. **1**, 245–276 (1966)
9. Kaiser, J.F.: The varimax criterion for analytical rotation in factor analysis. Psychometrika **23**(3), 187–200 (1958)
10. Crawford, C.B., Ferguson, G.A.: A general rotation criterion and its use in orthogonal rotation. Psychometrika **35**, 321–332 (1970)
11. Auflick, J.L.: Resurrecting driver workload, multivariate analysis of test track data. In: Proceedings of the 7th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences, Procedia Manufacturing (2016)

Road and Rail: Eco-Driving

Range Makes All the Difference? Weighing up Range, Charging Time and Fast-Charging Network Density as Key Drivers for the Acceptance of Battery Electric Vehicles

Julian Halbey, Ralf Philipsen(✉), Teresa Schmidt, and Martina Ziefle

Human-Computer Interaction Center (HCIC), RWTH Aachen University, Campus-Boulevard 57, 52074 Aachen, Germany
{halbey, philipsen, schmidt, ziefle}@comm.rwth-aachen.de

Abstract. Electric vehicles are currently seen as the most promising approach for the mobility sector to do it's bit for the reduction of greenhouse gas emissions. Although, there are still impediments to adoption, which may partly be overcome by a fast charging infrastructure that enables long-haul trips with battery electric vehicles (BEV). The presented work aims at weighing up user specific parameters of a charging infrastructure (range, charging time, network density, location) to understand public acceptance and reflect advice for planning a needs-based fast-charging network. Using an online questionnaire, 176 participants (users and non-users of BEVs) were addressed in a choice-based conjoint approach. Charging time emerged to be the most influential factor, followed by driving range and charging locations. Significant differences were found between BEV-users and non-users, with BEV-users putting more emphasis on the charging locations and less on the range.

Keywords: Electro mobility · Charging network · Infrastructure planning · BEV experience · Choice-based conjoint analysis

1 Introduction

The climate change is one of the major challenges humanity has to face today. The emission of greenhouse gases, in particular CO_2, is the most important human contribution to global warming. To stabilize the average global temperature, 197 nations acceded to the so-called Paris Agreement under the auspices of the United Nations in 2015 [1] and committed themselves to reductions of greenhouse gas emissions. The focus lies on the combustion of fossil fuels as key part of man-made CO_2 emissions. Apart from climate protection, the reduction of combusting fossil fuels can help modern societies to become independent from exhaustible resources. Alongside industrial production, motorized traffic, mainly based on internal combustion engines, is one important technological context to focus on. To both achieve the aforementioned goals and reduce particulate matter that becomes more and more problematic in metropolitan areas, the German Federal Government already started a national program to foster the development, research and dissemination of electric vehicles in 2009 [2].

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_90

In the meantime, this program has been supplemented by further measures to achieve the aimed goal of 1 million registered electric vehicles in Germany in 2020.

However, despite government incentives for electric vehicle users, e.g., tax breaks, buyer's premiums or traffic privileges, the German Federal Motor Transport Authority reported only a little bit more than 25.000 BEVs and 130.000 Plugin-hybrid electric vehicles (PHEV) for January 2016 [3]. Therefore, there must be major barriers preventing a widespread replacement of conventional cars with internal combustion engines by electric vehicles.

E-vehicle buyers in Germany differ from buyers of conventional cars especially by a higher average income [4]. The higher acquisition costs of BEVs are an important impediment to adoption [5]. This is partly counterbalanced by state buyer's premiums, but continuing low registration figures indicate more than only financial aspects as impedimental factors for e-vehicle acceptance. In particular, the limited range, the long charging duration of BEVs in comparison to combustion engines as well as the immature status of the charging infrastructure are the main barriers not to use BEVs. [6–9].

Upcoming fast-charging technologies that can currently use 150 A and will be able to utilize 300 or more Ampere charging current in near future, can address some of the challenges by reducing charging times significantly. But, the charging infrastructure has to be perceived as sufficient and needs-based to relieve existing impediments on potential buyer side. In this context, there are many unanswered questions regarding the acceptance-relevant factors, as preferred charging times, locations, network density and design and arrangement of charging stations.

2 Related Work and Questions Addressed

Electromobility is increasingly gaining attention in science. Research broached technical [10], economical [11], and environmental [12] issues. Another focus was on psychological reasons addressing the so-called range anxiety [13], or range comfort zones [14] to explain how e-vehicle users cope with the limited range of their vehicles.

Regarding the planning of charging infrastructure, current approaches are usually based on activity-based [15, 16] or discrete-choice models [17] to simulate the optimal charging grid. All approaches are quite coarse by only using basic demographics and movement profiles. To optimize those models, the data basis must be broadened by integrating further user factors and preferences. Therefore, two subject areas are relevant: First, the buying and usage intention of customers have to be taken into account. Next to the known barriers to adoption, further user factors influencing the usage intention have been identified, as the residential situation, the educational level in the household [18], or individuals' environmental awareness [19]. Furthermore, prior experience with electric vehicles was identified as acceptance-relevant [20–22]. BEV experience influences both the charging respective driver behavior as well as the general usage intention. This first thematic field helps to gain a better understanding about who will buy or use BEVs as successors of the current early adopter group to enable predictions about e-vehicles' future geographic distribution.

The second subject area deals with user preferences on charging stations' locations, designs and arrangements. In particular, motorway service stations were identified as

preferred location for fast-charging in long-haul scenarios, whereas the location for short-haul scenarios depends on potential parallel activities to spend time during charging. Supermarkets and shopping opportunities seemed to best fit users' needs [23]. Again, BEV experience influences the evaluation of charging infrastructure, and, as a consequence, the willingness to use BEVs. The availability of possible parallel activities was more important for people without BEV experience than for BEV-users.

Both subject areas must be factored in during infrastructure planning to establish a charging network that is both needs-based and accepted by existing and potential future electric vehicle users. Still, most of the prevailing research is adopted to current charging times and ranges, namely a range of about 100 km that can be utilized by charging in 30 min or less, as currently realized by BEVs like BMW i3 or Nissan Leaf. With an increase of possible ranges, charging currents and reduced charging times, users' preferences regarding charging infrastructure and potential trade-offs between charging net characteristics are likely to change. Therefore, understanding users' final decisions to use or not to use BEVs requires to empirically explore the weighing up of parameters of charging infrastructure, e.g., range, charging time and network density.

3 Methodology

To address these questions, a questionnaire-based conjoint study was conducted. In the following sections the study design of the conjoint analysis (CA), its development and final design, details on data acquisition and analysis as well as the sample description will be presented.

3.1 Questionnaire Design and Conjoint Analysis

Methodologically, a conjoint measurement approach was applied to explore users' attitudes towards BEVs. Conjoint analysis (CA) enables a holistic examination of decision scenarios, in which (product) attributes are weighed against each other – allowing to investigate, evaluate and understand how future products, in this case BEVs, are perceived by customers. CA opens the possibility to include a statistical estimation algorithm in a measurement model [24]. Through the assessment of beforehand identified relevant attributes and attribute levels, human decisions can be simulated. Also, preferences can be decomposed into separate part-worth utilities of the attributes and their levels [25] to reflect the nature of the decision process. For this work, the choice-based-conjoint (CBC) analysis approach was chosen, in which several attributes influence the final decision for a concrete scenario in a choice-based task. Based on multinomial logit, it is possible to develop a model about probability of choice.

The survey consisted of two main sections: The first section introduced the topic and questioned demographics. Here, age, gender, the mobility profile (e.g. mileage per year/day, travel behavior), the attitude towards novel technology [26] and the previous experience with e-vehicles was questioned. The second section introduced the CBC with four different attributes relating charging an e-vehicle:

- *Charging location.* The attribute indicates the attractiveness (and position) of the charging station. The following levels were used: Location (1) on a parking lot (only providing a restroom), (2) on a service area (equipped with a restaurant and playground), (3) on a parking lot off the motorway (necessary detour for charging) or (4) inner city (enabling a shopping tour and benefits of a city).
- *Charging time.* The attribute describes the duration for re-charging the vehicle and contains the levels 15 min, 30 min, 60 min and 90 min.
- *Network density.* The attribute indicates the distance between charging stations. It contains the levels 50 km, 100 km, 150 km and 200 km.
- *Range.* The introduced possible distances, which can be covered with a battery charge and which is recovered by fast charging, were 200 km, 300 km, 500 km and 800 km.

Participants had to select the preferred scenario out of four randomly composed alternatives based on predefined attributes and levels. This selection task was repeated ten times for each participant. The final research design looks as follows (see Fig. 1).

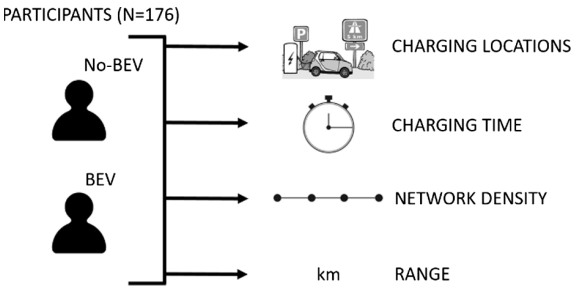


Fig. 1. Research design with focused user groups and identified attributes

3.2 Data Acquisition and Analysis

The study was conceptualized as web-based questionnaire. Participants were acquired in the authors' social environment and in specialist forums on the topics of motorized road traffic and electromobility. Subsequently, data was analyzed using Hierarchical-Bayes (HB) estimation to calculate the relative importance and part-worth utilities of the four attributes and the corresponding levels. The part-worth utilities represent the attractiveness of an attribute level relative to the other levels of the same attribute. The higher the utility score, the more likely the attribute level was chosen by participants. The relative importance of an attribute is calculated by the span of the part-worth utilities for that attribute divided by the spans of part-worth utilities of all attributes.

3.3 Sample

In total, 176 persons 68.8% men (n = 121) and 31.3% women (n = 55) participated, in an age range of 16 to 76 years (M = 34.82; SD = 12.93). The sample reported a rather high interest in new technical developments (M = 3.56 out of 5 points max.). On the base of the reported experience, participants were divided into two experience groups: BEV-users (24–65 years, n = 43), who drive an e-vehicle regularly or own an e-vehicle and non-users (16–76 years, n = 129), who never have driven an e-vehicle before.

4 Results

The results section will be structured as follows: First, the relative importance of charging network attributes will be presented. Second, results from choice simulations will be introduced. Last, the effects of BEV experience will be highlighted.

4.1 Relative Importance of Charging Network Attributes

The relative importances show how much the examined attributes contribute to the decision for a scenario. With a relative importance of 36,74%, charging time was the most decisive attribute for the choice of scenarios (see Fig. 2).

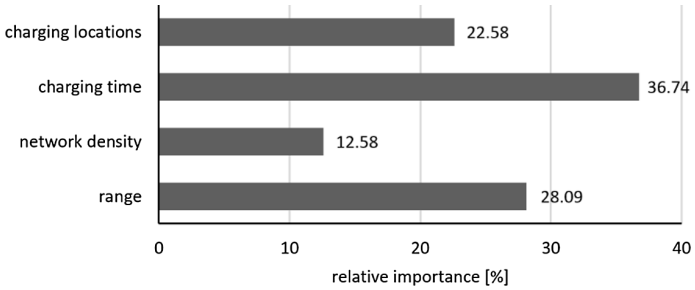


Fig. 2. Relative importances of e-vehicle and charging infrastructure attributes

Thus, charging time plays the most important role when people decide under which of the given conditions they would use BEVs for long-haul trips. The second most important attribute is the available range (per one battery charge), followed by the attractiveness of charging locations. The charging network density turned out least important.

More detailed insights into user's choice behavior and their preferences concerning a fast charging infrastructure for long-haul trips can be taken from part-worth utilities. It is important to note that negative part-worth utilities do not reflect rejection and positive scores do not mean approval of an attribute level. Rather, the differences between levels are meaningful.

As shown in Fig. 3, the service area is the most attractive *charging location* with a part-worth utility of 32.51. The two options that accompany leaving the motorway for charging clearly reached the least approval. Compared to the service area option, which was favored, the high differences in the utility scores of 49.92 for the inner-city option and 52.69 for the parking off the motorway symbolize quite low evaluations. In contrast, the simple parking with restroom – even though still less preferred than the service area – receives a higher acceptance than the options in which the motorway has to be left for charging.

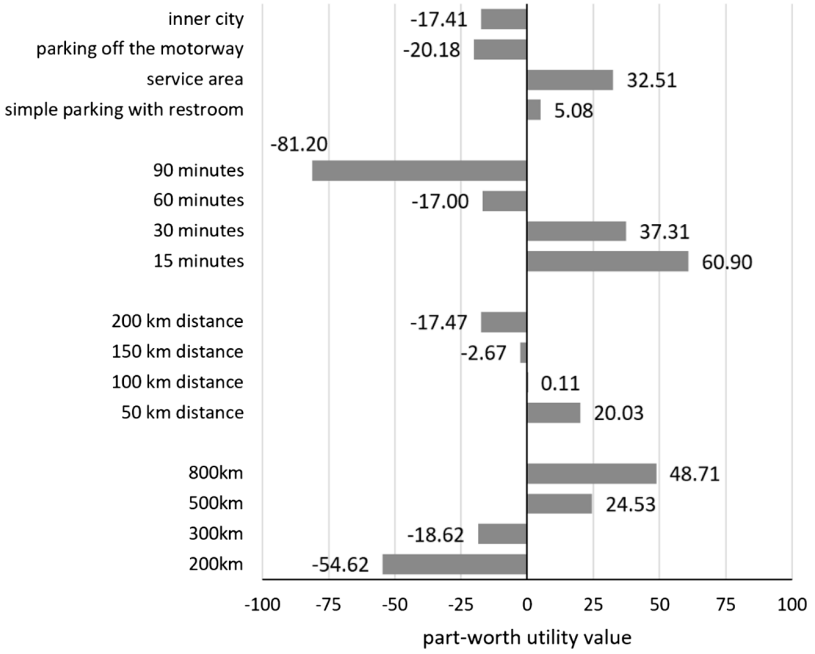


Fig. 3. Part-worth utility values (zero-centered diffs) for e-vehicle and charging infrastructure attributes

Concerning *charging time*, there are also two levels that seem acceptable, which are 15 min with a part-worth utility of 60.90 and 30 min with 37.31. The longer options fall back in approval.

With respect to the *charging network density* the medium distances of 100 and 150 km between the charging stations are rated almost equally around 0, whereas there are considerable differences in acceptance between the preferred distance of 50 km to the next charging station and the 200 km distance.

For *driving range*, a clear order of preferences can be seen: 800 km range is rated best, followed by 500 km, 300 km and last 200 km. The biggest step from one level to the following is between 500 and 300 km and the smallest from 500 to 800 km.

4.2 Choice Simulations

Choice simulations were used for a closer look at interactions between the attributes and to find out how the evaluation of scenarios develops with changing certain attributes. These simulations calculate a share of preference for each scenario from a given set of scenarios. As the current public debate is how far the distances between charging stations should be, three different scenarios were created with different levels of network density and charging location, keeping the other two attributes constant (Table 1).

Table 1. Shares of preferences for different choice scenarios

Scenario	Range	Network density	Charging time	Charging location	Share of preference	Std. error
1	200 km	50 km distance	30 min	Simple parking with restroom	36.16%	1.56
2	200 km	150 km distance	30 min	Service area	42.34%	1.72
3	200 km	100 km distance	15 min	Parking off the motorway	21.50%	1.51

The here defined scenarios number 1 and 2 reached rather similar shares of preference, indicating that participants would almost likewise choose scenario 1 or 2. Although the service area turned out to be the most attractive charging station, a higher network density can almost outweigh the relative unattractiveness of the simple parking with restroom. Or, on the opposite, if the charging locations are very attractive, users accept a much coarser meshed charging network with 150 km distance between the charging stations instead of the otherwise most preferred 50 km distance. Despite the shorter charging time, scenario 3 is least preferred, which shows the unattractiveness of leaving the motorway.

In a second simulation, the interaction between charging time and driving range was analyzed (see Fig. 4).

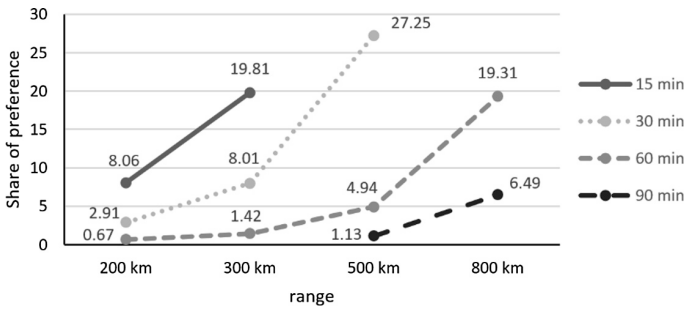


Fig. 4. Share of preferences for different e-vehicle ranges and charging times

It can be seen from Fig. 4 that, given the circumstance the most preferred scenario with 500 km range and 30 min charging time cannot be realized, the combination of 15 min charging time with 300 km range is almost equally as preferred as the combination of 800 km range with 60 min charging time. Or, if this is still not possible, 15 min with 200 km range and 30 min with 300 km range reach very similar shares of preference.

4.3 BEV Experience

Depending on participants' prior experience with BEVs, several differences between user groups were found. People who regularly use or own a BEV rated the driving range less important (M = 24.32, SE = 2.02) than people with no experience driving BEVs (M = 29.13, SE = 1.15), which was a significant difference (p = .050). Also, charging time turned out less important to the BEV experts (M = 35.68, SE = 1.40) than to non-users (M = 37.10, SE = 1.39), even though there was only a slight and not significant difference (p = .541, n.s.). The network density is slightly more important to the users (M = 13.28, SE = 1.11) than to non-users (M = 12.67, SE = .61), though this difference was not significant, too (p = .607, n.s.). In contrast, a significant (p = .000) and rather strong difference was found for charging locations: Users attach more importance to charging locations (M = 26.73, SE = 1.48) than non-users (M = 21.10, SE = .70). This means that the importances of network density and especially charging locations increase for the BEV-users, whereas range and charging time become less important, with range even falling back behind charging locations (see Fig. 5).

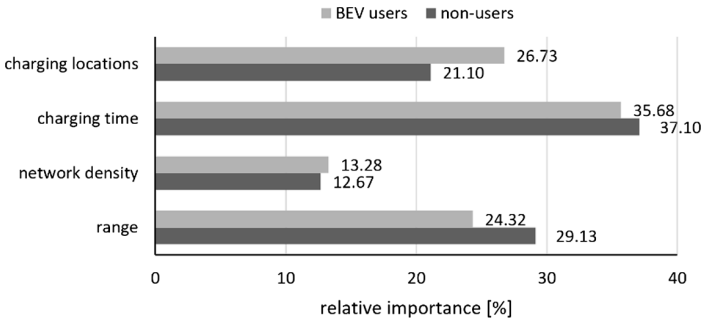


Fig. 5. Average importance of e-vehicle and charging infrastructure attributes distinguished by BEV experience groups

Even though the difference in relative importances for the network density is considerably low, there are still differences in the part-worth utilities between users and non-users of BEVs. In both groups, the charging network with 50 km distance between the charging stations reached a part-worth utility of about 20 (19.83 for users, 19.64 for non-users). For the 100 km distance, part-worth utilities reached −3.42 for the group of non-users and 4.22 for the BEV user group. For the 150 km distances, part-worth

utilities of −11.76 (BEV- users) and −0.89 (non-users) were found. Apparently, BEV-users distinguish stronger between 100 km and 150 km distance between the charging stations than the non-users, for which almost no difference could be identified.

Concerning the charging locations, users rated the service area even better, increasing the difference in approval to the second most preferred simple parking with restroom. Whereas the parking off the motorway is less preferred by the users, the inner-city option reached a slightly higher part-worth utility.

5 Discussion and Conclusion

The planning of a needs-based and user-centered fast charging infrastructure is a complex effort with many different aspects to consider. With weighing up some of these relevant attributes in close-to-reality decision scenarios, this research approach aimed at identifying the central acceptance factors that most strongly influence people in their decision whether to use an electric vehicle for long-haul trips or not and at deriving insights for the real infrastructure planning process.

As one of the key results of this study, the empirical data reveal that the range alone does not make all the difference. This is noteworthy as the debate about the limited range of BEVs dominates the public discussion. For long-haul trips, which were in the focus of this study, the charging time turned out the most influential factor. This finding might reflect charging behaviors of drivers that had been formed over a considerably long time: as refueling in conventional vehicles only takes a few minutes, people are not willing to tolerate longer charging times in BEVs. However, this finding is especially prominent in persons without experience with electric cars. BEV-users, in contrast, rate charging time, and especially range, as less important than people who have never driven a BEV. So, with more experience in driving electric vehicles the "classic" factors might lose importance, whereas charging locations and charging network density become more decisive.

Besides from that, several trade-offs in the acceptance decisions became obvious: Concerning charging time, there is a huge gap in approval between 30 and 60 min and a much smaller one between 15 and 30 min. This is a helpful insight for infrastructure planners, as it means that charging times longer than 30 min are very unlikely to be accepted for long-haul trips and the development of batteries and charging stations should aim at reaching these durations. Taking the current developments in fast charging like 300 A technology into consideration, this aim seems not too far away.

The charging locations are worth a closer look, too: The service area is clearly the most attractive location from the given options. In contrast, charging off the motorway, no matter the facilities of simple parking or inner city, were rated very badly. This result supports other studies where necessary detours for charging were seen as a major barrier [27]. The simple parking at the motorway can be seen as a backup option that still reaches some approval, but is especially suitable with shorter charging times. Therefore, time (lost through a detour or recharging) seems to be an essential factor, which is not surprising given the fact that most of the participants have only experience with refueling durations of conventional cars. Even by taking these results into account, a recommendation for an optimal configuration and equipment of charging locations

cannot be made at this point. There are other studies that focus more on this aspect [27]. However, necessary detours seem not acceptable, which is supported by the fact that these options even decrease in attractiveness with BEV-experienced persons. Here, an interesting insight is that although charging time is the most important attribute, leaving the currently driven motorway is so unattractive that even halving the charging time still does not make it comparable to other options that include charging at the motorway (as choice simulations showed).

Besides that, BEV-users focus more on the charging network density, too. Although the difference in relative importances is rather low, the gradations between the single levels provide useful insights: As BEV-users distinguish much more between the levels, especially between 100 km and 150 km distance between the stations, the network density seems to be more crucial in practice than the results from non-users indicate. Choice simulations showed that a denser network and more attractive charging locations can weigh up each other to a certain degree.

Concluding, the results provide interesting insights into what might really keep people from using BEVs for long-haul trips. First, charging times longer than 30 min turned out very unattractive, they might even be an exclusion criterion for many people. The other impediment addresses leaving the motorway and thereby accepting detours.

Although range is not the only important factor, it still needs to be considered. From the given options, which were rather optimistic from today's point of view, but showed up relevant during a qualitative pre-study [8], there was a strong decrease in approval from 500 km to 300 km driving range, especially for the non-users. For the BEV-users, this gap is much smaller, so maybe non-users need to be addressed differently or be given the chance to test an e-vehicle for a period of time in order to remove their range anxiety and accept smaller battery capacities.

6 Limitations and Further Research

Even though insightful findings regarding users' decisions to use or not use BEVs were revealed, still a lot of future research has to be done. In this context, some limitations of the present approach should be considered for future research efforts.

A first methodological limitation regards the selection and number of attributes. Even though the attributes –range, charging time, network density, location– were thoroughly determined in a qualitative pre-study beforehand [8] and reflected acceptance relevant attributes, still, there are other attributes which need to be examined. It should be kept in mind that additional factors could still change the shift of importances of the attributes which were under study here. As such, the attractiveness of charging stations should be looked at, varying the infrastructure and the distance within one attribute.

A second limitation regards the study setting and the measurement scale, which are focused on relative preferences. Thus, absolute "no-go-criteria" in BEV usage, although very important for policy decisions, cannot be derived from the current study and should be tackled in future work. Also, the sample size was quite small, with a comparably small number of persons with BEV experience. Besides that, the findings

represent only a German view, which might be quite special in case of its car history and industry. Future studies should therefore validate the findings with a larger sample size and users from other cultures and countries, likewise.

Another –very essential– focus of future work is the question where to optimally localize the charging stations. While there are mostly technical, infrastructural and economic reasons for positioning the localizations in inner-urban areas, valuable input could be collected by user-centered localization maps, to closely adapt the planning of citizens' life-styles and habits. Currently, an empirical study is run in our working group, in which citizens in Germany have the possibility to place charging stations in an interactive map of their home cities and to argue why they would place the charging station at the respective location. From their location choices and, also, the relative density, an "optimal" (in terms of users' decisions) planning rationale can be derived, that might inform public planners and communes where citizens would like to have their charging locations.

Acknowledgements. This research was funded by the German Ministry of Economics and Technology (Project SLAM, reference no. 01 MX 13007F). The authors thank all participants for their patience and openness to share opinions on a novel technology. Furthermore, authors owe gratitude to Dr. Sylvia Kowalewski for highly valuable research input.

References

1. United Nations, Framework Convention on Climate Change (UNFCCC)—03/08/17: Paris Agreement. http://unfccc.int/files/essential_background/convention/application/pdf/english_paris_agreement.pdf
2. Bundesregierung - 03/08/17: Nationaler Entwicklungsplan Elektromobilität. https://www.bmbf.de/files/nationaler_entwicklungsplan_elektromobilitaet.pdf
3. Kraftfahrt-Bundesamt – 03/08/17: Personenkraftwagen am 1. Januar 2016 nach ausgewählten Merkmalen. Kraftfahrt-Bundesamt. http://www.kba.de/DE/Statistik/Fahrzeuge/Bestand/Umwelt/2016_b_umwelt_dusl.html?nn=663524
4. Jarass, J., Frenzel, I., Trommer, S.: Early Adopter der Elektromobilitaet in Deutschland - Wer sie sind und wie sie fahren. Internationales Verkehrswesen **66**(2), 70–72 (2014)
5. Hidrue, M.K., Parsons, G.R., Kempton, W., Gardner, M.P.: Willingness to pay for electric vehicles and their attributes. Resour. Energy Econ. **33**(3), 686–705 (2011)
6. Egbue, O., Long, S.: Barriers to widespread adoption of electric vehicles: an analysis of consumer attitudes and perceptions. Energy Policy **48**, 717–729 (2012)
7. Ziefle, M., Beul-Leusmann, S., Kasugai, K., Schwalm, M.: Public perception and acceptance of electric vehicles: exploring users' perceived benefits and drawbacks. In: Design, User Experience, and Usability, pp. 628–639. Springer (2014)
8. Halbey, J., Kowalewski, S., Ziefle, M.: Going on a road-trip with my electric car—acceptance criteria for long-distance-use of electric vehicles. In: Design, User Experience, and Usability—Interactive Experience Design, pp. 473–484. Springer (2015)
9. Vassileva, I., Madlener. R.: Perceptions and adoption of EVs for private use and policy lessons learned. In: Veneri, O. (ed.) Technologies and Applications for Smart Charging of Electric and Plug-in Hybrid Vehicles, pp. 283–300. Springer, Cham (2017)

10. Frischknecht, R., Flury, K.: Life cycle assessment of electric mobility: answers and challenges. Int. J. Life Cycle Assess. **16**(7), 691–695 (2011)
11. Kley, F., Lerch, C., Dallinger, D.: New business models for electric cars—a holistic approach. Energy Policy **39**(6), 3392–3403 (2011)
12. Sourkounis, C., Bingchang, N., Broy, A.: Pollution of high power charging electric vehicles in urban distribution grids. In: Compatibility and Power Electronics (CPE), 7th International Conference-Workshop, pp. 33–39. IEEE (2011)
13. Rauh, N., Franke, T., Krems, J.F.: User experience with electric vehicles while driving in a critical range situation—a qualitative approach. IET Intell. Transp. Syst. **9**(7), 734–739 (2015)
14. Franke, T., Günther, M., Trantow, M., Rauh, N., Krems, J.F.: The range comfort zone of electric vehicle users—concept and assessment. In: Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems, Vienna (2014)
15. González, J., Alvaro, R., Gamallo, C., Fuentes, M., Fraile-Ardanuy, J., Knapen, L., Janssens, D.: Determining electric vehicle charging point locations considering drivers' daily activities. Procedia Comput. Sci. **32**, 647–654 (2014)
16. Shahraki, N., Cai, H., Turkay, M., Xu, M.: Optimal locations of electric public charging stations using real world vehicle travel patterns. Transp. Res. Part D Transp. Environ. **41**, 165–176 (2015)
17. Bernardo, V., Borrell, J.R., Perdiguero, J.: Fast charging stations: simulating entry and location in a game of strategic interaction. Energy Econ. **60**, 293–305 (2016)
18. Javid, R.J., Nejat, A.: A comprehensive model of regional electric vehicle adoption and penetration. Transp. Policy **54**, 30–42 (2017)
19. Jabeen, F., Smith, B., Olaru, D., Greaves, S.: Electric vehicles adoption: environmental enthusiast bias in discrete choice models. Transp. Res. Part D Transp. Environ. **51**, 290–303 (2017). Sydney
20. Rauh, N., Franke, T., Krems, J.F.: Understanding the impact of electric vehicle driving experience on range anxiety. Hum. Factors **57**(1), 177–187 (2014)
21. Zaunbrecher, B.S., Beul-Leusmann, S., Ziefle, M.: Laypeople's perspectives on electromobility—a focus group study. internet of things. In: Giaffreda, R., Cagánová, D., Li, Y., Riggio, R., Voisard, A. (eds.) IoT Infrastructures, pp. 144–149. Springer, Cham (2014)
22. Schmalfub, F., Mühl, K., Krems, J.F.: Direct experience with battery electric vehicles (BEVs) matters when evaluating vehicle attributes, attitude and purchase intention. Transp. Res. Part F Traffic Psychol. Behav. **46**, 47–69 (2017)
23. Philipsen, R., Schmidt, T., van Heek, J., Ziefle, M.: Fast-charging station here, please! User criteria for electric vehicle fast-charging locations. Transp. Res. Part F Traffic Psychol. Behav. **40**, 119–129 (2016)
24. Duncan, L.R., Tukey, J.W.: Simultaneous conjoint measurement: a new type of fundamental measurement. J. Math. Psychol. **1**(1), 1–27 (1964)
25. Orme, B.: Interpreting the Results of Conjoint Analysis, Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research, pp. 77–89. Research Publications LLC, Madison (2010)
26. Parasuraman, A.: Technology Readiness Index (TRI) a multiple-item scale to measure readiness to embrace new technologies. J. Serv. Res. **2**(4), 307–320 (2000)
27. Philipsen, R., Schmidt, T., Ziefle, M.: Well worth a detour? - Users' preferences regarding the attributes of fast-charging infrastructure for electromobility. In: Stanton, N., Landry, S., Di Bucchianico, G., Vallicelli, A. (eds.) Advances in Human Aspects of Transportation, pp. 937–950. Springer, Cham (2017)

Fast-Charging Stations or Conventional Gas Stations: Same Difference? - Variations of Preferences and Requirements

Ralf Philipsen(✉), Teresa Schmidt, and Martina Ziefle

Human-Computer Interaction Center (HCIC), RWTH Aachen University, Campus-Boulevard 57, 52074 Aachen, Germany
{philipsen, schmidt, ziefle}@comm.rwth-aachen.de

Abstract. A need-based fast-charging network for battery electric vehicles provides an opportunity to remove widespread impediments to adoption and thereby increase market penetration of electromobility. In that regard, it is crucial to know, whether experience with gas station planning, regarding position and business models, can easily be transferred to charging contexts. A large questionnaire study (N = 320) was conducted to evaluate potential commonalities and differences between conventional gas and novel fast-charging stations and provided on-site services. Therefore, both BEV users and drivers of internal combustion engine vehicles were questioned. It was revealed that most requirements for gas stations are transferable to charging stations, but brands lose importance in charging context. Payment preferences shift from payment by card and cash to fully automated car-based processes. It was also shown that even people without charging experience clearly differ between gas and charging stations and their requirements for charging stations converge towards BEV users' views.

Keywords: Electro mobility · Infrastructure planning · BEV experience · On-site services and amenities · Payment methods · Business models · Parallel activities · Human factors

1 Introduction

Currently, technical developments evolve very fast – likewise the number of economic challenges that must be addressed. One of the most important issues humanity needs to look at is global warming. Here, the amount of emission of greenhouse gases humanity is contributing to the challenge of global warming is alarming. A commitment to reduce these emissions, more precise CO_2, and to stabilize the average global temperature is focused in the so-called Paris Agreement, in which 197 nations acted on in 2015 under the auspices of the United Nations [1].

Addressing the reduction of combusting fossil fuels – a main contributor to CO_2 emissions made by humanity – could also help to diminish the dependence on exhaustible resources of modern societies. Industrial production and motorized traffic represent two of the main (internal) combustion-based contexts to look upon.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_91

Therefore, the German Federal Government published 2009 a national program to develop and promote the dissemination of electric vehicles [2]. The benefits of using more electrical vehicles instead of combustion-based engines are clear: reducing particulate matter in metropolitan areas and contributing actively with green technology to treat the environment with care. The high ambitioned goal of 1 million registered electric vehicles in Germany by 2020 seems therefore an appreciated measure in this matter. The government also promoted this goal by traffic privileges, buyer's premiums and tax breaks for (future) electric vehicle users. These measures seem to have worked to a limited extent, if one looks at the reported numbers of battery electric vehicles (BEV) and Plugin-hybrid electric vehicles (PHEV) by the German Federal Motor Transport Authority:

In 2009, about 1400 BEVs and 22.000 PHEVs were reported [3]. These numbers increased to about 34.000 BEVs and 165.000 PHEVs for 2017. Suggesting a healthy trend, the aimed goal is still far off - indicating major barriers, which prevent the replacement of conventional cars with internal combustion engines by electric vehicles.

Although the technology itself is constantly evolving [4], with e.g. higher ranges of electric vehicles or shorter charging times, and research broadly focuses on several fields of interest in this matter, like technical, economical [5] or environmental issues [6], conventional cars still dominate the market: in 2009 a total of 40.930.000 (gasoline and diesel) combustion engines were reported [3]. This number even increased up to 45 million in 2017, challenging the current developments to broaden research approaches towards a deeper understanding of the differences between both drive technologies, not only on a technical level, but also out of a user-centered social science perspective. Current user-centered studies about charging infrastructure focus on optimal planning of charging grids [7, 8] range anxiety [9] or take different user factors like individuals' environmental awareness [10] into account. The present work, however, will take a closer look on a one-on-one comparison between both technologies, focusing gas and charging stations.

2 Questions Addressed

Current simulations (e.g. about station placing) are often based on assumptions about charging behavior taking vehicle specifications into account, derive from gas station usage [11] or use system modelling [12]. At the same time, the large gas station chains are discussing whether or how the placing of charging stations at gas station sites can lead to a viable business model. But can one simply transfer gas station usage to charging stations? Are gas stations automatically attractive for e-vehicle users by installing charging stations? For a better understanding of the differences between gas and charging stations regarding preferences and requirements, it is necessary to include users of both facilities in research. That is, BEV users' views on charging stations should be compared with the attitudes towards gas stations of internal combustion engine vehicle drivers, hereafter referred to as non-users. In addition, it is crucial to gain deeper insights into the requirements towards charging stations of these conventional car drivers, as they could be future BEV users, if perceived barriers to adoption can be removed.

3 Methodology

To address the aforementioned questions, a web-based questionnaire study was designed with a focus on users' requirements for gas respectively charging stations and the provided on-site services.

3.1 Questionnaire Design

The questionnaire was based on previous focus group discussions and was structured as follows (see Fig. 1): First, all participants had to answer questions dealing with their demographic data, their mobility behavior in terms of car use and their commitment to technology [13]. In addition, BEV users were asked to provide information about their cars including usage, range and connector types. In the following, BEV users and non-users had to answer in different contexts: While BEV users were surveyed on their general requirements for charging stations, in terms of factors that influence their decision for a charging place, non-users had to answer the same questions in the context of conventional gas stations. Furthermore, information about payment methods and services at gas stations that are preferred by non-users were gathered.

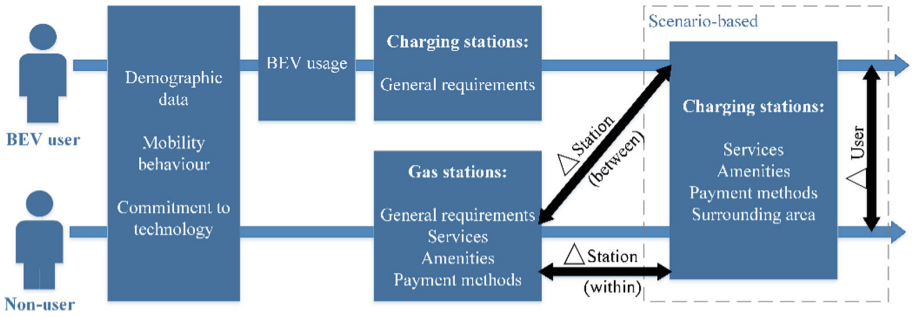


Fig. 1. Questionnaire design including setups for different user groups (Δ indicating conducted comparisons depending on station or user type)

After that, a scenario was presented to both user groups to calibrate their concept of BEV and charging processes, because there is a high inhomogeneity regarding charging speeds at currently used battery electric vehicles. The scenario was based on the current state of fast-charging development and determined that the participants should have in mind for the remaining questionnaire that their car's battery could be recharged during a charging stop that lasts 20–40 min. Concerning the given scenario and like the gas station context before, participants had to evaluate services, amenities, payment methods and additional opportunities to do something in the time spend during charging.

6-point Likert-scales were used for all questions aiming at participants' agreement or perceived importance (min = 0, no agreement | not important; max = 5, full agreement | highly important).

3.2 Data Acquisition, Preparation and Analysis

The participants were acquired at the university and predominantly in specialist forums dealing with motorized and electrified traffic. After finishing the acquisition phase, data was prepared by removing dropouts and unserious cases. Additionally, five participants without a driving license were excluded from analysis. The data was analyzed with both parametric (t-Tests, (M)ANOVA with Pillai's trace) and non-parametric (χ^2-Tests) inferential statistical methods to identify significant differences depending on user group or station type. The level of significance was set to $\alpha = .05$.

4 Sample

In total, 320 (N) participants were included in the analysis. 75.3% (n = 241) of them were male, 24.7% female (n = 79). The participants' age ranged from 18 to 75 years with a mean of 40.13 years (SD = 14.12). The sample was rather well educated: 57.2% (n = 183) stated that they achieved a university degree, 23.8% (n = 76) graduated from high school and 14.4% (n = 46) completed a vocational training.

Concerning car use, 247 participants (77.8%) owned a car and almost half of the participants (n = 157, 49.1%) used a car on a daily base, whereas 78 (24.4%) drove at least several times a week. 42.2% (n = 135) of the participants were users of battery electric vehicles, while 75.8% (n = 185) used only vehicles with internal combustions engines.

More than 90% of BEV users were male. They were significantly older ($t(318) = -7.355$, $p < .001$, $d = -0.833$) then non-users, had higher incomes ($U = 7841$, $p < .001$, $d = 0.670$), had a higher commitment to technology ($t(318) = -4.545$, $p < .001$, $d = -0.514$), were more frequent drivers ($U = 6270$, $p < .001$, $d = 0.940$), but did not differ concerning education ($p > .05$). Except for age and education, these differences are in accordance with early adopter characteristics in Germany [14]. For a detailed comparison see Table 1.

Table 1. Comparison of BEV users and users of internal combustion engine vehicles (non-users)

		BEV users	Non-users
N		135	185
Sex	Male	92.6%	62.7%
Age		M = 46.42 (SD = 10.78)	M = 35.54 (SD = 14.52)
Technology commitment		M = 4.09 (SD = 0.63)	M = 3.72 (SD = 0.78)
Education	University degree	49.7%	62.7%
	High school diploma	18.5%	27.6%
	Vocational training	23.7%	7.6%
Car usage	Daily	72.6%	31.9%
	Several times a week	22.2%	25.9%
	Several times a month	5.2%	21.6%
Income	>5.000€	34.1%	17.3%
	>4.000€	20.0%	11.9%

111 BEV users (82.2%) drove cars capable of fast-charging. The most frequently achieved range was between 100 and 150 km (n = 52, 38.5%), while 30 (22.2%) BEV users could realize even more than 300 km.

5 Results

The results section is structured as follows: First the comparisons between fast-charging stations and gas stations will be presented regarding participants' general requirements, perceived importances of both services and amenities as well as preferred payment methods. Second, particular demands on fast-charging stations will be outlined.

5.1 Contrasting Gas Stations and Charging Stations

The comparison between gas and charging station is based on four aspects: general requirements that influence the station selection, currently provided services, additional requested amenities and payment methods.

General Requirements. By looking at general requirements for gas respectively charging stations, it became obvious that the emerging picture looked quite different depending on the station type (see Fig. 2).

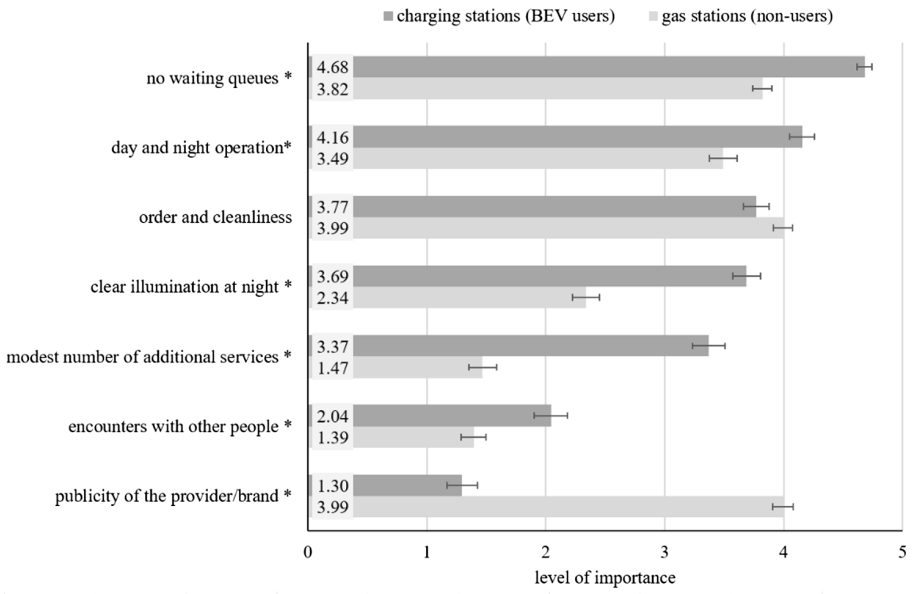


Fig. 2. Average importance and standard errors of general requirements on charging and gas stations (* indicating significant differences with $p < .001$)

The most important requirement for charging stations was the *absence of queues*, which indeed got high average importance scores for gas stations too, but differed significantly. In contrast, the *publicity of the provider/brand* was rated as most important for gas stations, while it was the most unimportant criteria for charging place decisions. Similarly, a *modest number of additional services* was rated as rather important for charging stations, while it was perceived as not important in the gas station context. In contrast, a *day and night operation* as well as *order and cleanliness* were rated as important for both gas and charging stations, whereas *encounters with other people* was rather unimportant in both contexts.

Multivariate testing indicated a significant main effect of station type on the general requirements with $V = .967$, $F(7, 312) = 1320.83$, $p < .001$. However, separate univariate ANOVAs revealed that there was no significant effect on the importance of order and cleanliness ($p > .05$).

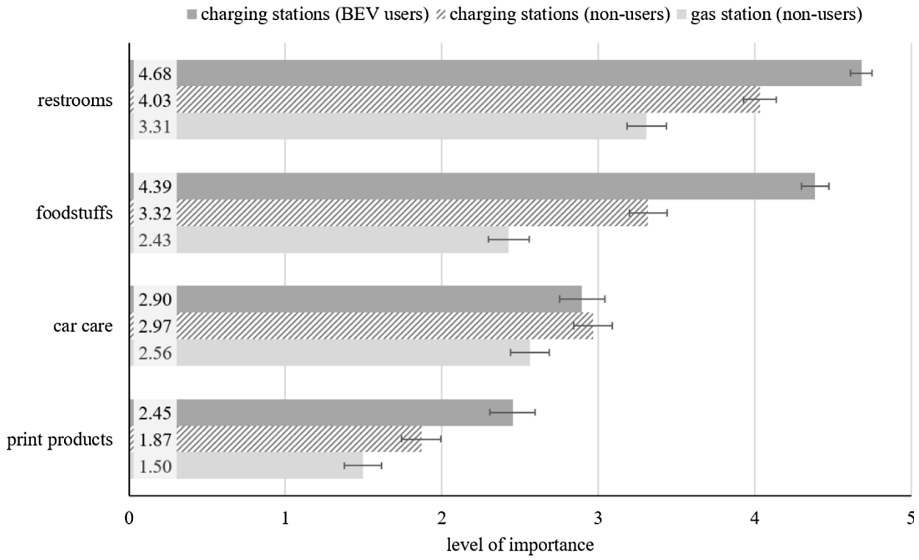


Fig. 3. Average agreement and standard errors regarding the question "What importance do you assign to current services at charging/gas stations?" distinguished by station type and user group

Provided Services. In view of provided services and sold goods at gas or charging stations, the study focused on *restrooms*, *car care* and the *sale of food and print products* as key factors identified in previous focus group studies. As can be seen in Fig. 3, *restrooms* got the highest importance levels in all contexts and for all user groups. In particular, BEV users assigned importance scores next to the scale maximum ($M_{BEV} = 4.68$, $SD_{BEV} = 0.79$). Same applies for the *sale of foodstuffs*, which was rated second important for charging stations ($M_{BEV} = 4.39$, $SD_{BEV} = 1.02$). In contrast, *car care* and the *sale of print products* were of considerable less importance. Especially at gas stations, *print products* were perceived as rather negligible ($M_{NON} = 1.50$, $SD_{NON} = 1.61$).

A significant effect of station type was found on the perceived importances (V = .911, $F(4, 315)$ = 804.18, $p < .001$) when looking at both user groups. All questioned items but *car care* ($p > .05$) were more important for charging stations than for gas stations. However, there were no completely opposing opinions. Looking at the differences between gas and charging stations mentioned by non-users only, repeated measurements revealed a significant effect of station type (V = .901, $F(4, 178)$ = 407.15, $p < .001$), too. Follow-up tests of single items showed that all products and services were significantly more important for charging stations than for gas stations ($p < .001$ for all items). In consideration of charging stations only, a significant main effect of user group was found (V = .937, $F(4, 312)$ = 1167.18, $p < .001$): BEV users stated significantly higher importances for *restrooms* ($p < .001$), *foodstuffs* (p. < .001), and *print products* ($p = .003$) than non-users, while there was no difference regarding *car care*.

Requested Amenities. Next to the evaluation of provided services, participants had to rate amenities they would like to see at gas respectively charging stations. Again, items derived from pre-studies had to be evaluated: *Wireless Local Area Network (WLAN) access*, *roofing of the station*, *information boards* presenting news and traffic data and *parcel stations* to receive and dispatch postal items. Considering gas stations, a *roofing* was the most and highly requested feature, followed by *information boards*, *WLAN* and *parcel stations*, that received neutral to slight approval rates. In contrast, the BEV users agreed that *WLAN* would be the most wanted amenity at charging stations, whereas *information boards* and *parcel stations* have been slightly refused. The high agreement to *roofing* did not differ in comparison to gas stations (see Fig. 4).

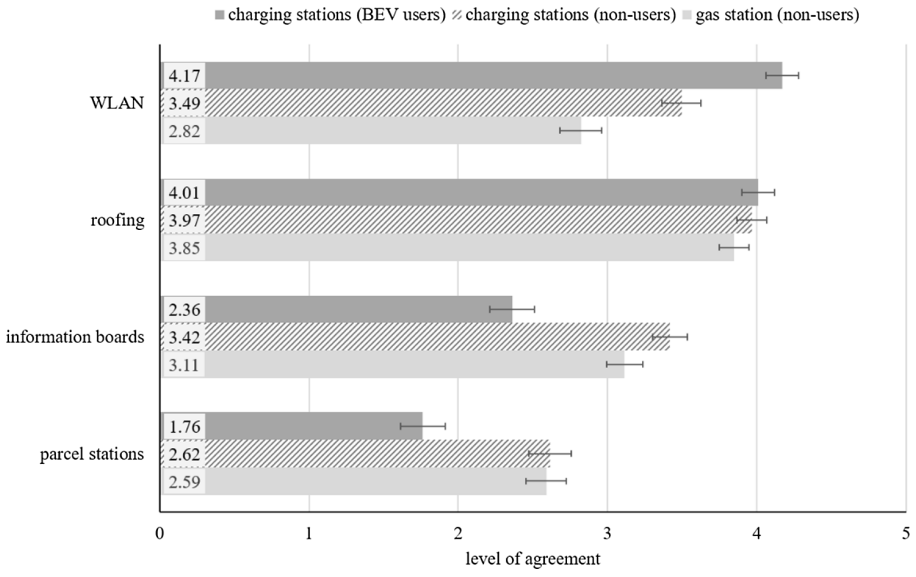


Fig. 4. Average agreement and standard errors regarding the question "What amenities would you like to see at charging/gas stations?" distinguished by station type and user group

Considering both user groups, multivariate testing revealed a significant main effect of station type ($V = .916$, $F(4, 315) = 860.45$, $p < .001$). *WLAN* got significantly higher agreement levels for charging stations by BEV users than for gas stations by non-users, while *information boards* and *parcel stations* were rated vice versa ($p < .001$ for all items). There was no significant difference regarding the *roofing* of stations ($p > .05$). Likewise, there was a significant effect of station type concerning non-users ($V = .926$, $F(4, 178) = 558.95$, $p < .001$). However, individual tests showed that in this user group there was no significant importance difference regarding *roofing* and *parcel stations* ($p > .05$), while non-users rated *WLAN* ($p < .001$) and *information boards* ($p = .001$) significantly more important in the charging station context. Focusing only on charging stations, there was a significant effect of user group ($V = .924$, $F(4, 312) = 941.80$, $p < .001$). The groups differed significantly regarding all amenities ($p < .001$) but *roofing* ($p > .05$).

Preferred Payment Methods. The ranking of payment methods is similar for both charging stations and gas stations (see Fig. 5). For both contexts *bank or credit cards* are the preferred payment method, followed by a *fully automated payment process* handled by the vehicle itself. Down in the third place is *cash* use. *Fuel or charging cards with PINs*, *customer accounts* at the actual provider and *prepaid cards* were the three least preferred methods to pay for fuel or power, that were preferred by four or less percent of the participants.

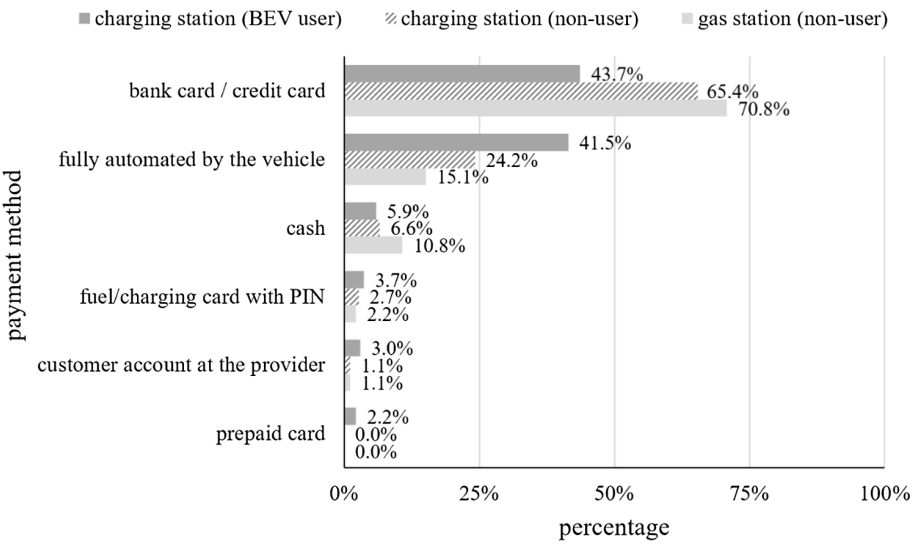


Fig. 5. Preferred payment methods (in percent) distinguished by station type and user group

Although the absolute ranking was comparable for gas and charging stations, there were major differences regarding the scope of preference. Again, there was an significant main effect of both BEV use ($\chi^2(5, N = 317) = 19.59$, $p = .001$, $V = .249$) and station type ($\chi^2(16, N = 182) = 285.78$, $p < .001$, $V = .627$): Approximately

two-thirds of the non-users preferred a *bank or credit card* as payment method irrespective of whether the payment takes place at a gas or charging station. In contrast, less than 50 percent of BEV users stated that this method to pay for charging is their favorite one. Their preference was also not unequivocal, because a *fully automated payment* received comparable scores. Focusing on non-users, it was found that their scores for payment methods at charging stations converge towards BEV users' preferences, though they do not reach their absolute extents. For example, non-users that preferred *cards* and *cash* at gas stations changed their preference in the charging context to a *fully automated payment*. Thus, *fully automated payment* was least preferred by non-users at gas stations (15.1%), got higher preference scores for charging stations by the same user group (24.2%) and was most preferred by BEV users regarding charging (41.5%).

5.2 Requirements on Fast-Charging Stations

Regarding preferred opportunities to spend time during the fast-charging process, it became obvious that only *shopping activities* in supermarkets or shopping centers and *visits to nature* in parks or on short walking trails got high average importance scores (see Fig. 6). Instead, *short-term working*, *cultural offerings*, *sports* and other *recreational services* were rated rather ambiguous, ranging from slight rejection (*sports*, M = 2.20, SD = 1.73) to slight agreement (*short-term working*, M = 2.74, SD = 1.72). The lowest approval rates were found for using *health care services* during charging time.

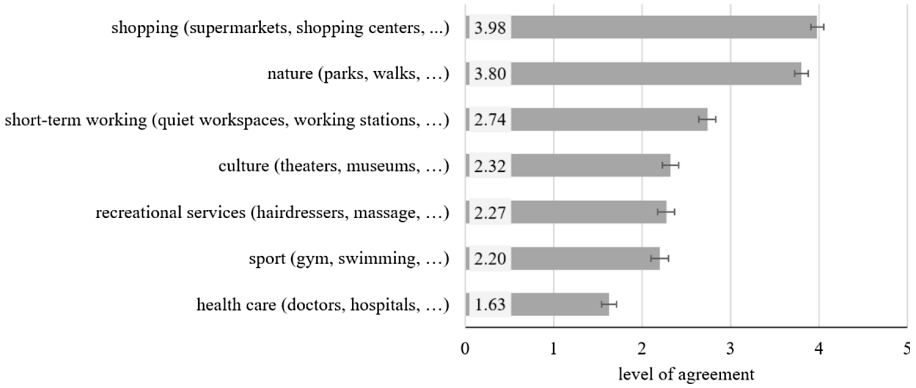


Fig. 6. Average agreement and standard errors to opportunities to spend time in the surrounding area of fast-charging stations (min = 0, max = 5).

In contrast to previously mentioned findings, BEV experience had no significant effect on all but one item: BEV users showed a lower average approval rate (M = 1.97, SD = 1.80) on *sports* as possible activity to spend charging time than non-users (M = 2.37, SD = 1.66) with $t(316) = 2.026$, $p = .044$, $d = 0.23$. However, the statements of both groups must be classified as (slight) rejection.

6 Discussion and Conclusion

The present study aimed at quantifying potential differences between fast-charging stations and conventional gas stations to factor them in during traffic simulations, charging infrastructure planning processes and development of business models. To understand differences or commonalities is crucial for establishing a need-based charging grid on a large scale as well as for offering optimal services and amenities on single charging station level that should be used and accepted. Indeed, both levels have to be taken into account to remove impediments to adoption.

It was revealed, that the general requirements for gas and charging stations differ in several ways. As expected, the absence of waiting queues was more important for charging stations, because the longer charging times compared to conventional fueling aggravate the negative consequences of queues. Same applies for a day and night operation, whose importance might be influenced by range anxiety and the current still deficient charging network. Most interestingly, the providers' brands played almost no role in charging station context. It is not clear, whether this is a general effect of technology or, more likely, of the current state of the market, which is highly fragmented. Except from some bigger electricity providers there are a lot of independent, small vendors and a consolidation towards a few market leaders has not occurred (yet). When the big gas station companies enter the charging market, it is questionable, whether the future electromobility adopters will maintain both their brand preferences and competence attribution and transfer them to the new technology, or if those companies will have to completely re-establish their reputation.

Looking at the preferred on-site services, it became obvious that users of both station types have comparable preferences. Although the attributed importance was almost always significantly higher in charging station context, the absolute ranking did not differ. Therefore, providing restrooms and selling foodstuff might be useful to increase the attractiveness of a charging location. Car care, which was rated similarly for both station types, might become more important for charging stations as option to bridge charging time, when washing and charging in parallel will become possible.

As expected due to the long charging stops, provided WLAN would be more important for charging stations. The lower agreement to information boards at charging stations might be related to the higher commitment to technology in BEV user group and a suspected better IT equipment. Smartphones and in-car infotainment can make information boards obsolete, if internet access is available.

The preferred payment methods identified in the present paper clarify that the currently fragmented payment systems at charging stations are not future-proof. Indeed, the preference of customer-account based payment, that is unusual for gas stations, but widespread in charging context, tends towards zero. Providers of charging stations should focus on payment by card and support the automated payment by car as soon as a reliable operation is technically possible. Especially BEV users showed high preference levels for this 'plug-in-and-forget'-payment.

Due to short fueling stops, the bypassing traffic is more important for gas stations than opportunities to spend time in the surrounding area. Nevertheless, this can be a crucial criterion to decide, which gas station location is suitable for charging station

placement. In accordance with [15], participants preferred locations with shopping facilities and public green spaces nearby.

Next to the previously mentioned differences, the station type differences within the group of combustion engine vehicle users were notably. It was revealed, that not only real usage experience influenced the evaluation, but also the given context itself: Even participants which had no experience with charging stations made clear distinctions between gas and charging stations regarding many of the questioned requirements, services and amenities. With a few exceptions, the preferences of non-users tend to conform to the ones of BEV users, which indicate that requirements for gas stations are not transferable ono-to-one into the charging station context and even non-users can have a realistic mental model of charging processes. Therefore, this user group should not be excluded during the planning of charging infrastructure.

7 Limitations and Outlook

The present study focused on preferred services, amenities, and payment methods and identified differences and similarities between gas and charging stations. However, it is still unclear how the actual use of both charging stations in general and on-site services fits to the users' individual preferences. It is likely that the decision to carry out a charging stop will be influenced by further factors that must be considered in future research, e.g., the length and type of trips, financial aspects, and the charging behavior in terms of embedding charging in the daily routine or personal handling of remaining range capacities.

The decision to use a scenario-based approach not only for non-users, but for BEV users, increased comparability, but limited the considered part of electric vehicles to a clearly defined technical setup. However, there is a broad spectrum of BEVs varying especially in range, if premium brands are considered. Therefore, higher ranges that lead to a lower frequency of charging stops might lead to a shift of requirements for charging stations. It is worth a (long-term) look, whether the requirements for charging stations will conform to the ones for gas stations, when technological development lead charging speeds that enable charging times comparable to conventional fuel times.

Therefore, the findings of the present study can only be an initial step towards both a need-based location planning and viable business models. Concerning the research question, whether fast-charging stations and conventional gas stations are equal or must be treated in separate ways, the answer for now has to be: It depends!

Acknowledgements. This research was funded by the German Ministry of Economics and Technology (Project SLAM, reference no. 01 MX 13007F). The authors thank all participants for their patience and openness to share opinions on a novel technology. Furthermore, thanks go to Dr. Philipp Brauner for valuable inspiration and to Alexander Kwiatkowski for his research assistance.

References

1. United Nations, Framework Convention on Climate Change (UNFCCC) - 03/08/17: Paris Agreement, http://unfccc.int/files/essential_background/convention/application/pdf/english_paris_agreement.pdf
2. Bundesregierung - 03/08/17: Nationaler Entwicklungsplan Elektromobilität, https://www.bmbf.de/files/nationaler_entwicklungsplan_elektromobilitaet.pdf
3. Kraftfahrt-Bundesamt – 03/01/17: Personenkraftwagen am 1. Januar 2009 nach ausgewählten Merkmalen. Kraftfahrt-Bundesamt, http://www.kba.de/DE/Statistik/Fahrzeuge/Bestand/Umwelt/2009/2009_b_emi_eckdaten.html?nn=663488
4. Kraftfahrt-Bundesamt – 03/01/17: Der Fahrzeugbestand am 1. Januar 2017. Kraftfahrt-Bundesamt, http://www.kba.de/SharedDocs/Pressemitteilungen/DE/2017/pm_06_16_bestand_01_16_pdf.pdf?__blob=publicationFile&v=3
5. Lu, J., Chen, Z., Ma, Z., Pan, F., Curtiss, L.A., Amine, K.: The role of nanotechnology in the development of battery materials for electric vehicles. Nat. Nanotechnol. **11**(12), 1031–1038 (2016)
6. Frischknecht, R., Flury, K.: Life cycle assessment of electric mobility: answers and challenges. Int. J. Life Cycle Assess. **16**(7), 691–695 (2011)
7. Kley, F., Lerch, C., Dallinger, D.: New business models for electric cars - a holistic approach. Energy Policy **39**(6), 3392–3403 (2011)
8. Sourkounis, C., Bingchang, N., Broy, A.: Pollution of high power charging electric vehicles in urban distribution grids. In: Compatibility and Power Electronics (CPE), 7th International Conference-Workshop, pp. 33–39. IEEE (2011)
9. González, J., Alvaro, R., Gamallo, C., Fuentes, M., Fraile-Ardanuy, J., Knapen, L., Janssens, D.: Determining electric vehicle charging point locations considering drivers' daily activities. Proced. Comput. Sci. **32**, 647–654 (2014)
10. Shahraki, N., Cai, H., Turkay, M., Xu, M.: Optimal locations of electric public charging stations using real world vehicle travel patterns. Transp. Res. Part D Transp. Environ. **41**, 165–176 (2015)
11. Rauh, N., Franke, T., Krems, J.F.: User experience with electric vehicles while driving in a critical range situation - a qualitative approach. IET Intell. Transp. Syst. **9**(7), 734–739 (2015)
12. Jabeen, F., Smith, B., Olaru, D., Greaves, S.: Electric vehicles adoption: environmental enthusiast bias in discrete choice models. Transp. Res. Part D Transp. Environ. **51**, 290–303 (2017). Sydney
13. Kitamura, R., Sperling, D.: Refueling behavior of automobile drivers. Transp. Res. Part A Gen. **21**(3), 235–245 (1987)
14. Lam, A.Y., Leung, Y.W., Chu, X.: Electric vehicle charging station placement: formulation, complexity, and solutions. IEEE Trans. Smart Grid **5**(6), 2846–2856 (2014)
15. Neyer, F.J., Felber, J., Gebhardt, C.: Kurzskala zur Erfassung von Technikbereitschaft (technology commitment). Zusammenstellung sozialwissenschaftlicher Items und Skalen (2016). doi:10.6102/zis244
16. Jarass, J., Frenzel, I., Trommer, S.: Early Adopter der Elektromobilitaet in Deutschland - Wer sie sind und wie sie fahren. Int. Verkehrswesen **66**(2), 70–72 (2014)
17. Philipsen, R., Schmidt, T., van Heek, J., Ziefle, M.: Fast-charging station here, please! User criteria for electric vehicle fast-charging locations. Transp. Res. Part F Traffic Psychol. Behav. **40**, 119–129 (2016)

Fuzzy Logic Based Merging Gap Acceptance Model Incorporating Driving Styles and Drivers' Personalities

Chen Chai[1(✉)], Xuesong Wang[1], Yiik Diew Wong[2], and Yidan Gao[2]

[1] College of Transportation Engineering, Tongji University, 4800, Cao'an Road, Shanghai, China
chaichen@tongji.edu.cn
[2] Center for Infrastructure Systems, Nanyang Technological University, 50 Nanyang Avenue, Singapore, Singapore

Abstract. Cognitive ability and decision-making process in the driving task vary among drivers. Differences in driving styles and driver's personality affect driving behavior, risk-taking and responses under hazardous scenarios. An innovative fuzzy logic based gap acceptance simulation model is proposed to extract and predict gap acceptance behavior with driving styles and drivers' personalities. The proposed approach models driver's decision-making process by a fuzzy set of the input factors. Cognitions, such as perceptions of front gap and current velocity, are modelled as fuzzy inputs while decisions, such as deceleration, are modelled as fuzzy outputs. For each driver, weighting factors are applied to differentiate driver's sensitivity to operational factors, such as front gap and relative velocity during gap acceptance behavior.

Keywords: Human factors · Driving styles · Fuzzy logic · Merging behavior · Gap acceptance

1 Introduction

Driving style and driver's personality vary among the drivers. The underlying style and personality are related to drivers' characteristics, such as age and gender and driving experience [1–3]. Differences in driving styles and driver's personality affect driving behavior, risk-taking and responses under hazardous scenarios. Therefore, applying a generalized model in simulating microscopic movements for all simulated vehicles is a relatively coarse approach [4]. Instead, it is desirable to analyze driving style and driver's personality of the drivers and classify them by driver type, so as to enhance accuracy and flexibility of microscopic simulation models.

Driving style and personality have been analyzed and modeled in several recent studies [5, 6]. Johnson and Trivedi [7] classified driving style into two broad categories: aggressive and non-aggressive. They proposed a Dynamic Time Warping (DTW) model and used smartphone sensors to detect and deduce driving patterns from driving experiments. Chen et al. [4] developed a driving habit graph (BHG) to model driving styles. The BHG model uses arcs and nodes to represent each driving behavior

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_92

in which driving style and driver's personality are extracted from experimental data from a driving simulator under hazardous scenarios. Although these previous studies have recognized and modeled driving style and driver's personality, a more adaptive and flexible approach is needed to incorporate driver's personality into existing microscopic simulation models.

Whenever several road lanes are merged to continue into fewer lanes, a merging area is defined. These merging areas, at expressway on-ramp as well as at lane-drop (decrease in number of lanes) zone at signalized junctions, are critical bottle-necks due to lesser capacity by virtue of fewer numbers of the merged lanes. This study develops a fuzzy logic based car-following model which analyzes driving style and driver's personality. In fuzzy logic, parameters are computed by varying degrees of truth. Herein, the uncertainty and approximation inherent to fuzzy logic is a well-suited approach to model complex decision-making process and driving behavior. Moreover, linguistic terms that are used in fuzzy logic can be used to represent drivers' cognitions, attitudes, and reactions [8]. For example, Al-Shihabi and Mourant [9] applied fuzzy logic in modelling car-following behavior. Furthermore, fuzzy logic can be incorporated into existing approach to extend current simulation model. In particular, Fuzzy Cellular Automata (FCA) model, that embeds fuzzy logic in conventional Cellular Automata models, has been developed and applied in traffic simulation [10, 11]. It has been found that the FCA approach is able to consider cognitions and decision-making process of each driver with a high computation efficiency.

2 Fuzzy Logic Based Model with Drivers' Personalities

The decision-making process of each driver of merging traffic is described in Fig. 1. Three sets of input factors, defined as lane configuration, operational attributes of subject stream, and operational attributes of other stream, all have impact of decision-making of vehicles. For each driver, decisions on entering or not entering the merging area must be made

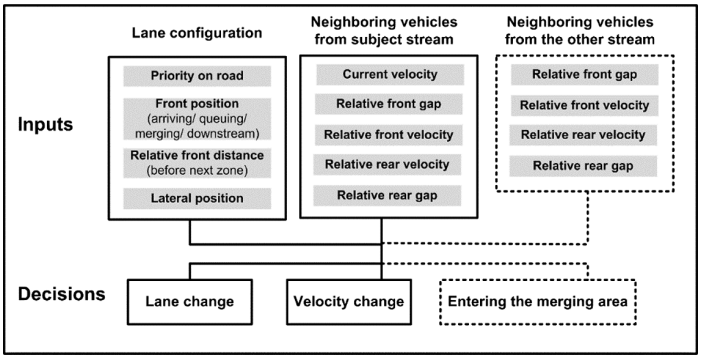


Fig. 1. Decision-making process of merging

In this study, the cognition system is modelled by fuzzy inputs while driver decisions are modelled by fuzzy outputs, as shown in Fig. 2. According to the decision-making process of each individual driver, three fuzzy sets are defined to model vehicle movements, as summarized in Fig. 3.

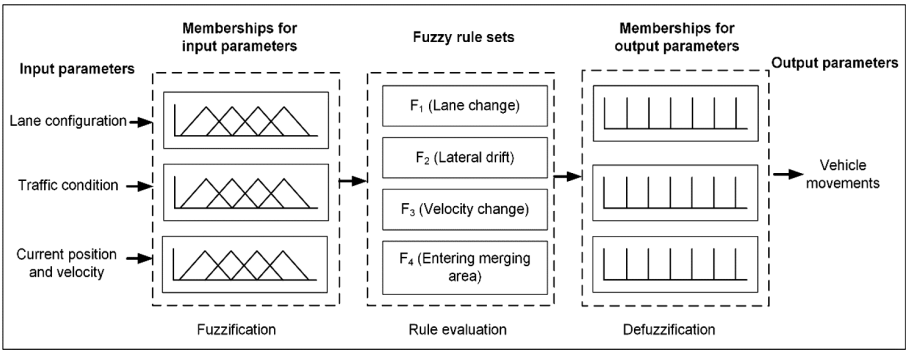


Fig. 2. Operation diagram of fuzzy sets and rules

Subset	Inputs
Lane configuration ($F_{1\text{-}1}$/ $F_{2\text{-}1}$/ $F_{3\text{-}1}$)	**Priority on road** {Yes; No} **Front position** {Arriving; Queuing; Merging; Downstream} **Distance to entering next zone** {Far; Medium; Close} **Lateral position** {Lane 1; Lane 2;…}
Neighbouring vehicles from subject stream ($F_{1\text{-}2}$/ $F_{2\text{-}2}$/ $F_{3\text{-}2}$)	**Current velocity** {Slow; Medium; Fast} **Relative front gap** {Close; Medium; Far} **Relative front velocity** {Opening fast; Opening; About zero; Closing; Closing fast} **Relative rear gap** {Close; Medium; Far} **Relative rear velocity** {Opening fast; Opening; About zero; Closing; Closing fast} **Relative lateral gap** {Close; Medium; Far}
Neighbouring vehicles from the other stream (F_4)	**Relative front gap** {Close; Medium; Far} **Relative front velocity** {Opening fast; Opening; About zero; Closing; Closing fast} **Relative rear gap** {Close; Medium; Far} **Relative rear velocity** {Opening fast; Opening; About zero; Closing; Closing fast} **Relative lateral gap** {Close; Medium; Far}

Fig. 3. Linguistic terms in the proposed fuzzy sets

Therefore, this study attenuated existing rules in fuzzy merging model to involve driver's style and style. In the new model, instead of using intersections of linguistic inputs, the final decision is combined from three decisions computed of each input, as shown in Table 1 Furthermore, weight factors are involved to model different impacts of the three input factors, as in Eqs. 1 to 2.

Table 1. Fuzzy rules in the proposed merging model

	Input	Gap acceptance
1	Current velocity is Slow	GA_1
2	Relative front velocity is Closing	GA_2
3	Front gap is Close	GA_3
4	Rear gap is Close	GA_4
5	Relative rear velocity is Closing	GA_5

$$(\mathrm{GA}) = (\omega_1 * GA_1 + \omega_2 * GA_2 + \omega_3 * GA_3 + \cdots \omega_\mathrm{n} * GA_n) \tag{1}$$

$$\omega_1 + \omega_2 + \omega_3 + \cdots \omega_\mathrm{n} = 1 \tag{2}$$

Where GA is combined gap acceptance for particular driver, GA_1, GA_2 are the gap acceptance computed from each input factor. ω_1, ω_2, are the weight factors.

Fuzzy sets and rules are embedded with CA model by using the Fuzzy Logic Toolbox of Matlab. By using the 'Evalfis' function in Matlab, response of each vehicle at each time step will be computed according to dynamic traffic flow and signal timing. The output response will be used to provide vehicle velocity and position at the next time-step, as shown in Fig. 4.

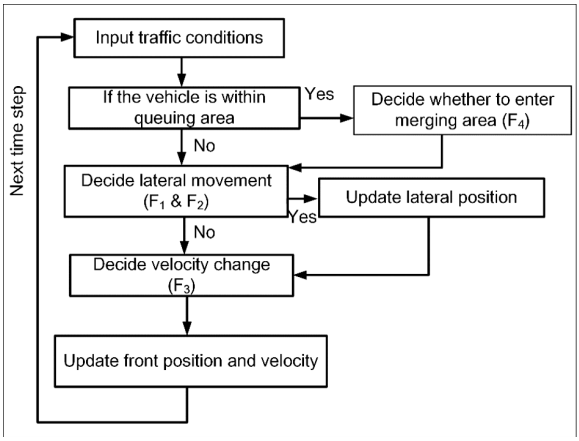


Fig. 4. Structure of the FCA model

3 Model Calibration and Validation

Field observations are conducted at eight survey sites at different junctions to collect merging vehicle movements as summarised in Table 2. Among the four departure legs, four have lane-drop along the departure lanes. The field observations were made on for a peak hour (6:00–7:00pm) during weekdays. Automatic vehicle detection and tracking algorithm based on video processing was applied to extract trajectories of vehicles [11]. Movement characteristics, including maximum velocity, maximum acceleration and deceleration rates, and minimum gap between vehicles (front, rear and lateral) in each zone are calibrated based on these field observations.

Table 2. Lane configurations of observed sites

Site No.	No. of lanes			Signal control
	Major	Minor	Merging	
1	3	1	3 to 2	Signalised
2	3	1	3 to 2	Signalised
3	4	2	4	Signalised
4	3	1	3	Signalised
5	2	1	2 to 1	Unsignalised
6	2	1	2 to 1	Unsignalised
7	2	1	2	Unsignalised
8	2	1	2	Unsignalised

Furthermore, membership functions are calibrated based on observed vehicle movements. It is assumed that driver responses can be achieved using input factors and defined fuzzy rules. In this study, fuzzy rules and sets are defined to describe driving behaviour. Moreover, relationship between input factors and driver response can be obtained from field observations. As average perception-response time (PRT) at signalised junctions is set as 1 s in this study, driver response at each time shall be determined by traffic and signal conditions 1 s earlier. Using the back-stepping technique, membership functions of input factors are derived. Detailed calibration processes are described by way of an example of calibrating relative front gap within the merging area.

The methodology to calibrate membership functions of each input consists of three steps, as shown in Fig. 5. Figures 5c to 5e show the relationship between front gap and frequencies of making each decision. For each decision, membership functions are derived according to the shape of frequency histograms. Therefore, membership function for the front gap is derived as Fig. 5f.

In this study, the proposed FCA model is validated by observed vehicle trajectories at microscopic level. Firstly, simulated vehicle trajectories are compared against field data from the survey sites. In order to generate the same initial headway distribution, observed arrival distribution and vehicle density are used to generate vehicles in the simulation, and density of observed/simulated vehicles (ρ) is plotted as shown in

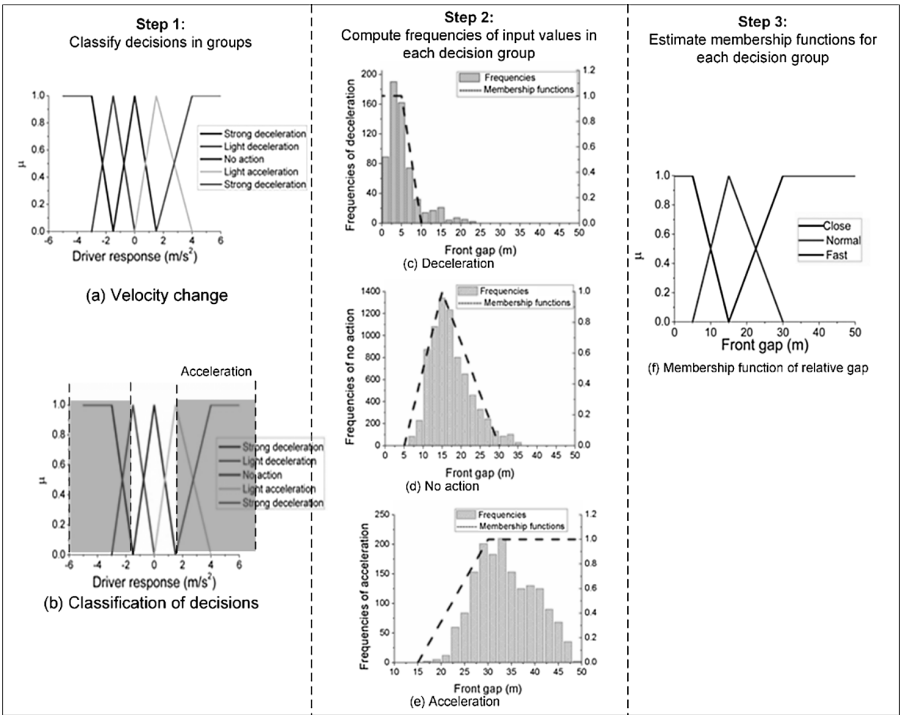


Fig. 5. Calibration of membership functions

Figs. 5a and 5b. An error assessment for each cell is also done for all the eight survey sites. The Root Mean Square Error (RMSE) and Mean Percentage Error (MPE) are relatively small and acceptable errors (around 8%) in comparison with the simulation and field data which present evidence that the CA model can well describe traffic dynamics at the microscopic level (Fig. 6).

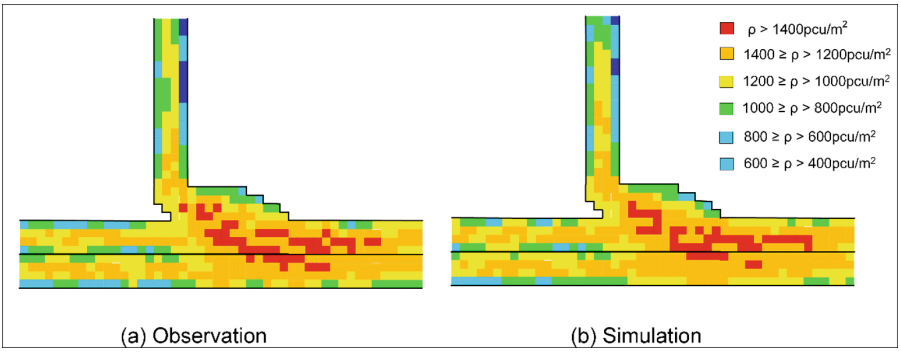


Fig. 6. Comparison of observed and simulated vehicle density (ρ)

4 Conclusions

This study improves existing fuzzy logic based merging gap acceptance model by involving driver's personality. Different from fuzzy rules which are conditional statement in previous models, the model proposed in this paper involves weight factors to represent the different impact of input factors on individual driver.

Compared to existing microscopic simulation models, contributions of the proposed fuzzy logic based car-following model can be summarized as follows: First, the proposed fuzzy logic based approach is able to explain decision-making process in a realistic and straight-forward way. As linguistic terms are used to define each input factor, graded cognitions are simulated from surrounding situations. Moreover, compared to behavioral simulation models, the proposed fuzzy logic based approach is able to model individual differences of each driver. Last but not least, by leveraging on calibrating weight factors and membership functions, the model is able to provide an individual analysis on the risk-taking behavior and the impact of different cognitions.

Acknowledgments. This work is supported by Singapore Ministry of Education Academic Research Fund Tier 2 MOE2013-T2-2-073.

References

1. Clarke, S., Robertson, I.: A meta-analytic review of the Big Five personality factors and accident involvement in occupational and non-occupational settings. J. Occup. Organ. Psychol. **78**, 355–376 (2005)
2. Taubman-Ben-Ari, O., Yehiel, D.: Driving styles and their associations with personality and motivation. Accid. Anlys. Prev. **45**, 416–422 (2012)
3. Garrity, R.D., Demick, J.: Relations among personality traits, mood states, and driving behavior. J. Adult Dev. **8**, 109–118 (2001)
4. Chen, S., Fang, C., Tien, C.: Driving behaviour modelling system based on graph construction. Transp. Res. Part C Emerg. Technol. **26**, 314–330 (2013)
5. Wang, J., Zhu, S., Gong, Y.: Driving safety monitoring using semi-supervised learning on time series data. IEEE Trans. Intell. Transp. Syst. **11**(3), 728–737 (2010)
6. Wang, W., Zhang, W., Guo, H., Bubb, H., Ikeuchi, K.: A safety-based approaching behavioural model with various driving characteristics. Transp. Res. Part C Emerg. Technol. **19**, 1202–1214 (2011)
7. Johnson, D.A., Trivedi, M.M. Driving style recognition using a smartphone as a sensor platform. 14th International IEEE Conference on Intelligent Transportation Systems, pp. 1609–1615 (2011)
8. Hurwitz, D.S., Wang, H.Z., Knodler Jr., M.A.: Fuzzy sets to describe driver behavior in the dilemma zone of high-speed signalized intersections. Transp. Res. Part F **15**(2), 132–143 (2012)

9. Al-Shihabi, T., Mourant, R.R. Toward more realistic driving behavior models for autonomous vehicles in driving simulators. In: 82nd Annual Meeting of the Transportation Research Board, Washington, DC (2003)
10. Chai, C., Wong, Y.D.: Fuzzy cellular automata (FCA) traffic models at signalized intersections. Comput. Aided Civ. Infrastr. Eng. **30**(12), 951–964 (2015)
11. Chai, C., Wong, Y.D., Er, M.J., Gwee, E.M.T.: Fuzzy logic based observation and evaluation of pedestrians' behavioral patterns by age and gender. Transp. Res. Part F Traffic Psychol. Behav. **40**, 104–118 (2016)

Road and Rail: Rail Passengers

Providing Improved Crowding Information to Provide Benefits for Rail Passengers and Operators

James Pritchard(✉)

Transportation Research Group,
University of Southampton Boldrewood Campus, Room 4001, Building 176,
Burgess Road, Southampton SO16 7QF, UK
J.A.Pritchard@soton.ac.uk

Abstract. This paper outlines the fact that crowding is an increasing problem on train services across the world. New ways of tackling the problems are needed, and providing improved information to help passengers make informed decisions is one option. This paper summarizes some stated preference work undertaken in the UK, which shows that improved information provision could in some cases be effective, and considers how it could be presented. The challenges of providing such information are then considered; one option is to rely on existing data sources, and the merits of a number of different sources are discussed.

Keywords: Rail · Information provision · Data acquisition

1 Introduction

There is evidence across the globe (for example in Australia [1], the UK [2] and the USA [3]) that travel habits are changing and that demand for rail travel is increasing. Overcrowding of rail services in major cities has become a worldwide problem, for which finding solutions is a big challenge for authorities [4]. Studies suggest that when between 50% and 70% of the seats on the train are occupied then passengers experience a disutility such as reduced physical comfort, lower productivity, or increased stress levels [5, 6]. Crowding, and related factors, are consistently shown to be of importance for transit passengers, and are occasionally even the dominating factor when it comes to customer satisfaction [5]. It has also been found that delays on commuter trains invariably increase with passenger density [6], which in addition to the further negative effect on passengers, makes it harder for operators to provide a reliable service.

A number of behavioral responses to crowding have been observed, including: Departing earlier or later to avoid crowding, waiting for a less crowded service, departing from a different station, or choosing a less crowded carriage [5]. However, "such behavior does not seem as typical as high standing valuations might suggest" [7]. There may be a number of good reasons for this discrepancy, including a lack of sufficient flexibility (time constraints when travelling) or insufficient service frequency (the gap between trains may be too long to justify waiting), but insufficient information is likely to be a key factor. It is hard for passengers to make a reasoned choice about

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_93

whether it is worth making alternative plans (such as waiting for a later train) when crowding levels are rarely communicated [8]. This is backed up by observations that the provision of crowding information can influence passenger behavior; for example, the communication of expected crowding levels achieved favorable results during the Sydney Olympic Games [4].

By providing crowding information, rail operators could encourage behavioral change amongst their passengers. As well as potentially benefitting those who are willing and able to change their travel behavior, it may additionally allow those who stick with their travel habits to benefit from reduced crowding levels. Rail operators also stand to benefit if information provision is successful in encouraging behavioral change – a uniform spread of passengers means better capacity utilization and a reduced risk of excessive boarding and alighting times. The first part of this paper summarizes recent stated preference studies conducted on train services in the UK between London, Gatwick Airport and Brighton [8, 9], which reinforce the idea that information provision could encourage behavioral change. Insights are provided in to the presentation of such information, and the impact it may have on different passenger groups.

In the few instances where crowding information are already provided in the UK [10], it is typically estimated based on historical data and is infrequently updated. Emerging technologies and a desire to extract more value from data [11] have led to an interest in providing more accurate, real-time, crowding information. The second part of this paper discusses the range of data sources that could be used to provide the level of information required to have an impact on passenger behavior. The advantages and disadvantages of the different sources are outlined, and consideration is given to possible additional benefits (and concerns) for passengers and operators alike.

2 Methodology: Stated Preference Surveys

On-train surveys were conducted on-board Gatwick Express services between London Victoria, Gatwick Airport and Brighton, with the full co-operation of the train operator (Govia Thameslink Railway). The surveys were undertaken using tablet PCs, and used stated preference (SP) methodology to ascertain how passengers might change their behavior when presented with crowding information. A pilot study was conducted in February 2016 [9], and after some refinements, a full study was conducted in April 2016. In each case, participants were asked some questions about their journey and were given three SP exercises. The first SP exercise was concerned with train crowding information that might be presented to passengers when planning or booking a journey in advance. It is assumed that such information would be based on historical estimates, and is not considered in this paper. The other SP exercises were concerned with information provided at the station, in real-time. One considered participants' willingness to wait for a later, less crowded, train if the first one due was shown to be full and standing. The other considered participants' willingness to move along the platform in order to board the next train at a less crowded carriage.

When considering the willingness to wait for a later train, survey participants were given a set of six pairs of upcoming departures from their origin station. The first train (given as 'Due') was always shown to be crowded, with no seats available. A second

train, assumed to follow the same route and journey timings, was shown departing a number of minutes later and with a reduced level of crowding. Participants were asked to choose one of the departures, with the wording in the full study making it clear that the choice was between boarding the current train or waiting at the station. Each of the six cases – shown to participants in a random order – contained a variation on either the time gap until the second train or on its level of crowding. Some participants were randomly chosen to be given crowding information in a graphical format, whilst the others were given numbers of seats in a text-based format; examples of the displays used in the full study are shown in Fig. 1.

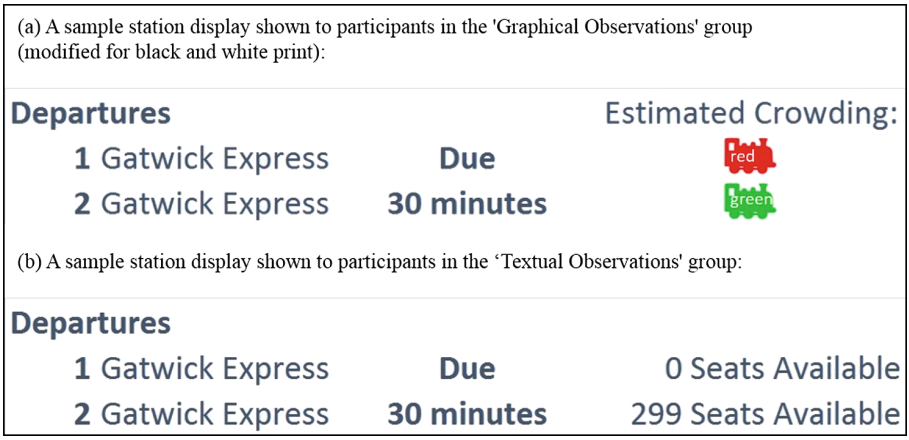


Fig. 1. Sample station displays shown to survey participants [8]

In the graphical case, red was used to represent no available seats, orange represented an occupancy level of 60%, and green represented an occupancy level of 10% of the number of seats. When text-based information was given instead, the numbers of seats given followed this same pattern, and was based upon the rolling stock used on the route at the time. The intermediate level of crowding was set at an occupancy of 60% of the number of seats, because it corresponds to the threshold at which crowding impacts typically begin to be felt [7].

When considering the choice of carriage, the survey was based on a simple three-carriage train formation. An illustration of the formation (Fig. 2) was provided to participants at the start of the SP exercise, giving the relative positions of the platform entrance at the origin station and the platform exit at the destination station. The SP exercise comprised nine example display boards (in a random order) in which Carriage B was always given as having no seats available, and the crowding levels of carriages A and C were varied. Consistent with the exercises concerning the whole train, red represented a crowded carriage (no seats available), orange represented a busy carriage (60% of seats occupied) and green represented an empty carriage (10% of seats occupied). A sample display board is shown in Fig. 3. All participants were shown graphical information; there was no text-based alternative.

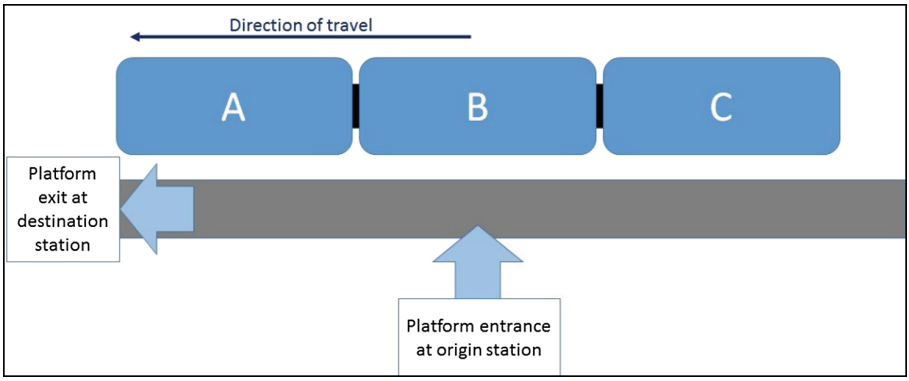


Fig. 2. Train formation shown to survey participants in the full study

In each case, the statistics package Stata was used to fit alternative specific constant logit models (McFadden's Choice Model) [12, 13] to the data collected [8, 9].

In summary, the model assumes that, given a choice of trains i, the probability of choosing train T is described as follows, where Ui is the utility function of train i.

$$P(T) = \frac{e^{U_T}}{\sum e^{U_i}}. \tag{1}$$

The earlier train was chosen as the base alternative, with a utility function given by:

$$U_{FirstTrain} = 0. \tag{2}$$

For the full study, the utility function for the later train was constructed as follows:

$$U_{LaterTrain} = \beta_{some_seats}s1 + \beta_{many_seats}s2 + \beta_{offset}t + \beta_{text_based}x. \tag{3}$$

Where:

- s1 is a Boolean which is 1 when the crowding information displayed 40% of seats on the later train as being available, and 0 otherwise.
- s2 is a Boolean which is 1 when the crowding information displayed 90% of seats on the later train as being available, and 0 otherwise.
- t is the time offset of the later train, in minutes
- x is a Boolean which is 0 when crowding information is displayed graphically, and 1 when text-based crowding information is given.

The model for choosing a carriage c has the same form:

$$P(c) = \frac{e^{U_c}}{\sum e^{U_i}}. \tag{4}$$

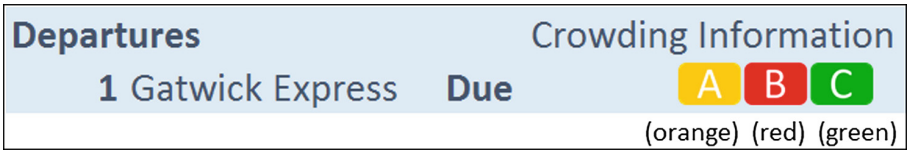


Fig. 3. Sample station display shown to survey participants

The Utility Function for Carriage i, U_i, was defined in this case as:

$$U_i = \beta_{some_seats} s1 + \beta_{many_seats} s2 + \beta_{first} A + \beta_{last} C. \quad (5)$$

s1 and s2 were again defined as mutually exclusive Booleans reflecting 40% and 90% of available seats respectively, whilst A and C are Booleans used to denote whether Carriage i is the first or last in the formation respectively.

3 Analysis of Stated Preference Surveys

3.1 Willingness to Wait for a Later Train

When considering the propensity of participants to wait for a less crowded train, the data were divided in to the following five groups, where the term "airport users" includes those who were flying in or out of Gatwick Airport, but not those who used the station at Gatwick for other reasons:

(1) Morning Commuters. Commuters travelling in the morning peak period between 6am and 9am. Airport users were excluded.
(2) Evening Commuters. Commuters travelling in the evening peak period between 4 pm and 7 pm. Airport users were excluded.
(3) Other Business Travelers. Non-commuting business travelers, irrespective of time of day. Airport users were excluded.
(4) Non-airport Leisure Travelers. Leisure travelers who were not airport users.
(5) Leisure Travelers with a flight to catch. Leisure travelers who had specifically stated that they were flying out of Gatwick Airport.

For each group, the willingness to wait for a later train was calculated using the following metric:

$$W = \left| \frac{\beta_{some_seats} s1 + \beta_{many_seats} s2 + \beta_{text_based} x}{\beta_{offset}} \right|. \quad (6)$$

W is an estimate of how long, in minutes, passengers would be willing to wait for a less crowded train. It is a useful metric because it helps to understand whether crowding information at the station would make a realistic difference for a given service frequency (or, alternatively, to ascertain the level of service frequency needed for waiting for the next train to be a realistic proposition). Values of W are given for the different passenger

groups in Table 1. β_{text_based} was not significant for either group of commuters, and the model was recalibrated to take account of this. Similarly, β_{some_seats} was insignificant for leisure travelers with a flight to catch.

Table 1. Values of the willingness to wait metric for different passenger groups

Passenger group	Willingness to wait (min)			
	Some seats, graphical info.	Some seats, text-based info.	Many seats, graphical info.	Many seats, text-based info.
Morning commuters	14	–	20	–
Evening commuters	17	–	17	–
Other business travelers	19	25	22	28
Non-airport leisure travelers	12	26	19	33
Leisure travelers with a flight to catch	–	–	9	28

The Gatwick Express has a peak-time service frequency of 15 min; on the assumption that stated preference surveys overstate the likely reality [8], it could be argued that the impact of crowding information for commuters at peak time will be minimal. However, there are additional factors that might influence this. Firstly, further investigation revealed a marked difference between morning commuters boarding at Brighton (a well-equipped terminal station) and those boarding at the subsequent intermediate stations (typically small and lacking in facilities) [8], with the willingness to wait at Brighton being as high as 23 min if the next train was showing as having many seats available. Secondly, many participants commented verbally that they perceived the reliability of the service to be poor, and indicated that this had an impact on their willingness to change their travel habits. The fact that β_{text_based} is not significant for commuters might reflect the fact that they expect the train to be crowded, and any perceived benefits of reduced crowding (e.g. less discomfort) are not tied to a guarantee of a seat.

On the other hand, other business travelers showed a greater willingness to wait. They responded much better to text-based information about seating availability, and were also more willing to wait if the later train had many seats available. This suggests that getting a seat is more important than simply having a reduced level of crowding, and may be tied to a need to use the travel time productively, and the fact that this sort of business trip may be longer than a regular commute. Similar patterns are seen in Table 1 for leisure travelers, although they valued their time less highly (β_{offset} was smaller) and the availability of seats more. Leisure travelers would seemingly much prefer text-based information than graphical information, but it is likely that this result highlights some unintended consequences of information provision. A high proportion of those using the

airport were not UK residents, did not have English as their first language, and may not have been familiar with the system. For those given graphical information, the colors may have implied other issues (e.g. orange is often associated with a warning) whilst those given text-based information may have (incorrectly) inferred that zero availability of seats meant that it was not possible to board the train at all.

3.2 Carriage Choice

The passenger groupings used to analyze the carriage choice exercise differed slightly from those used when considering the choice of train. Chiefly, because the exercise didn't include a time-dependency (which may have been perceived differently by those traveling home from work than it would be by those traveling to work), morning and evening commuters were treated as a single group.

The willingness to choose a carriage based on crowding levels varies between passenger groups in much the same way as the willingness to choose a different train does, with non-commuting business travelers and leisure travelers being the most responsive to crowding information. β_{first} was only found to be significant for commuters; this is probably because it is linked with being closest at the ticket barrier at the destination, and for those who don't travel at peak times, queues at the barrier are rarely an issue. β_{last} was always negative; given that choosing the last carriage involved walking a greater distance overall, this is not surprising.

Although the model predicts that some people would move down the platform to a less crowded carriage before boarding, it is hard to ascertain what would happen if the carriage closest to the platform entrance were not so crowded (although in this case there may be limited benefit from spreading the passengers out anyway). Furthermore, the location of the station entrance and exit relative to the carriages was explicitly stated here, and passengers may not be familiar enough with the train formation to make an informed choice on this basis in reality. Such a lack of information or familiarity may however be advantageous as there would be fewer factors taken in to account when choosing a carriage; passengers may be more willing to spread themselves out even if it did ultimately result in a longer walk or being further from the ticket barrier at the destination.

4 Use of Data to Provide Crowding Information

It is evident that providing crowding information could influence passenger behavior, but (in the UK), the static information currently available on a limited number of routes (e.g. [10]) has limitations. Firstly, although it may be a useful guide for some passengers, there is no expectation that it will accurately reflect the situation in real-time (especially in the event of disruption). Secondly, this information is only presented on a per-train basis, and doesn't help passengers to make an informed choice about which carriage to board. Emerging technologies mean that it is now possible to monitor and report passenger loading levels on a per-carriage basis in real-time. Although there is some evidence globally that rail operators have already begun to invest in such systems [14], they are far from ubiquitous at this point in time. There are a number of potential barriers to implementation, including cost, reliability, and the availability of supporting infrastructure.

A possible way of overcoming some of these barriers is to make use of existing systems to monitor passenger loading levels. This is in line with the aim of the Rail Delivery Group in the UK to extract more value from existing data [11] and has potential advantages in terms of both a lower cost-base and the fact that much of the required infrastructure is already in place. Reliability remains a key challenge, however; for passengers to confidently make informed decisions, the information provided must be both accurate and perceived as such. As part of an ongoing project in the UK, investigations are underway in to a range of possible data sources which could be used to count passengers on a train. This section aims to provide some initial insights in to the available data sources, highlighting some of the issues and limitations.

There are three basic components which must be addressed in order to provide useful train crowding information in real-time. Firstly, the number of people physically on a given train (the loading data) must be known. Secondly, the location of this train must be known, and it must be linked with a known service, so that the loading data can be reported in a timely and relevant manner. Finally, the system should have some predictive capabilities; to make informed decisions, passengers actually need to know how many people will be on the train when they board, not how many passengers are on it currently (especially if it is still a few stops away). This section is concerned mainly with the first component, and seeks to outline how the number of people on a train may be counted.

The available data sources can be categorized as follows:

(1) Sources tied to a physical train (e.g. data from onboard equipment).
(2) Sources tied to a particular station, which may help inform the number of people on a train (e.g. data from ticket barriers which counts the number of people going on to the platform).
(3) External data sources.

Advantages and disadvantages of some on-train data sources are given in Table 2; at station sources are considered in Table 3. The main type of external data sources which could be useful in this case is cellphone data. Cellphone providers can provide a rich set of additional demographic data, and offer a useful insight in to regular journey patterns. However, such data have limitations when it comes to actually counting people on trains – the data may not be available in real-time, and it can be difficult to infer the actual number of people from the number of devices identified on a given cellphone network. If a train journey is made entirely within one cell area it could also be difficult to correctly identify whether a device was on the train or not.

It is likely that a combination of data sources will be required to accurately estimate the number of people on a train – although the required accuracy will be dependent on the format of the information required. For example, loadweigh sensors may on their own be sufficient to provide graphical crowding information using a ‘traffic-light’ system or similar with only a few levels. In this instance, the aim is simply to provide an indication of crowding levels, so not being able to distinguish between, for example, four people or three people with luggage is likely to be immaterial. In fact, in some cases, it could be argued, that luggage and other items also contribute to the negative effects of crowding (e.g. lack of personal space), and load weight alone may be a good metric.

Table 2. Advantages and disadvantages of on-train data sources which could be used to count passengers

Data source	Data description	Key advantages	Key disadvantages
Loadweigh sensors	Weight carried in each carriage	Reported on a per-carriage basis. A standard feature of a lot of new rolling stock. Can be reported in real-time	Difficult to distinguish between people and luggage. Subject to variation around the mean weight of a person. Can only provide a net change at a station stop, not actual numbers alighting and disembarking
Onboard WiFi	Number of devices connected to on-board WiFi system	Potential to report on a per-carriage basis. Can be used to infer actual numbers alighting and disembarking at each station. May allow regular users and journey patterns to be monitored, providing a richer source of information and potentially aiding prediction capabilities	Potential privacy concerns. Likely to require users to be registered to use WiFi (and to potentially have paid for the service). Counts devices not people; some people won't have a WiFi enabled device, others may have several (e.g. phone and laptop)
CCTV	Number of people captured on CCTV images	Should allow number of people to be counted, irrespective of weight, number of mobile devices or amount of luggage. May allow passenger distribution to be analyzed on a per-carriage or even per seating area basis, depending on the number and the position of CCTV cameras	Large amount of data, with the potential to require a significant amount of processing. Processing could be costly, and likely to be 'offline', making this unsuitable for real-time updates

It was clear from the stated preference survey work, however, that providing numbers of available seats is important, particularly for some passenger groups. In this case, it is more important to be able to accurately estimate actual numbers of people on a train (although, if the trains are not designed with sufficient separate luggage provision, other items onboard could also restrict access to seats and hinder productivity). The disadvantages of the loadweigh data become more of a concern, and it may be necessary to supplement it with other sources. In this case, there are more implications for cost, and supporting infrastructure (additional systems, such as WiFi and CCTV need to be present), and additional concerns, such as data privacy. Questions are raised

Table 3. Advantages and disadvantages of at station data sources which could be used to count passengers

Data source	Data description	Key advantages	Key disadvantages
Ticket barrier/gateline	Number of tickets passing through ticket barrier on/off platform	Knowing the number of people on a platform helps predict likely loadings once the train has departed. Some ticketing systems will allow a richer source of information (e.g. journey patterns) to be collected	Not directly linked with the number of people on a given train; people may arrive on the platform some time ahead of the train they intend to catch. In many cases, more than one platform can be accessed through the same ticket barrier
At station WiFi	Number of devices connected to station WiFi system	May help predict likely loadings and provide information about journey patterns	As per on-train WiFi. Additionally, there is no guarantee that someone connected to the WiFi at the station will be catching a particular train, or even catching a train at all
CCTV	Number of people captured on CCTV images	With sufficient CCTV coverage and image recognition capability, the number of people boarding/alighting a given service ought to be obtainable	As per on-train CCTV

about whether it is right to use WiFi or cellphone data to infer journey patterns by tracking people – and whether passengers would be so keen to sign in to on-board WiFi systems if they knew that the data were being used for these purposes. It is clear that some of the data sources could provide hugely beneficial information to the train operators; for example, knowing regular journey patterns could help with decision making during times of disruption, whilst knowing demographic data from the cell-phone providers could help them with their marketing.

5 Conclusions

With a global growth in rail travel, crowding issues need to be addressed. Encouraging passengers to change their behavior by providing information, allowing them to make more informed decisions, may help alleviate some of the issues. As well as potentially reducing crowding and improving capacity utilization by spreading passengers out, the very ability of being able to make an informed decision may limit some of the negatives, such as stress.

Stated preference work conducted in the UK found (in accordance with existing literature) that passengers would be prepared to change their travel habits, and wait for a later train, or board a different carriage, in response to crowding information. Insights were gained in to the potential impact on different passenger groups, with leisure and non-commuting business travelers being the most responsive. Consideration was also given to the presentation of such information.

Providing improved crowding information is a challenge, as it requires the number of passengers on trains to be monitored in real-time. This paper considers the use of existing data sources to gather the necessary information, which could help overcome some of the barriers to implementation, such as cost. It is thought that some level of crowding information could be provided fairly easily from loadweigh sensors, which are fitted as standard to a lot of new rolling-stock in the UK. Many of the disadvantages of the data provided are unimportant if the crowding information only needs to be simple. However, the stated preference work found that some passengers responded better to more detailed information, such as the number of seats available, which is more difficult to provide. Use of some data sources raises other concerns, such as privacy issues, and it may be that operators are faced with trade-offs between the effectiveness of the information provided and the costs (financial and otherwise) of the implementation.

Acknowledgements. The authors would like to thank the Rail Safety and Standards Board (RSSB) and Govia Thameslink Railway (GTR) Ltd. for supporting this research.

References

1. Mees, P., Groenhart, L.: Travel to work in Australian cities: 1976–2011. Aust. Plan. **51**(1), 66–75 (2014)
2. Metz, D.: Peak car and beyond: the fourth era of travel. Transp. Rev. **33**(3), 255–270 (2013)
3. Kamga, C.: Emerging travel trends, high-speed rail, and the public reinvention of U.S. transportation. Transp. Policy **37**, 111–120 (2015)
4. Henn, L., Karpouzis, G., Sloan, K.: A review of policy and economic instruments for peak demand management in commuter rail. In: ATRF 2010 33rd Australas. Transp. Res. Forum, No. October, pp. 1–20 (2010)
5. Pel, A.J., Bel, N.H., Pieters, M.: Including passengers' response to crowding in the Dutch national train passenger assignment model. Transp. Res. Part A Policy Pract. **66**(1), 111–126 (2014)
6. Cox, T., Houdmont, J., Griffiths, A.: Rail passenger crowding, stress, health and safety in Britain. Transp. Res. Part A Policy Pract. **40**(3), 244–258 (2006)
7. Wardman, M., Whelan, G.: Twenty years of rail crowding valuation studies: evidence and lessons from british experience. Transp. Rev. **31**(3), 379–398 (2011)
8. Preston, J., Pritchard, J., Waterson, B.: Train overcrowding: investigating the use of better information provision to mitigate the issues. In: 96th Annual Meeting of the Transportation Research Board, vol. 44, no. 0, pp. 1–15 (2017)

9. Pritchard, J., Preston, J.: Investigating the potential to mitigate crowding issues on trains by providing improved information to passengers. In: Stanton, N.A., Landry, S., Di Bucchianico, G., Vallicelli, A. (eds.) Advances in Human Aspects of Transportation: Proceedings of the AHFE 2016 International Conference on Human Factors in Transportation, July 27–31, 2016, Walt Disney World®, Florida, USA, pp. 77–87. Springer, Cham (2017)
10. London Midland: Finding a seat. www.londonmidland.com (2016). http://www.londonmidland.com/your-journey/your-journey/finding-a-seat/. Accessed 21 June 2016
11. Rail Delivery Group: Capability Delivery Plan (2017)
12. StataCorpLP, "Stata Manual - asclogit." (2016)
13. McFadden, D.: Conditional logit analysis of qualitative choice behavior. Front. Econom. 105–142 (1973)
14. Japan Today: JR East launching app that will allow for real-time tracking of train services. www.japantoday.com (2014). https://www.japantoday.com/category/technology/view/jr-east-launching-app-that-will-allow-for-real-time-tracking-of-train-services. Accessed 01 Mar 2017

User Experience Evaluation on Ticket Gage of Subway Station: A Repertory Grid Approach

Linna Hu(✉), Yimu Yang, Long Liu, and Hua Dong(✉)

College of Design and Innovation, Tongji University,
Siping Road. 1239, Shanghai 200092, China
hulinna25657@126.com, donghua@tongji.edu.cn

Abstract. Methods of user experience evaluation play an important role in guiding designers in designing for a desired user experience, but methodological research on how to evaluate experience effectively is inadequate. The main aim of our study was to evaluate user experience with the ticket gate of subway station. This paper discussed Repertory Grid Technique (RGT), a method adopted from psychology that has been used in product development to help unveil users' subjective feelings about the product, and empirically illustrated a way in which the RGT may be conducted as an approach in field study for dealing with user experience evaluation. In conclusion, the application of this technique elicited valuable insight from respondents and provided promising direction for further research on user experience evaluation.

Keywords: User experience · Experience evaluation methods · Repertory Grid Technique · Subway station · Ticket gate

1 Introduction

Rapid urbanization and motorization have led to the increasing subway commuters on a large scale. With the rapid development of the subway construction, the ticket gate seems everywhere along with the subway station. The ticket gate is an important facility used to automatically examine tickets by means of its intelligent recognition system [1]. According to the different door mechanism, the ticket gates can be categorized into three types: tripod turnstile gate, flap-stile gate and slide-stile gate [1], among which slide-stile gate works best in terms of the transit of a large number of passengers [2]. However, the tripod turnstile gate is most widely used for the subway traffic in Shanghai. Hence, we were convinced of the need to figure out how people feel about the different ticket gates of subway station.

When we conducted a study aimed to evaluate user experience and by that to improve a product, we should figure out users' perceptions and subjective evaluative judgments of the product. In particular, when the core of user experience evaluation in our study was to investigate how people feel about using the ticket gate, the traditional usability methods are hardly reliable. Hence, practical methods to support the evaluation of user experience are needed. In this paper, we discussed the use of the

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_94

Repertory Grid Technique (RGT), a method adopted from psychology that has been used in product development to help unveil users' subjective feelings about the product. RGT, as a method of eliciting and collecting both qualitative and quantitative data of interacting with technology, enables an immediate and spontaneous assessment done by respondents including a comprehensive impression of user experience in terms of usability, aesthetics, emotions, pleasure, etc.

2 Background

According to the results of a survey, most respondents characterized UX as dynamic, context-dependent and subjective [3]. One element is always served as a part of experience, that is, a momentary feeling of pleasure and pain in various intensities [4]. To further specify, this feeling is the common "psychological currency" that allows user to qualitatively compare different experiences [5]. In addition, products trigger emotional and behavioral reactions through attributes that are perceived by the user, such as easy to use/hard to use, convenient/inconvenient, etc. [6]. A wide variety of models and frameworks have been proposed in order to represent the diverse elements that compose the UX. A model on the basis of cognitive psychology defined user experiences as an overall reflection of many identifiable and measurable constructs, thus breaking down the complexity of experiences into evaluative constructs [6]. According to Hassenzahl's assumption, product characters can be described by two attribute groups: pragmatic and hedonic attributes [7]. Pragmatic quality is associated with the product's usability and utility. Hedonic quality is related to the users themselves. Pragmatic aspects mainly concern the product utility while the hedonic aspects aim at the fulfillment of underlying psychological needs [7]. In our study we shifted the emphasis from the product to users' subjective feelings of interaction with the product, system, service or environment. The changed emphasis on the UX poses the question for evaluation of users' subjective feelings. Users construct their own personal version of the product quality based on the product per se and on their individual standards and expectations. This personal perception of the product brings judgments about the product's attributes.

3 Methods

3.1 An Overview of User Experience Evaluation Methods

A wide range of UX evaluation methods have already been applied in industry and academia. From 2004 to 2006, the European Union ENGAGE project collected a multitude of UX evaluation methods [8]. The collected methods were further classified into Generative and Evaluative, and the Evaluative methods were then categorized in three groups according to the focus of the method: Sensory characteristics, Expression or Meaning, and Emotional reactions [9]. The diverse methods presumably have been designed for different phase of design process, such as evaluating early concept and ideas, prototypes, or final products. Some methods aim at quantitative, some others qualitative data, which produce data on different levels of detail [10]. Traditional user experience methods often highlight usability criteria, which roughly in accordance with

the concepts of pragmatic quality [11]. More recent approaches increasingly focus on the subjective reactions, especially the emotional aspects of the user's experience, which are also referred to as hedonic quality aspects [11].

3.2 Repertory Grid Technique (RGT)

Repertory Grid Technique (RGT), as a methodological extension of Kelly's Personal Construct Theory [12], was first developed by a clinical psychologist named George Kelly in the 1930s and was later used to understand how people think and feel [13]. The operation of RGT is based on the assumption that everyone has a particular series of personal constructs, which means people have their own way to interpret the real world, and these constructs could be identified [12]. For example, we judge a person whom we first met through modeling constructs, say unfriendly or unfriendly, strong or weak, good looking or ugly, etc. According to these findings, the user experience varies from person to person since they all have different construct system. Consequently, a construct is a bipolar evaluation system, allowing two or more objects to be compared [14]. The goal of RGT is to explore and identify individual's construct, completed by a form of structured and reflective interview process where the respondents reveal their subjective experience. Szwejczewski et al. [15] and Lemke et al. [16] found that RGT could bring a greater depth and width of construct elicitation than did questioning alone. Moreover, much evidence has shown that the RGT is effective for evaluating user experience. Over the last decades, RGT has been utilized to consumer research, where it could help respondents verbalize the way they think about the particular product or service. It provides deep insight into user perceptions, offering an unbiased account of how they view a product or service. The sets of data gleaned from a collection of interviews served as input to the later analysis.

The core of the technique is the Repertory Grid, conducted by use of 'How' questions during a one-to-one interview. A Repertory Grid session involves a research and a number of participants who are familiar about the product being researched. The process of RGT can be divided into five steps, i.e. elements selection, constructs elicitation, rating, analysis of Repertory Grid data, results presentation [12]. For starters, respondents are presented with a certain number of pictures, namely elements in this case, and then invited to pick out which one is different from the rest. Alternatively, in many cases, interviewee provides a random group of three elements, and the respondent is asked to choose which two elements are similar to each other and at the same time different from the third one. The quality which makes the chosen one or two stand out from the rest is noted as one pole of a pair of constructs and should be described in the respondent's own words. To complete the bipolar constructs, the respondent needs to propose the opposite of the elicited pole. Next, each element as the row heading is rated on a Likert scale, together with each pole of the construct (as column headings), which represents the extremes of the scale. As a result, each element is quantitatively linked with the qualitative constructs by the use of rating system [14]. Figure 1 provides an overview of the process. The process of constructs elicitation and rating is repeated as long as the participants or time allows and until no new constructs are produced.

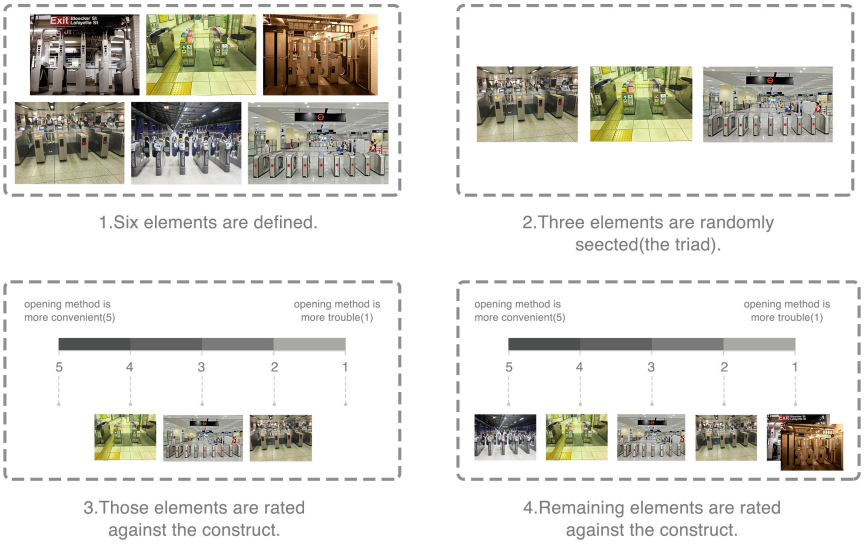


Fig. 1. The process of RGT

3.3 An Example Study Using RGT: User Experience Evaluation on Ticket Gate of Subway Station

Two questions were of interest in our study: first, how people feel and think about when they are interacting with the ticket gate of subway station; second, in what way we can evaluate the user experience of ticket gate of subway station. Interviews were made on three project days, and a total of 5 h of face-to-face interviews were organized in the café, located in College of Design and Innovation, Tongji University. Interviews were structured lasting from 15 to 30 min and researchers wrote notes and took photographs during and after the interview.

Participants. Altogether 11 respondents individually participated in the research. Five of them were female and six male. The median age was 25, ranging from 22 to 36. They all had experience using the ticket gate of subway station.

Procedure. A typical RGT interview starts with selecting elements. The elements are selected from a given domain and offered by the interviewees (researchers). In our example, we selected six different ticket gates from five countries. The name and picture of each element was written on a card shown to the respondents to stimulate responses. The group of six cards were then given to a respondent with the 'How' question. The question to be used to elicit the comparison that was repeated throughout the interview was established by the researchers before the interview process and designed to match the goal of the research. In our example study, which was focused broadly on evaluating the user experience of ticket gate of subway station, the question was, "Please think about how this particular ticket gate distinguishes itself from the others" or "How two of these ticket gates are similar to each other and different from the third in terms of your experience." This resulted in a pair of construct consisting of

a construct and its counterpole written on a form called a repertory grid. Each element was measured with specific number of bi-polar constructs, employed in semantic differential scales. Next step, the interviewee was asked to rate each element using the 5-point scale [17]. By asking the interviewee which attribute (the construct or the counterpole) he/she preferred, the preferred one was assigned a score of 5, which means that the element closely resemble the construct in the left hand column, while the counterpole scored 1, which means that this element closely matches the counter construct given on the right. Intermediate values suggested that the element was regarded as being somewhat between the extremes. Note that more than one element are possible to have the same score, but the score of 1 and 5 must be given in each row. The interviewer must maintain the consistency of the scale used during the interview. For example, in our study, the construct "Clean" was scored 5 point and the contrasting construct "Dirty" was scored 1 point. Next, the respondents needed to rate all remaining elements, which was a total of six in this case, on the same scale. As a result, the preferred elements will have higher average scores. In this example (Fig. 2), Shanghai has the highest average score, indicating that this interviewee thought Shanghai was the best overall product based on the constructs identified. In addition, the rating process requires the interviewees to fill in a score in the specific cell of the grid as immediately and spontaneously as possible.

	Tokyo	Shanghai	Hong Kong	New York	Paris	London	
Contrast pole (5)							**Contrast pole (1)**
Easy to use	4	4.5	4	2	1	4	Difficult
High quality	3	4	4	5	5	4	Bad quality
Affinity	5	5	3	3	3	4	Depression
Spacious and bright	3	5	3	4	3	5	Compact and dark
Good management	3	5	4	4	5	5	Bad public security
Clean	2	5	3	4	4	5	Dirty
Wide channel	3	5	3	3	3	5	Narrow channel
Soft materials	3	3	3	5	5	4	Hale materials
Materials-saving	4	5	4	2	2	5	Material-wasting
Modernization	2	5	4	4	3	5	Obsolescence
Intuitive visual guide	4	5	4	2	4	5	Chaotic visual guide
Average Score	3.11	4.68	3.55	3.45	3.45	4.64	**Average Score**

Fig. 2. An example of Repertory Grid

The information acquired from the interview was written into electrical form for further analysis and the data was analyzed by content analysis. If the interviewee could not provide a contrast to the construct immediately, the interviewer probably asked, "If this one is easy to use, how is this other one relatively dissimilar?" Presumably, the interviewees might explain the problems they faced with the ticket gate in Shanghai, providing more instructive insights. It is essential that the interviewer directly use the interviewee's own words in the Repertory Grid result, and in doing so, the interviewer's bias is eliminated. Once the results recorded completely in the repertory grid, a matrix made up of the elements and the constructs is formed [18]. This process repeated until the respondent could not identify any more new constructs, which

usually occurred after about 10 constructs have been identified. The whole process led respondents to uncover constructs which they might not have been consciously aware of before the interview. Each respondent independently contributed a facet to our understanding of user experience with ticket gate.

4 Results

4.1 Elements Selection

A Repertory Grid is a matrix consisting of rows and columns, in which elements are listed in each row, and constructs in columns. Generally, the grid should include 5–10 concrete elements. The elements should be representative, precise and homogeneous examples within that domain in that they could help the respondents to establish their personal constructs regarding the research topic. In this study, we singled out six ticket gates (Fig. 3) respectively from Shanghai, Hong Kong, Tokyo, Paris, New York and London, as these cities are all metropolis known for their developed and convenient public transportation. We presented to the respondents six A4-sized papers containing a picture of an existing ticket gate; the name of the city where the device was used; and the identification number.

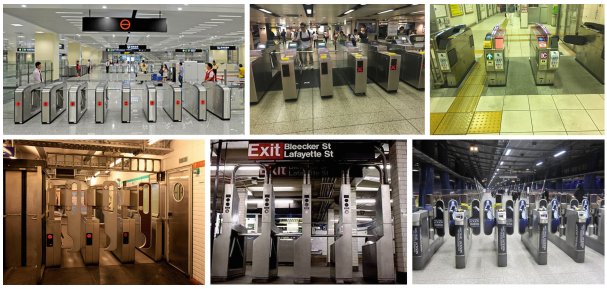


Fig. 3. Selected elements

4.2 Constructs Elicitation

The meaning of a construct illustrated by the respondent can be better understood when the opposite of the construct is provided with a brief description of the characteristic. For further discussion and analysis, as many relevant product features as possible should be taken into account. However, if a construct contradicted the rest, this construct should not be considered as a key factor [19]. Two kinds of data were collected through the use of RGT. First, personal constructs related to experience were elicited, which provides the study with qualitative data, i.e. respondents' own perceptions and preferences. Since each participant worked independently, the constructs were different from each other. A total of 71 pair of constructs was eventually formulated. Second, quantitative data was also gathered about the degree to which respondents thought the elicited constructs correspond to an element.

4.3 Analysis of Repertory Grid Data

Each Repertory Grid in our study consisted of multiple bipolar constructs, six singled out elements and a fixed rating system. There are various ways in which all participants' Repertory Grid can be synthetically analyzed. Every single grid could be summarized by calculating the average score of elements and constructs. To begin with, in order to statistically analyze how the participants perceive the given elements of the study as a whole, an average score for each pair of construct on all elements was calculated, as shown in Table 1. For instance, the average score for the constructs 'comfort vs. repression' is 3 point, which means the participant's overall evaluation of the six ticket gates in terms of the sense of comfort is moderate. Furthermore, the fact that content analysis entails the identification of content and context unit of analysis, say text or key word, has been pointed out by Holsti, the constructs are thus the foundation of analysis [20].

Table 1. The average score of each pair of construct (a sample Repertory Grid)

Contrast pole (5)	Tokyo	Shanghai	Hong Kong	New York	Paris	London	Contrast pole (1)	Average score
Convenient to open	5	1	1	1	1	5	Inconvenient to open	2.33
Warm feeling	5	2	1	2	3	5	Cold feeling	3.00
Similar color system	1	5	3	1	4	5	Multiple color system	3.60
Wide channel	5	1	1	1	2	5	Narrow channel	2.50
Simple structure	5	4	2	1	1	1	Complex structure	2.33
Materials-saving	4	4	1	1	1	1	Material-wasting	2.00
Feel safer in opening way	2	4	2	5	5	4	Feel dangerous in opening way	4.00
Comfort	5	1	3	1	3	5	Repression	3.00
Rational use of space	5	5	2	1	1	5	Irrational use of space	3.17
Harmonious with the environment	3	5	1	2	2	5	Not harmonious with the environment	3.00

To be able to analyze data between grids, a simple method to determine preferred elements is to calculate the average score for each element's construct scores individually, and then to average the set of average scores calculated from 11 Repertory Grids to finally get a cumulative average for the element. As seen in Table 2, the ticket gates of subway station in London received the highest accumulated average score whereas Paris got the lowest. Accordingly, the example of London ranked as the best ticket gate for its good user experience in our study.

Table 2. The accumulated average score of each element

	Tokyo	Shanghai	Hong Kong	New York	Paris	London
Average score	3.45	3.47	2.76	2.56	2.52	3.85

Some similar bipolar constructs appeared more than one time. To find out which of the constructs were most discussed, we counted the number of occurrences of each construct, for example, 'Intuitive visual guide vs. Chaotic Visual Guide' and 'Wide channel vs. Narrow channel' occurred in Repertory Grids for six times, which suggested that participants paid much attention to these attributes of ticket gate, making these two features key success factor. The quantitative outcomes of the study were not intended to provide a complete picture of user experience evaluation, but benefited to the development of a further qualitative analysis of the respondents' answers.

4.4 Interpretations and Presentation of the Result

Although Group-level analysis of repertory grid data brings a major challenge, researchers have found an analysis method to combine the data from several grids and uncover the valuable findings, that is, the constructs categorization. We used Hassenzahl's model of pragmatic and hedonic aspects of product qualities as a guiding theory in the analysis phase [7]. The perceptions of both pragmatic and hedonic qualities are subjective, in relation to a person's overall judgments of their experience. In our study, users described the pragmatic aspects with words like easy to use, learnable, effective, etc. Hedonic quality contributes to positive experience which stems from fulfilling a basic psychological need. Participants described the hedonic quality with words such as comfortable, attractive, intimate, frustrating, etc. These adjectives were obviously related to the user's own experience and expectation of using the ticket gate.

The process of categorizing was started with an interpretation of the semantical content of the constructs [14]. The constructs were firstly grouped into pragmatic or hedonic attribute, based on a connection between constructs through meaning. The constructs were then sorted into similar groups, and several labels were selected to characterize for the group as a whole, and reformat a new bipolar construct to represent the group. Next, we discussed the reformed results, teasing out the similarities and differences between different groups, and a couple of improved sub-category definitions were created. Lastly, each of us grouped all constructs into pragmatic and hedonic category and multiple sub-categories to test the reliability and validity of the categorization frame, as described above (Fig. 4). To further specify, in our case, 71 constructs were assigned to the pragmatic group or the hedonic group in the first step, and then separately sub-categorized the two groups into 9 and 10 improved category definitions. Obviously, 19 categories are easier to analyze than 71 individual constructs. This process contributed to summarizing the perceptions of all participants through capturing the data in a more accessible way [17]. The analysis result also showed which categories mattered most to interviewees. The count of how many constructs are in each category can highlight issues that are more important.

The qualitative analysis enabled us to develop a deeper understanding of the existing problems. This understanding together with insights obtained from earlier quantitatively analysis led towards an overall user experience evaluation of the ticket

Category	Sub-category	Definition	Numbers
Pragmatic	Visual Communication	The gate has intuitive visual information to guide passengers clearly./ The gate has chaotic visual information that easily misguide passengers.	11
	Usability	The gate is useful./ The gate is not useful.	13
	Materials	The kind of materials used in gate are appropriate./ The kind of materials used in gate are inappropriate.	4
	Color system	Different color of floor tiles can help passengers better identify the each channel of gate./Same color of floor tiles may not help passengers better identify the each channel of gate.	4
	Directness	Some informations or situations will be found directly./ Some informations or situations are hard to find.	6
	Management	The management is good in station./ The management is bad in station.	2
	Structure	The structure of gate is simple and reasonable./ The structure of gate is complex and unreasonable.	14
	Accessibility	The gate is accessible./ The gate is inaccessible.	9
	Quality	The quality of gate is high./ The quality of gate is bad.	1
Hedonic	Inclusiveness	The gate reflects the inclusive design./ The gate doesn't reflect the inclusive design.	6
	Sense of security	The gate has a stronger sense of security./ The gate has no sense of security.	2
	Attractiveness	Some informations or advertisements attract individuals./ Some informations or advertisements are uninviting.	12
	Sense of intimacy	The gate feels intimate./ The gate feels alienated.	20
	Harmony	The gate is in harmony with surrounding environment./ The gate is not in harmony with surrounding environment.	4
	Materials	The materials of gate are moderate and soft./ The materials of gate feel uncomfortable.	2
	Sense of alienation	The gate is alienated. / The gate is intimate.	2
	Shape	The shape is beautiful and moderate./ The shape is uncomfortable.	6
	People	There is very few people in station./ There are too many people in station.	1
	Sense of trust	The gate has a sense of trust to passengers./ The gate has no trust to passengers.	2

Fig. 4. Constructs Categorization

gate. It has already been revealed in Table 2 that London, Shanghai and Tokyo were ranked as the top 3 in our user experience evaluation study, but we had no clue as to why they were judged as the best and why most respondents disliked the example of Paris. Figure 5 provides an overall user experience evaluation of the six selected ticket gates, according to the nine most focused product characters (D1–D9) that has been identified through the analysis process mentioned above. From this table, we realized which characters were most closely and consistently related to the elements that the respondent judged as being the best or worst ticket gate of subway station. These results also provide empirical evidence that the evaluative judgments of the ticket gate are subjective and relative to the usage situations.

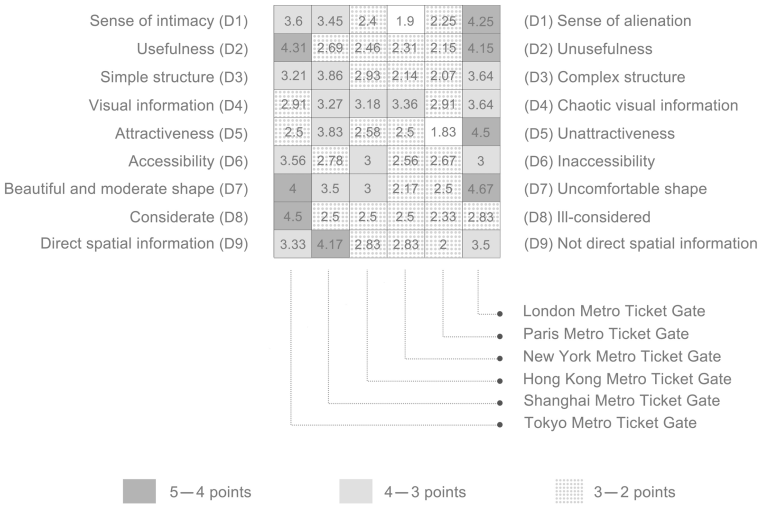


Fig. 5. An overall user experience evaluation on ticket gate of subway station

5 Discussion

What we were evaluating in our study is the experience of each individual. The essence of Repertory Grid Technique is in understanding how individuals view the given product, service, environment or system. Understanding the users' feelings and expectations rather than the product itself offers more possibilities. By the use of RGT, we not only gleaned valuable and novel insight from respondents into user experience of the ticket gate, but also the degree to which they thought the product quality matched the description articulated by themselves. Accordingly, RGT could be described as both qualitative and quantitative research [14]. The technique can elicit not only product attributes but also the problems they have with a product or service, as well as their feelings and emotions about it and the context in which it is used. When the technique was used to investigate the use experience of ticket gate in subway station, interviewees discussed specific attributes, but also mentioned their unhappy experience. The constructs elicitation session confirmed much of what had been known in previous work but also identified problems that had been not.

The interview process is not complicated but requires much attention to details. For example, when we conducted RGT interviews, we as interviewers probably used to interpret the interviewee's answer with our own language and sometimes we were prone to replace the respondents' understanding with our own terms. It is essential to put the interviewee in the first place through capturing the voice of them in their own terms, and in doing so, the risk of different participants interpreting questions in different ways is avoided. This method allowed participants to identify the attributes of the ticket gate that they considered as important and to provide their own meanings rather than being constrained by terms provided by a researcher.

A couple of limitations of the study have to be noted here. Firstly, RGT, to an extent, rely on how users perceive products instead of their actual use. Secondly, the results of RGT are associated with the given elements and personal constructs involved in the study, and therefore the results are intended to be relational instead of absolute [14]. Additionally, RGT is more applicable and effective for developing product improvements, rather than entirely new products. Another issue that can be seen as a limitation to the validity of RGT is that only a small number of interviews were conducted in our study, and the respondents were selected from a narrow section of the community. For future research, the sample should contain more members who may hold diverse values. RGT can compete with other user experience evaluation methods such as questionnaires, observation or interviews, for the data collection phase of RGT takes less time and effort. To further develop our RGT findings, we could make additional use of observations, which means RGT can be supplemented by other methods. In addition, further research could be done to provide a more detailed picture of user experience of supporting facilities of public transportation from a methodical as well as from a practical point of view.

6 Conclusion

Our study aimed to evaluate user experience of the existing ticket gate of subway station, but traditional usability testing methods are not adequate for doing it. Hence, there is an increasing need of developing systematic user experience evaluation methods. We started our research with a literature review on user experience. According to Hassenzahl's definition of user experience, the qualities of products can be grouped into two categories, pragmatic and hedonic quality [7]. Pragmatic quality related to the product's usability and utility, and hedonic quality related to users themselves. Based on Hassenzahl's model, we tentatively put the focus of UX evaluation on investigating users' subjective feelings of interaction with the product.

George Kelly's Personal Construct Theory assumes that people create a personal construct system of how the world works based on their experiences and examination of everything around them [21]. In addition, Kelly [12] proposed that people build their understanding of the world upon a multitude of bipolar constructs. To support the theory, the Repertory Grid Technique (RGT) has been taken as a useful technique for eliciting user's experiences of interacting with technological products. In this paper, we presented the results of exploring a user experience evaluation possibility by the usage of RGT to form an overall evaluative judgment on the ticket gate of subway station. We assumed that people perceive the ticket gate along two different dimensions, namely pragmatic and hedonic quality. An example study of RGT was used to learn how respondents feel and think about the selected ticket gates, through the process of elements selection, constructs elicitation, rating, data analysis and presentation.

To calculate the average score for each construct and element can yield notably important insights for user experience evaluation and product development. The results stemming from the process of constructs categorization is that in total 71 constructs were finally categorized into 19 sub-categories. Taking frequency count into consideration, the most focused constructs were identified, and which aspects of the ticket

gate would most affect users' overall evaluation were presented as well. To conclude, the example of London led the rankings, with the accumulated average score of 3.85, followed by Shanghai, Tokyo, Hong Kong, New York and Paris.

We found that with the aid of repertory grid respondents can articulate their feelings and impressions that occurred in the interaction with the ticket gate. This method allows the respondents in a very simple and immediate way to express how they judge user experience quality, using their own words and reflecting their own experience. Compared with other evaluation methods, the strength of RGT lies in its focus on user's subjective and personal feelings about the product. However, there are several limitations in the application of the method. The development of user experience evaluation thus still requires further research.

Acknowledgments. Many thanks go to every interviewee who took part in our research and everyone who offered helpful advice on ways to manage this study.

References

1. Fan, W., Yin, J.: Station Devices in Urban Railway Traffic, Chinese. China Railway Publishing House, Beijing (2000)
2. Qu, R., Bao, Y., Ren, C.: The study of event recognition technology of ticket gate in urban railway traffic. Sens. Actuators, A **134**(2), 641–649 (2007)
3. Law, L.C., Roto, V., Hassenzahl, M., Vermeeren, A.P.O.S., Kort, J.: Understanding, scoping and defining user experience: a survey approach. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 719–728. ACM, Boston (2009)
4. Kahneman, D.: Objective happiness. In: Kahneman, D., Diener, E., Schwarz, N. (eds.) Well-Being: The Foundations of Hedonic Quality, pp. 3–25. Sage, New York (1999)
5. Russell, J.A.: Core affect and the psychological construction of emotion. Psychol. Rev. **110** (1), 145–172 (2003)
6. Pucillo, F., Cascini, G.: A framework for user experience, needs and affordances. Des. Stud. **35**(2), 160–179 (2014)
7. Hassenzahl, M.: The thing and I: understanding the relationship between users and product. In: Blythe, M.A., Overbeeke, K., Monk, A.F., Wright, P.C. (eds.) Funology: From Usability to Enjoyment, pp. 31–42. Kluwer, The Netherlands (2003)
8. ENGAGE: Report on the evaluation of generative tools and methods for 'emotional design'. Deliverable D15.3. EU project Engage 520998 (2006)
9. Vermeeren, A.P.O.S., Law, L.C., Roto, V., Obrist, M., Hoonhout, J.: User experience evaluation methods: current state and development needs. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, pp. 521–530. ACM, Reykjavik (2010)
10. Obrist, M., Roto, V.: User experience evaluation: do you know which method to use? In: CHI 09 Extended Abstracts on Human Factors in Computing Systems, pp. 2763–2766. ACM (2009)
11. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (ed.) Symposium of the Austrian HCI and Usability Engineering Group, vol. 5298, pp. 63–76. Springer, Berlin, Heidelberg (2008)
12. Kelly, G.: The Psychology of Personal Constructs. Routledge, London (1955)

13. Goffin, K., Lemke, F.: Uncovering your customer's hidden needs. Eur. Bus. Forum **18** (Summer), 45–47 (2004)
14. Fallman, D., Waterworth, J.: Dealing with user experience and affective evaluation in HCI design: a repertory grid approach. Workshop Paper CHI. 2–7 (2005)
15. Szwejczewski, M., Goffin, K., Lemke, F., Pfeiffer, R., Lohmüller, B.: Supplier management in German manufacturing companies: an empirical investigation. Int. J. Phys. Distrib. Logist. Manag. **31**, 354–373 (2001)
16. Lemke, F., Goffin, K., Szwejczewski, M.: Investigating the meaning of supplier-manufacturer partnerships: an exploratory study. Int. J. Phys. Distrib. Logist. Manag. **33** (1), 12–35 (2003)
17. Baxter, D.I., Goffin, K., Szwejczewski, M.: The repertory grid technique as a customer insight method. Res. Technol. Manag. **57**(4), 35–42 (2014)
18. Easterby-Smith, M.: The design, analysis and interpretation of repertory grids. Int. J. Man Mach. Stud. **13**(1), 3–24 (1980)
19. Alexander, P., Loggerenberg, J.V., Lotriet, H., Phahlamohlaka, J.: The use of the repertory grid for collaboration and reflection in a research context. Group Decis. Negot. **19**(19), 479–504 (2010)
20. Holsti, O.R.: Content analysis. In: Lindzey, G., Aronson, E. (eds.) The Handbook of Social Psychology, vol. 2, pp. 596–692. Addison Wesley, London (1968)
21. Home, R., Bauer, N., Hunziker, M.: Constructing urban green spaces: an application of kelly's repertory grid. Tour. Rev. **62**(62), 47–52 (2007)

A Study on Quantitative Evaluation of Relationship Between Longitudinal Shock and Ride Comfort by Train Braking

Toshiro Miura(✉) and Toshiyuki Izumi

West Japan Railway Company, 2-4-24, Shibata, Kita-ku, Osaka-shi, Osaka-fu, Japan
{toshirou-miura,toshiyuki-izumi}@westjr.co.jp

Abstract. The purpose of this study is to quantitatively evaluate the relationship between a longitudinal shock during braking of a train and ride comfort. The authors attempted to develop a method of quantitatively evaluating the main braking patterns envisaged (40 in total) with a view to establishing an evaluation method that would be readily applicable to actual driving and control. A multiple regression model and a multinomial logistic model were derived for a one-step braking process and the usefulness of same was confirmed on the basis of the high value of the coefficient of determination. Based on the model formula/formulae derived and the results of evaluating the braking patterns, it was possible to evaluate the 40 braking patterns by means of an ordinal scale. Train driving/control is evaluated with respect to various aspects from various perspectives, but the development of a method of evaluating ride comfort, which was partly unsolved, means that it is now possible to comprehensively evaluate train driving. Future application of the evaluation method in improving train-driving skills and in automatic train-driving controls can be expected.

Keywords: Railway · Ride comfort · Driving · Longitudinal shock · Multiple regression model · Multinomial logistic model · Ordinal scale

1 Introduction

1.1 Background

Train driving is evaluated on the basis of various factors such as “on-time performance”, “stop precision”, “energy saving”, “ride comfort”, etc. Thanks to the advances in information technology in recent years, it has become possible to measure factors assessed based on physical quantities such as on-time performance, stop precision, energy saving, etc. and therefore, to evaluate such factors quantitatively. However, ride comfort is a psychological factor, which is evaluated based on the feelings/perceptions of passengers and is distinctive insofar as it cannot be evaluated directly from a measured physical quantity. Regarding this ride comfort, various studies have been done in the field of ergonomics [1, 2]. Among the shocks in the left-right, up-down, and longitudinal directions that affect the ride comfort [3–5], the longitudinal shock is largely influenced by train driving. Longitudinal shocks are mainly caused by powering and braking in train

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_95

driving, and passengers feel uncomfortable when braking. Omino et al. proposed a method for quantitatively evaluating ride comfort for longitudinal shocks during one-step braking [6–8]. In a train operated by a train driver, the train driver performs various braking operations as well as one-step braking. Therefore, in order to evaluate the ride comfort by train braking, it is necessary to clarify the relationship between longitudinal shocks and ride comfort in various brake patterns envisaged in actual driving maneuvers.

1.2 Purpose

The purpose of this study is to quantitatively evaluate the relationship between longitudinal impulses and ride comfort during braking. This study aimed to construct a method of evaluating the ride comfort/discomfort of various braking patterns by ordinal scale [9], as a level that could be used in the actual railway business. Such an evaluation method will make it possible not only to evaluate ride comfort but also to comprehensively evaluate train driving. Future application of the evaluation method in improving train-driving skills can be expected.

1.3 Study Structure

Figure 1 shows the scope and the investigative process of this study.

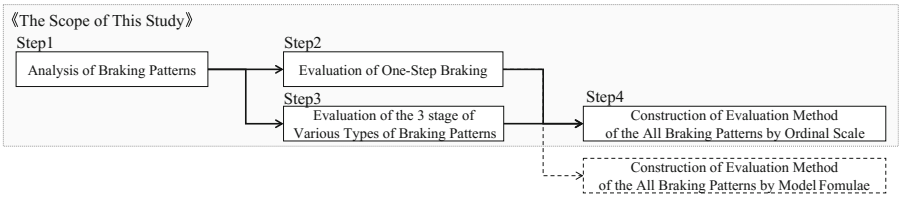


Fig. 1. The scope and investigation process of this study

2 Analysis of Braking Patterns

To develop an evaluation method that would be relevant to actual driving and control, it was important that the method take account of the various driving patterns seen in the actual field.

As shown in Fig. 2, the entire braking process from the start of braking to the stopping of the train was split into three "stages," namely "start of braking," "braking midway" and "just before stopping;" and then based on interviews with drivers and instructors, each stage was divided into the main conceivable braking patterns identifiable by distinctive properties. As shown in Fig. 3, that exercise made it possible to sort braking activity into two patterns in the start-of-braking stage (S1, S2), five patterns in the braking-midway stage (M1, M2, M3, M4, M5) and four patterns in the just-before-stopping stage (E1, E2, E3, E4). This gives 40 different combinations of braking patterns in the entire braking.

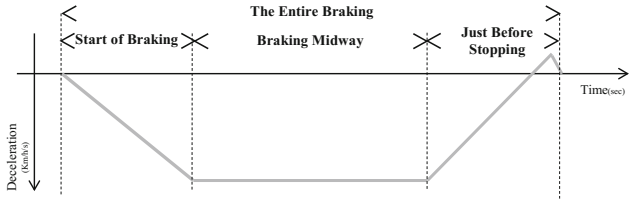


Fig. 2. The entire braking process and the relationships of each stage of braking

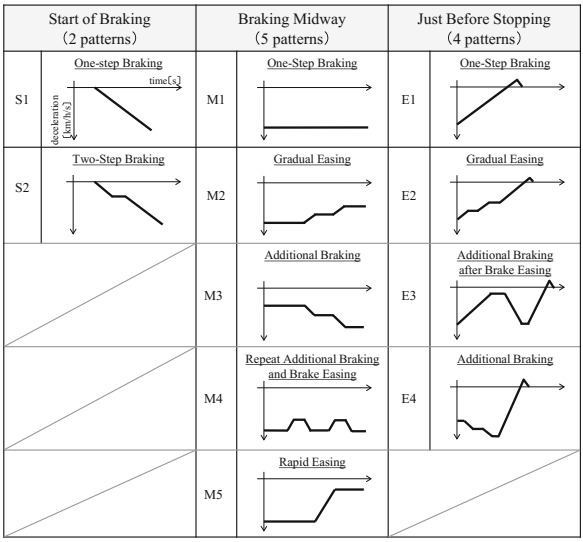


Fig. 3. The main braking patterns envisaged

3 Construction of an Evaluation Method

3.1 Experiment I: Evaluation of One-Step Braking

Using an experimental method, the authors analyzed the relationship between physical quantity and ride comfort assessments for single-step braking (S1, M1, E1) the simplest braking patterns, and studied the findings with a view to deriving model formulae.

Experimental Methodology. The experiment was conducted using suburban direct current electric trains (type of rolling stock: 223 series and 225 series) held by West Japan Railway Company.

An accelerometer (a built-in GPS train navigation accelerometer manufactured by KRE and an MTi type of high-performance vibration acceleration sensor manufactured by Xcens) was placed in the center of the carriage in which the subjects were boarded, and the deceleration rate was measured as shock data. Jerk values were calculated by differentiating the deceleration rate measurements [10, 11].

There was a total of 30 subjects, 15 male and 15 female subjects. The age ranges of both male and female subjects were as follows: aged 18–30 - 4 people, aged 31–40 - 3 people, aged 41–50 - 3 people and aged 51–60 - 4 people. All subjects were regular users of rail transport. The subjects were directed to stand so that their sides were positioned in the direction of travel [*i.e. facing either side of the train*] with their feet about shoulder-width apart.

The process used to obtain the subjects' ride comfort assessment value was to have the subjects fill in their assessment of each stage of braking and the entire braking. The ride comfort assessment value was set four-point rating scale [12]. The rating points were set as follows: "Rating value of 1 = absolutely no problem" "rating value of 2 = slightly uncomfortable" "rating value of 3 = uncomfortable but within tolerance" and "rating value of 4 = uncomfortable and unbearable" [12, 13]. Moreover, the subjects were also requested to maintain the original posture insofar as they could when assessing the ride comfort.

The conditions of the experiment were set as follows to simulate the circumstances of actual braking operations envisaged: seven patterns of initial braking speeds were set: 15 km/h, 25 km/h, 45 km/h, 60 km/h, 75 km/h, 100 km/h and 120 km/h. Moreover, five patterns of braking notches were set: 1 N, 2 N, 3 N, 4 N and 5 N for initial braking speeds up to and including 25 km/h, and 4 N, 5 N, 6 N, 7 N and 8 N (emergency brake) for initial braking speeds of 45 km/h and higher. Furthermore, two conditions were set for the just-before-stopping: the driver eased off the brakes, the driver did not ease off the brakes. Moreover, the experiment was conducted on a straight line and along a section with a gradient of ±5‰ to exclude factors other than longitudinal shock. The order in which the experiment was conducted was randomized to exclude order effects [12].

Experiment Results. The physical quantities shown in Fig. 4 were measured and calculated in the three stages of braking. The correlation between the average value of the ride comfort assessment values and the the physical quantities in the three stages is shown in Table 1. In the start-of-braking stage, the coefficient correlation of average value of C_{S1} (mean to the ride comfort assessment values in S1) with d_S (means to absolute value of deceleration) is 0.86, indicating a high correlation, and the coefficient correlation with j_S (means to absolute value of jerk) is 0.59, indicating a correlation. In the braking-midway stage, the coefficients correlation of average value of C_{M1} (mean to the ride comfort assessment values in M1) with $d_{M\text{-}ave}$ (means to average value of absolute value of deceleration) and $d_{M\text{-}max}$ (means to maximum value of absolute value of deceleration) are 0.85 and 0.84, respectively, indicating a high correlation for both. In the just-before-stopping stage, the coefficients correlation of average value of C_{E1} (mean to the ride comfort assessment values in E1) with d_E (means to absolute value of deceleration) and α_E (means to absolute value of swing-back-acceleration) are 0.81 and 0.76, respectively, indicating a high correlation for both quantities, and the coefficient correlation with j_S (means to absolute value of jerk) is 0.69, which shows a correlation exists.

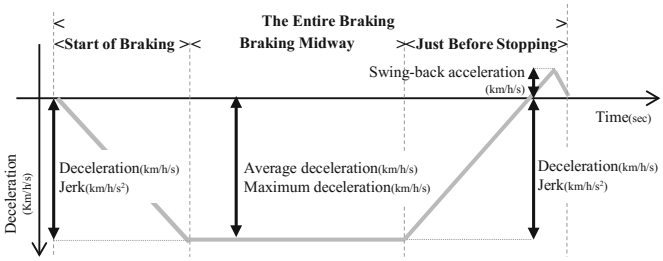


Fig. 4. The physical quantities measured/calculated in the three stages

Table 1. The coefficients correlation between the average value of the ride comfort assessment values and the physical quantities in the three stages of one-step braking

Start of braking		Braking midway		Just before stopping	
d_S (km/h/s)	0.86	$d_{M\text{-}ave}$ (km/h/s)	0.85	d_E (km/h/s)	0.81
j_S (km/h/s^2)	0.59	$d_{M\text{-}max}$ (km/h/s)	0.84	j_E (km/h/s^2)	0.69
				α_E (km/h/s)	0.76

Study of Model Formulae in the Three Stages of One-Step Braking. Model formulae were derived in relation to the relationship between ride comfort assessment values and the physical quantities in the three stages of one-step braking: a model formula derived from linear multiple regression analysis (hereinafter referred to as "the multiple regression model,") and another derived from a nonlinear multinomial logistic regression analysis (hereinafter referred to as "the multinomial logistic model.") The multiple regression model is a model formula derived by linear multiple regression analysis using each physical quantity in Table 1 as the explanatory variable/s and the average value of the ride comfort assessment values as the target variable. The multinomial logistic model is a model formula that calculates an expected value of ride comfort assessment (value) from the ride comfort assessment value and its estimated proportion P_n (n = 1, 2, 3, 4). The estimated proportion P_n (n = 1,2,3,4) is calculated from a regression equation derived by nonlinear multinomial logistic regression analysis using the absolute values of the physical quantities in Table 1 as the explanatory variable/s and the ride comfort assessment values as the target variable.

Moreover, since a correlation is deemed to exist between ride comfort assessment values and physical quantities having equal intervals, for the purposes of this study only it was deemed reasonable to handle the ride comfort assessment values like an interval scale [9]; and the ride comfort assessment values were applied in the linear multiple regression analysis and used in the derivation of the multinomial logistic model to calculate the expected values.

Formulas (1), (2) and (3) is the multiple regression model.

$$C_{S1} = 0.37d_S + 0.13j_S + 0.02. \tag{1}$$

$$C_{M1} = 0.26d_{M\text{-}ave} + 0.16d_{M\text{-}max} + 0.06. \tag{2}$$

$$C_{E1} = 0.25d_E + 0.01j_E + 0.29\alpha_E + 0.34. \tag{3}$$

Formulas (4), (5) and (6) is the multinomial logistic model.

$$C_{S1} = 1\,P_{S1} + 2P_{S2} + 3P_{S3} + 4P_{S4},$$

$$P_{Sn} = \begin{cases} \dfrac{e^{a_{Sn}\,d_s + b_{Sn}j_s + c_{Sn}}}{1 + \sum_{i=1}^{3} e^{a_{Si}d_s + b_{Si}j_s + c_{Si}}}, & n = 1, 2, 3, \\ \dfrac{1}{1 + \sum_{i=1}^{3} e^{a_{Si}d_s + b_{Si}j_s + c_{Si}}}, & n = 4, \end{cases} \tag{4}$$

$$a_{S1} = -2.22,\ b_{S1} = -0.52,\ c_{S1} = 14.1,\ a_{S2} = -1.81, b_{S2} = -0.27,$$
$$c_{S2} = 11.90,\ a_{S3} = -1.04,\ b_{S3} = -0.23,\ c_{S3} = 7.75$$

$$C_{M1} = 1P_{M1} + 2P_{M2} + 3P_{M3} + 4P_{M4},$$

$$P_{Mn} = \begin{cases} \dfrac{e^{a_{Mn}\,d_{M-ave} + b_{Mn}d_{M-}\max + c_{Mn}}}{1 + \sum_{i=1}^{3} e^{a_i\,d_{M-ave} + b_i d_{M-}\max + c_i}}, & n = 1, 2, 3, \\ \dfrac{1}{1 + \sum_{i=1}^{3} e^{a_i\,d_{M-ave} + b_i d_{M-}\max + c_i}}, & n = 4, \end{cases} \tag{5}$$

$$a_{M1} = -0.87,\ b_{M1} = -1.15,\ c_{M1} = 11.81,\ a_{M2} = -0.16, b_{M2} = -0.88,$$
$$c_{M2} = 7.83, a_{M3} = -0.06,\ b_{M3} = -0.49,\ c_{M3} = 3.97.$$

$$C_{E1} = 1P_{E1} + 2P_{E2} + 3P_{E3} + 4P_{E4},$$

$$P_{En} = \begin{cases} \dfrac{e^{a_{En}\,d_E + b_{En}j_E + c_{En}\alpha_E + d_{En}}}{1 + \sum_{i=1}^{3} e^{a_{Ei}\,d_E + b_{Ei}j_E + c_{Ei}\alpha_E + d_{Ei}}}, & n = 1, 2, 3, \\ \dfrac{1}{1 + \sum_{i=1}^{3} e^{a_{Ei}\,d_E + b_{Ei}j_E + c_{Ei}\alpha_E + d_{Ei}}}, & n = 4, \end{cases} \tag{6}$$

$$a_{E1} = -1.41,\ b_{E1} = -0.07, c_{E1} = -1.30, d_{E1} = 10.94, a_{E2} = -0.94,$$
$$b_{E1} = -0.03, c_{E2} = -0.38, d_{E2} = 7.42$$
$$a_{E3} = -0.41,\ b_{E3} = -0.0004, c_{E3} = 0.22, d_{E3} = 3.53.$$

As shown in Table 1, it was confirmed that the coefficients of determination of formulas (1) to (6) are high, all being greater than 0.5. Accordingly, based on the study results above, the multiple regression models and the multinomial logistic models in the three stages of braking, can be described as useful models with which it is possible to quantitatively evaluate ride comfort.

Study of Model Formulae for One-Step Braking in the Entire Braking. A multiple regression model and a multinomial logistic model were derived in relation to the relationship between C_{all} (means to the ride comfort assessment values in the entire braking) and the physical quantities in the entire braking.

The multiple regression model is a model formula derived by linear multiple regression analysis using the average value of the ride comfort assessment values for each stage as the explanatory variable/s and the average value of the ride comfort assessment values for the entire braking process as the target variable. The multinomial logistic model is a model formula that calculates an expected value of ride comfort assessment (value) from the ride comfort assessment values and their estimated proportions P_n ($n = 1, 2, 3, 4$). The estimated proportion P_n ($n = 1, 2, 3, 4$) is calculated from a regression equation derived by nonlinear multinomial logistic regression analysis using the ride comfort assessment value for each stage as the explanatory variables and the ride comfort assessment value for braking as a whole as the target variable.

Formula (7) is the multiple regression model.

$$C_{all} = 0.30E_{S1} + 0.31E_{M1} + 0.61E_{E1} + 0.02. \tag{7}$$

Formula (8) is the multinomial logistic model.

$$\begin{aligned} &C_{all} = 1P_{all1} + 2P_{all2} + 3P_{all3} + 4P_{all4}, \\ &P_{alln} = \begin{cases} \dfrac{e^{a_n E_{S1} + b_n E_{M1} + c_n E_{E1} + d_n}}{1 + \sum_{i=1}^{3} e^{a_n E_{S1} + b_n E_{M1} + c_n E_{E1} + d_n}}, & n = 1, 2, 3. \\ \dfrac{1}{1 + \sum_{i=1}^{3} e^{a_n E_{S1} + b_n E_{M1} + c_n E_{E1} + d_n}}, & n = 4. \end{cases} \\ &a_1 = -3.28, b_1 = -3.85, c_1 = 6.05, d_1 = 27.97, a_2 = -1.33, \\ &b_2 = -1.81, c_2 = -3.13, d_2 = 17.00 \\ &a_3 = -0.16, b_3 = -0.85, c_3 = -1.33, d_3 = 7.55 \end{aligned} \tag{8}$$

As shown in Table 2, it was confirmed that the coefficients of determination of formulae (7) and (8) are very high, both being greater than or equal to 0.9. Accordingly, based on the study results above, the multiple regression model and the multinomial logistic model for the entire braking process, can be described as useful models with which it is possible to quantitatively evaluate ride comfort. Moreover, it can be confirmed from the coefficients of the explanatory variables in formula (7) that the ride comfort assessment value of the just-before-stopping stage has approximately double the effect on the ride comfort assessment value for the braking process as a whole, comparing with the ride comfort assessment values for the start-of-braking and braking-midway stages. This suggests that the operation of the brakes in the just-before-stopping stage is particularly important from the perspective of ride comfort.

Table 2. The coefficients of determination of the multiple regression model and the multinomial logistic model in each of the three stages of one-step braking and in the entire braking process

	Start of braking		Braking midway		Just before stopping		The entire braking	
Multiple regression model	Formula (1)	0.78	Formula (2)	0.76	Formula (3)	0.77	Formula (7)	0.96
Multinomial logistic model	Formula (4)	0.72	Formula (5)	0.60	Formula (6)	0.76	Formula (8)	0.96

3.2 Experiment II: Evaluation of the Three Stages of Various Braking Patterns

Using an experimental method, the authors analyzed the relationship between physical quantity in the three stages of various braking patterns and ride comfort assessments, and studied an evaluation method based on an ordinal scale.

Experimental Methodology. Apart from the experimental conditions, a methodology similar to that used in Experiment was used.

The experimental conditions were as follows: The initial braking speeds were set at 75 km/h and 90 km/h, and the braking notches were set for each braking pattern as laid out in Table 3. Furthermore, two conditions were set for the just-before-stopping stage: the driver eased off the brakes, the driver did not ease off the brakes. Moreover, to prevent subjects from rating the wrong braking stages, the subjects were informed of a switch in braking stage just before the transition to the new stage.

Table 3. The experiment conditions for each braking pattern

Start of braking		Braking midway		Just before stopping	
S1	5 N	M1	5 N	E1	5 N
S2	2 N → (1 s after brake operation) → 5 N	M2	5 N → (30 km/h) → 4 N →(20 km/h) → 3 N	E2	5 N → (10 km/h) → 4 N →(5 km/h) → 3 N
		M3	5 N → (40 km/h) → 6 N →(30 km/h) → 7 N	E3	5 N → (10 km/h) → 4 N →(5 km/h) → 3 N
		M4	5 N → (50 km/h) → 3 N →(40 km/h) → 5 N →(30 km/h) → 3 N →(20 km/h) → 5 N	E4	5 N → (10 km/h) → 4 N →(5 km/h) → 3 N
		M5	5 N → (30 km/h) → 1 N		

Experiment Results. Ride comfort assessment values of the three stages in various braking patterns were compared. Ride comfort assessment values of the start-of-braking stage under conditions whereby there was no difference in the deceleration rate S1 and deceleration rate S2 in Fig. 5, were compared. The result of that comparison was that C_{S2} (means to the ride comfort assessment value in S2) was lower than C_{S1}, and the difference was found to be significant ($p < 0.01$). Ride comfort assessment values of the braking-midway stage under conditions whereby there was no difference in the deceleration rates M1, M2, M3, M4 and M5 in Fig. 5, were compared.

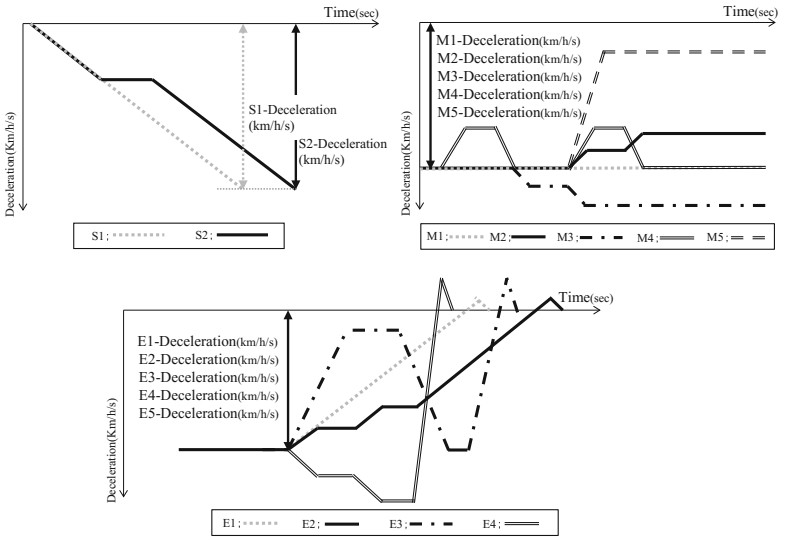


Fig. 5. Comparison conditions of Braking Patterns (S1 and S2 in start-of-braking stage, M1 to M5 in braking-midway stage, E1 to E4 in just-before-stopping stage)

The result of that comparison was that no significant difference was found between C_{M2} (means to the ride comfort assessment value in M2), C_{M5} (means to the ride comfort assessment value in M5) and C_{M1}. C_{M3} (means to the ride comfort assessment value in M3) and C_{M4} (means to the ride comfort assessment value in M4) were higher than C_{M1}, and the difference was found to be significant ($p < 0.01$). Next, C_{M3} was higher than that for C_{M4}, and the difference was found to be significant ($p < 0.01$). Ride comfort assessment values of the just-before-stopping stage under conditions whereby there was no difference in the deceleration rates E1, E2, E3 and E4 in Fig. 5, were compared. The result of that comparison was that no significant difference was found between C_{E2} (means to the ride comfort assessment value in E2) and C_{E1}, C_{E3} and C_{E4} were higher than C_{E1}, and the difference was found to be significant ($p < 0.01$). Then, The result of that comparison was that no significant difference was found between C_{E3} and C_{E4}.

Study of an Evaluation Method Based on an Ordinal Scale. If the ride comfort assessments of the three stages of various braking patterns are evaluated by means of an ordinal scale, in the order of ascending (*worsening*) ride comfort, the result is as follows: The existence or lack of statistical significance was confirmed for every rating, and therefore one can say that the ordinal scale is a useful tool.

- Start-of-braking stage: $C_{S2} < C_{S1}$
- Braking-midway stage: $C_{M1} = C_{M2} = C_{M5} < C_{M4} < C_{M3}$
- Just-before-Stopping stage: $C_{E1} = C_{E2} < C_{E3} = C_{E4}$.

3.3 Study of an Evaluation Method Based on an Ordinal Scale, for All 40 Braking Patterns

The authors studied to construct a method of evaluating the 40 braking patterns by means of an ordinal scale, from the results of Experiment I and Experiment II.

In developing the evaluation method, it was necessary to investigate the effect/s of the combination/s of braking patterns. The authors investigated whether or not the braking patterns in the start-of-braking stage affected the ride comfort assessment values of the braking patterns in the breaking-midway stage. Since there is no difference between S1-deceleration and S2-deceleration in Fig. 5, no difference attributable to a difference in the braking patterns at the start of braking will arise between M1-deceleration to M5-deceleration in Fig. 5. It can be said that a difference in the braking pattern in the start-of-braking stage does not affect the ride comfort assessment values of the braking-midway stage.

Next, the authors investigated whether or not the braking patterns in the braking-midway stage affected the ride comfort assessment values of the just-before-stopping stage. It was noticed that differences arose between E1-deceleration to E4-deceleration in Fig. 5 due to differences in the braking patterns in the braking-midway stage. It can be seen from formulae (3) and (6) that C_{E1} varies depending on the deceleration rate. Hence, a difference attributable to a difference in the braking pattern in the braking-midway stage arises, and the assessment value can be said to be in the order of $M5 < M2 < M1 = M4 < M3$, in ascending (*worsening*) order. It could be confirmed that there was a correlation between average value of C_{E2} and average value of E2-deceleration, the correlation coefficient being 0.60. Hence, a difference attributable to a difference in the braking pattern in the braking-midway stage arises, and the assessment value can be said to be in the order of $M5 < M2 < M1 = M4 < M3$, in ascending (*worsening*) order. It could be confirmed that the correlation between average value of C_{E3} and average value of E3-deceleration was very low, the correlation coefficient being 0.04. Hence, a difference attributable to a difference in the braking pattern in the braking-midway stage is thought not to arise. It could be confirmed that there was a high correlation between average value of C_{E4} and average value of E4-deceleration, the correlation coefficient being 0.78. Hence, a difference attributable to a difference in the braking pattern in the braking-midway stage arises, and the assessment value can be said to be in the order of $M5 < M2 < M1 = M4 < M3$, in ascending (*worsening*) order.

Based on the above findings, the one-step braking patterns (S1, M1, E1) were set as "0," and points were allocated using a system with intervals of "1" and positive integer/s for low (*good*) ride comfort assessment values and negative integers for high (bad) ride comfort assessment values as set forth in Table 4. Based on the results of formula (7), each of the points in Table 5 was weighted according to the weighting: start-of-braking stage: braking-midway stage: just-before-stopping stage = 1:1:2. The ride comfort assessment values for all 40 braking patterns were ranked based on the totals of the weighted points. The result of that ranking exercise is shown in Table 5. The rank-correlation coefficient (Spearman's) between the ranking in Table 5 and the ranking in Experiment II is 0.75. It could thus be confirmed that the correlation between the two is high. Therefore, the evaluation method based on the ordinal scale for all 40 braking patterns can be described as a useful scale.

Table 4. The points of the ride comfort assessment valeus in the three stages of various braking patterns (taking into account the effect of the combination of braking patterns)

Start of braking		Braking midway		Just before stopping		
Braking pattern	Points	Braking pattern	Points	Braking pattern (precondition)		Points
S1	0	M1	0	E1	(M1, M4)	0
					(M2)	1
					(M3)	−1
					(M5)	2
S2	1	M2	0	E2	(M1, M4)	0
					(M2)	1
					(M3)	−1
					(M5)	2
		M3	−2	E3	(all)	1
		M4	−1	E4	(M1, M4)	−1
					(M2)	0
					(M3)	−2
					(M5)	1
		M5	0			

Table 5. The ranking of all 40 braking patterns

Ranking	Braking pattern			Ranking	Braking pattern		
1	S2	M5	E1	7	S1	M4	E1
	S2	M5	E2		S1	M4	E2
2	S1	M5	E1		S2	M1	E3
	S1	M5	E2		S2	M1	E4
3	S2	M2	E1		S2	M2	E3
	S2	M2	E2	8	S1	M1	E3
	S2	M5	E3		S1	M1	E4
	S2	M5	E4		S1	M2	E3
4	S1	M2	E1		S2	M4	E3
	S1	M2	E2		S2	M4	E4
	S1	M5	E3	9	S1	M4	E3
	S1	M5	E4		S1	M4	E4
5	S2	M1	E1		S2	M3	E1
	S2	M1	E2		S2	M3	E2
	S2	M2	E4		S2	M3	E3
6	S1	M1	E1	10	S1	M3	E1
	S1	M1	E2		S1	M3	E2
	S1	M2	E4		S1	M3	E3
	S2	M4	E1	11	S2	M3	E4
	S2	M4	E2	12	S1	M3	E4

4 Conclusion

This study explored the construction of a method of evaluating ride comfort using an ordinal scale, applicable to all 40 braking patterns elucidated by an analysis of braking patterns. In Experiment I, multiple regression models and multinomial logistic models were derived for one-step braking patterns, and the high coefficients of determination showed the models to be useful models. In Experiment II, a method of evaluating the three stages of braking of various braking patterns by means of an ordinal scale was developed, and the method was shown to be useful on the basis of the existence or lack of statistical significance. A method of evaluating all 40 braking patterns by means of an ordinal scale (Table 5) was developed from the results of Experiments I and II, and its usefulness elucidated by the high rank-correlation coefficient calculated. Moreover, a method of evaluating the ride comfort during all 40 braking patterns by means of a model formula was beyond the scope of the present study and therefore was not investigated.

As a railway company, the result of this study could be usefully deployed in actual operations. There are plans to develop an evaluation tool for carrying out a general evaluation of train driving, and to advance efforts to use the tool for the improvement of train-driving skills.

References

1. Suzuki, H.: Train ride and comfort. J. Meteorol. Soc. Jpn. **35**(6) (2000)
2. Powell, J.P., Palacin, R.: Passenger stability within moving railway vehicles: limits on maximum longitudinal acceleration. Urban Railw. Transit. **1**(2), 95–103 (2015)
3. Enomoto, M., Sasaki, K., Shiroto, H.: Aiming at improving ride comfort. Regen. Reconstr. Restor. **67**(3) (2010)
4. Takai, H.: Changes in evaluation method of ride comfortableness. RTRI Rep. **9**(8), 61–66 (1995)
5. Miyoshi, K.: The study background of ride comfort control standards. J. Railw. Eng. Study **38**(3), 127–134 (1981)
6. Omino, K., Shiroto, H., Endoh, H., Seino, H.: A study of braking patterns in terms of ride-comfort. Jpn. J. Ergon. **42**(3) (2006)
7. Endoh, H.: Exploring optimal braking patterns from the perspectives of safety and ride comfort. Regen. Reconstr. Restor. **71**(8) (2014)
8. Shiroto, H., Fujinami, K., Omino, K.: Appropriate assessment of deceleration practice in the train operation. RTRI Rep. **8**(12), 43–46 (1994)
9. Stevens, S.S.: On the theory of scale of measurement. Science **3**(2684), 677–680 (1946)
10. Suzuki, H., Shiroto, H.: Method for implementing a test to evaluate ride comfort on the basis of sensations. RTRI Rep. **17**(1) (2003)
11. Masanobu, N.: Brake control aiming to improve ride comfort. Regen. Reconstr. Restor. **57**(9), 16–19 (2000)
12. Omino, K.: Examination of ride comfort of brake. Regen. Reconstr. Restor. **59**(6), 8–11 (2002)
13. Suzuki, H.: Railway vehicles and ride. In: Human Measurement Handbook, 4.4. Asakura Publishing (2003)

Guidelines for Electronic Systems Designed for Aiding the Visually Impaired People in Metro Networks

Eliete Mariani[1(✉)] and Marcelo Eduardo Giacaglia[2]

[1] The São Paulo Metropolitan Company, Operation Management, Rua Vergueiro 1200, São Paulo, SP 01504000, Brazil
eliete.mariani@gmail.com
[2] Faculty of Architecture and Urbanism, University of São Paulo, Rua Do Lago 876, São Paulo, SP 05508080, Brazil
mgiacagl@usp.br

Abstract. The research described herein sought to contribute to the theoretical field pertaining assistive technology systems for the visual impaired. It aimed to investigate by what means the visual impaired interact within the metro environment, seeking to understand their abilities, limitations and fears, in light of their cognitive variables. The study was conducted through participant observation of users with visual impairment and also survey questionnaires to specialists, instructors of orientation and mobility for the visual impaired, and members of the focused users, in São Paulo (Brazil), and Porto (Portugal), which include, respectively, the assessment of two electronic guidance systems, NavGATe and Navmetro. It resulted in a set of guidelines for design, implementation and operation of mobile based assistive systems for the visual impaired within metro networks that fill existing gaps in current knowledge.

Keywords: Visually impaired · Spatial awareness · Navigation · Metro

1 Introduction

Public spaces in general, and transportation and information systems in particular, are still generally inaccessible to the visual impaired, despite this being the most common personal disability worldwide. In Brazil, the last National Census [1] verified that 18.8% of the population have some degree of visual impairment, while motor, hearing and mental/intellectual impairment account less, respectively 7.0%, 5.1% and 1.4%. Visual impaired persons have greater difficulty in understanding large public spaces, because of the lack of visual references commonly used by sighted people [2].

Metro networks account for the commute of millions of passengers daily. Metro stations are designed as transition spaces within such systems. In order to cope with such high flows and gatherings of passengers, intersecting service lines and other urban infrastructures, stations become non-standard complex buildings. Flow of passengers vary, not only, daily, but over time, as there are changes to the network. Also, different metro car dimensions, layouts, and door mechanisms makes metro networks confusing to the autonomous movement of the visual impaired.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_96

Adaptation of existing metro networks as well as making of accessible new stations and lines is a major challenge, as verified in the day to day work of technical and customer support teams of the São Paulo Metropolitan Company (CMSP) [3]. In Brazil, despite the adjustments made, as dictated by legislation [4] and norms [5, 6], one notes that many visual impaired users feel unsafe and require the aid of personnel while moving within metro stations [7]. Current specifications comprise basically, installation of tactile mats, Braille plates or raised lettering as alternatives to the existing visual cues.

In recent years, visual impaired users have asked metro companies to provide electronic means of guidance. Although there are no official records of such claims, they can be easily related to widespread use of the Internet, accessible software, and availability of mobile phones. Such technology has enabled the visual impaired to access information and to interact with the rest of the world, greatly reducing their dependency on other people, allowing them to experiment a desired freedom.

Such state of things was the motivation for the research project [8] described herein, that investigated the means the visual impaired interact with the environment while travelling inside metro systems. The study involved the assessment of the São Paulo (Brazil) and Porto (Portugal) metro systems. The Porto metro system assessment included a review of Navmetro, an electronic information and navigation system for the visual impaired, in operation since 2009 [9].

This paper also includes the continuity of the research, mainly the assessment of NavGATe [10], a proposed orientation and navigation system for the visually impaired, under evaluation by CMSP.

2 Method

Use of a transportation system is a socially organized activity, in which, real behaviour can differ from stated preference. According to Flick [11], qualitative research is best indicated to interpret and identify user needs, not acknowledged by themselves, and to identify peculiarities not covered in academic publications. Also, Creswell [12] affirms that qualitative research provides for the investigation of complex problems, also for situations involving the visual impaired. And, according to Moresi [13], qualitative research must be used when seeking to understand in detail why an individual exhibits a particular behaviour. Therefore a qualitative approach was adopted in this study. This was also influenced by Angrosino [14] who recommends such an approach when the experiments occur in their natural context, which was the case.

This research was oriented by instruments observation and interviews, of qualitative research, favorably used by Cohen [15], Valentini [16], Queiroz [17] and Barbosa [18].

In face of the difficulty in finding specific significant data compilations, research on the problem domain sought the multidisciplinary information specific to researchers and other concerned on the subject. It was necessary to build a cognitive map of the targeted users, identify the relevant characteristics of metro networks, and investigate the available technology, of interest to orientation and navigation for the visually impaired.

Descriptive analysis was made by systematic reading of the data to seek for emerging behavioral patterns, as proposed by Angrosino [14], so as for them to be organized in categories for the analysis.

Content analysis was done by reading the transcriptions of spoken interviews, testimonials and observation protocols. This material is summarized after treatment and classification, to guide the selection of the relevant information in the recorded transcripts, as indicated by Gerhardt and Silveira [19]. A similar procedure was applied by Rose [20] to contribute to the categorization of results, from moving images. The analysis of the interviews considered the interpretive context, since, according to Gill [21], all discourse is circumstantial.

3 Previous Work

Previous work considered are Hatwell [22]; Bradley and Dunlop [23]; Taylor, Peplau and Sears [24]; Bins Ely [25]; Rogers, Sharp and Preece [26]; WHO [27]. Based on these studies it was identified that visually impaired people exhibit particular corporal structure. They organize their world according to the ways they walk, their routes and its access points are configured by their bodies in movement.

In order to orient themselves and move in a quick, efficient and independent manner, the visual impaired rely on a cognitive process that includes perception, coding, learning and environmental information memory [28]. However, ACAPO [29] observes that 70 to 80% of the visual impaired are not completely blind, and, according to Harper [30] and Honorato [31], this influences the way they interact with public circulation spaces. According to WHO [27] and Lora [32], despite using other means of perception, like aural, haptic, kinesthetic, muscular memory, vestibular (or labyrinthine) and/or olfactive, people tend to use as much as possible whatever visual capability they possess.

In order to learn how to orientate themselves without the visual sense special training is necessary and courses exist, on using the remaining senses, through systematic learning techniques. These are referred to as Orientation and Mobility (OM) [33]. In addition, Almeida [34] proposes the use of audio guides, especially in the case of congenital visual impairment, as many cannot pinpoint their location based spatial information, e.g. tactile maps.

Notable experiences on the design and use of electronic guidance systems, especially for indoor environments, are Crandall, Bentzen, Myers and Mitchell [35]; Du Buf et al. [36]; Almeida, Orduña and Castillejo [37]; López-De-Ipiña et al. [38]; Virtanen and Koskinen [39]. Although prototypes may have been produced and tested, the great majority of these devices did not reach production. Hersch and Johnson [40] speculate on the possible reasons for this: (a) excessive complexity of operation allied to insufficient training; (b) odd-looking device; (c) cumbersome or too heavy to carry; (d) fails to address the user's needs or even; (e) less useful than the walking stick.

The key aspect regarding safety and reliability is to secure a user's next steps. Blindness or reduced vision makes all spaces hazardous to some degree. The issue for

any electronic guidance system is the instantaneous determination a user's position, as emphasized by Hersh and Johnson [40]. According to them, this is a complex process because: (a) references are to be acquired through sensors, and this can involve communication between devices, possibly of differing technology; (b) references are subject to changes, i.e., new construction and placement of objects that become interventions to the original design; and (c) interpretation of reference information data can be quite complex.

The majority of navigation devices studied are based on radio waves. However, radio triangulation errors can be of several meters, according to Furukawa and Bruno [41]; and Pereira [42], and this is unacceptable regarding a metro system.

In what regards comfort, devices must carried with ease, most importantly, they mustn't require permanent handling during use, e.g. holding in one hand, considering that the other is already in charge of the walking stick. One hand should be free for other interactions, as recalls ONCE [43].

For studies on navigation, those of Harper [30] and Hersh and Johnson [40] stand out. Harper's findings, despite almost twenty years, are still up to date, regarding navigation, because they refer to the natural behavior of people; and the scarcity of newer references.

The contemporary meaning for navigation, used also as the determination of the route through terrestrial paths and built environment was explored in [8] with significant importance, as, according to Rogers, Sharp and Preece [26], the ways in which a person is oriented is crucial to a well succeeded trip. In this sense, navigation comprises a few simple actions: (a) route planning; (b) orientation based on reference points; (c) orientation during dislocation; and (d) detecting/avoiding obstacles and hazards.

Yatani et al. [44] verified that the most effective communication interface is one that combines audible and tactile (or haptic) information, but that further studies are necessary to determine communication style preferences before actual usage.

Almeida [34], based on Harper [30], summarized some concepts related to non-visual navigation, two of which are noteworthy: (a) a route to be followed will require more segments in relation to what would be described for sighted people; (b) despite the necessity for more information, they must simple and objective.

Hersh and Johnson [40] also state that a navigation system should provide information in all situations, both normal and abnormal or in face of an emergency.

Gedawy [45] studied navigation instructions based on necessities of the visually impaired, and emphasizes that some instructions can be stated in different ways, according to user preference. Booch et al. [46] recommend that such systems be designed to adapt to user expertise.

Rogers, Sharp and Preece [26, p. 85] recommend that "provide additional hidden information that is easy to access for users who wish to understand more about how to carry out an activity more effectively". And also that menus and voice instructions be minimal, as people find it hard to deal with menus having more than three or four options, as well as remembering multipart instructions.

4 Development

The research was initiated by the study of the perceptive experience of the target users. In this first phase, that culminated in [8], defined the objectives: (a) identify and analyse the senses and references used by the visually impaired when moving in metro stations and trains; (b) understand how they perceive space, train arrivals and departures, and other people, in order to orientate themselves; (c) observe user experience in the use of a orientation and navigation system when moving inside a metro station.

The instruments used in the first phase were observation and interviews, organized as follows:

(a) São Paulo (Brazil) metro, from September to October, 2015

- Individual observation, without researcher interference, of six frequent users, with different types of visual impairment, going through different stations, on their usual routes to work and/or education;
- Individual interviews, semi-structured and face-to-face, with eleven operational or station design staff of CMSP, and two orientation and mobility instructors;

(b) Porto (Portugal) metro, in March, 2015

- Group interview with two architects, employees of the *Metro do Porto* (Porto metro company) and developers of the Navmetro orientation and navigation system; and an individual interview with its project leader, also lecturer at FEUP – Faculty of Engineering of the University of Porto;
- Observation, without researcher interference, of a frequent user on his usual route, oriented by the Navmetro system inside a station.

Navmetro is a free orientation and guidance system for use in the Porto metro, which provides information such as lines, schedules and fares, through the use of voice commands on a regular mobile phone. Orientation is provided separately through bird sounds emitted at different points inside stations.

Observations compiled [8] were of fundamental importance to the understanding of the ways visually impaired people: do to know metro itineraries, locate entrances and exits; manage to reach a station not known to them and perceive their spatial location, among others, for the development of electronic guidance systems.

The second phase of the research, occurred from June to August, 2016, made use of the instruments of observation and interview of seven voluntary participants for testing NavGATe [10], an integrated system in which users carry receivers of information transmitted to them from devices distributed inside stations. The system provides exact user location and directions to follow. It consists of several transmitters conveniently placed inside stations, marked as (A) on Fig. 1. A receiver provided to each user, marked as (B) on Fig. 1, captures the location or other information, in the form of short range light signals; and sends them to his smart-phone via Bluetooth, where an application software transforms them into audible form. The smart-phone need not be handled during the whole of the course.

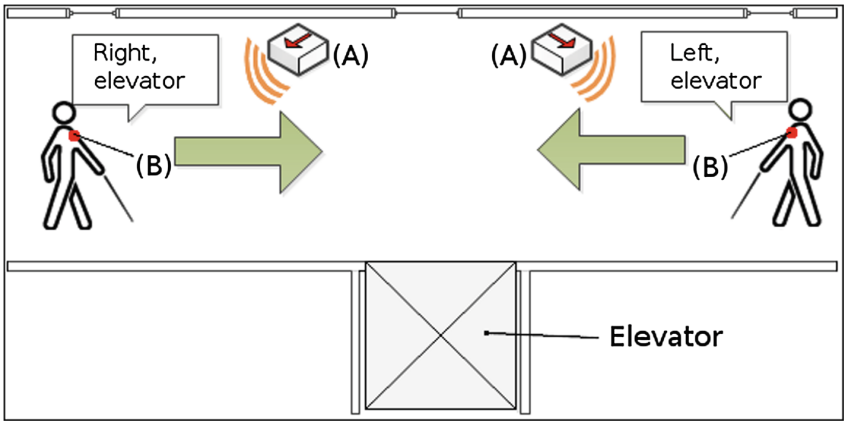


Fig. 1. NavGATe working example.

Fig. 2. Visual impaired participant using NavGATe to find a staircase, not reachable via tactile mat.

In this experiment, NavGATe was programmed to orient users in a manner consistent with information already provided by tactile mats, while offering added information and alternative paths, from the alighting from trains at Vergueiro station, to one of its exits to the São Paulo Cultural Center, that hosts a library (with books in Braille) much used by visually impaired people.

Figure 2 illustrates a participant using such system to reach a staircase that cannot be reached through tactile mat. In this station, the tactile mat is directed towards an elevator on the platform (Fig. 2 "A"). However, at an appropriate point during the course, NavGATe informs the user of the alternative use of stairs, to her left, which she accepts (Fig. 2 "B"). Beyond this point, having no tactile mat, the user relies only on NavGATe (Fig. 2 "C") and manages to reach the stair case (Fig. 2 "D")

5 Results

Results and discussions comprise the whole of the research, including [8]. The interviews with metro operation staff and systems designers came to the conclusion that metro stations are great circulation spaces, and flows vary during the day and also in time, and this is an important factor in their design. In addition, metro stations embody systems not perceived by its users, but that can affect and be affected by any new system that may be introduced. This requires careful planning regarding the choice of electromagnetic frequencies, cabling or transmission systems used. Another important consideration is the diversity of users, and their varied needs.

Based on the above, plus observation and interviews with visual impaired users, further conclusions are that any such system design must seek to achieve: (a) safety and reliable information; (b) effectiveness – provide the relevant information and; (c) comfort – ease of use and customizable.

The interviews with the orientation and mobility instructors [8] showed favourable opinions to the use of electronic systems as orientation and navigation aids. These instructors also provided useful information, regarding spatial perception and cognition, not found in literature and that hadn't been identified through user observation.

In the assessment of the existing Navmetro in the Porto metro [8], the interviews of personnel and the faculty researcher involved in its implementation, and observation of a visual impaired user, evidenced both the necessity and the difficulty in creating a system capable of determining the exact location of the user in order to provide proper guidance; also the possibility of use of differentiated sounds (e.g. bird sounds), that are clearly identified by the visually impaired, while not disturbing others; the importance of flexible user interfaces that enable users to refer to places in different ways.

In the Porto assessment, the respondents, involved in its implementation, also revealed the limitations of the use of Wi-Fi and Bluetooth in metro environments, due to its inherent inaccuracy in positioning, as well as the interference it potentially causes and suffers from other systems required in normal operation. This corroborates the affirmations of Furukawa and Bruno [41]; and Pereira [42], regarding the use of radio wave technology.

The respondents also informed that Navmetro needed improvements, the ability to adapt to user preferences and expertise. Such had been, respectively, cited by Gedawy

[45], and Booch et al. [46]. This was confirmed during the observation of an experienced user who got annoyed at having to repeat always the same steps in its operation.

Other important conclusions, from literature review, user and specialists (in metro operations and design) interviews, and user observation, that form the first phase of this research [8], were that, for any proposed user orientation and navigation system:

(a) user safety comes first;
(b) its usage mustn't cause perturbation to existing systems nor disturbance to other users;
(c) there mustn't be a limit to the number of simultaneous users;
(d) if a device is to be carried by the user, it should be of simple use, able to provide information through audio, preferably (though not mandatory) using vibration - as recommended by Yatani et al. [44] and stated in user and mobility instructors interviews - to confirm commands, indicate left or right, etc.;
(e) it must include information regarding trains; their destination, door states, next station and, alert users on platforms of existing gaps between cars;
(f) it must be planned for continuous evaluation and updates;
(g) its boundaries are station entrances and exits and it must encompass all possible paths, in association or not with tactile mats; also inside trains and in transfers between trains;
(h) user location must be given accurately;
(i) any possible interference to or from existing systems must be foreseen and avoided;

Guiding visual impaired people through metro stations and in and out of trains is a complex and challenging real time task, which can be tackled by dividing its path into segments derived from orientation and mobility requirements. This was cited by Harper [30] and confirmed in the observations of Navmetro and NavGATe users. In this regard the second phase of the research confirmed earlier findings considering the design of electronic guidance systems:

(a) it must inform users of the correct direction of escalators, elevator door status, escape routes and other operational information – as stated in the users and specialists interviews;
(b) it must be able to detect route deviations and provide the necessary corrective instructions – from observations of the use of Navmetro and NavGATe;
(c) if a device is to be carried by the user, e.g. a smart-phone, it mustn't require to be held full time, i.e., one hand must be free for other needs - as stated by ONCE [43] and verified in the experiments using Navmetro [8] and NavGATe - as stated in user and mobility instructors interviews and during all user observations;
(d) the advantages of adding the needed functionality to the user's own smart-phone; choice of narrator voices; usage of a single earphone, for guidance or other uses, e.g., listen to music, or receive calls – as stated in user and mobility instructors interviews and during all user observations;
(e) and function also as an audio guide – as recommended by Almeida [34], verified through the interviews in [8] and also experimentally while testing NavGATe.

6 Conclusion

Despite the difficulties posed by current metro systems to the visually impaired, which leads to 85% of these users resorting to CMSP staff in helping them, 72% of whom would rather be independent [8]. It was initially thought that this can be achieved through the use of electronic orientation and navigation systems, since people have become accustomed to smart-phones and apt to absorb information technology. However, there were few and inconclusive studies in the field.

Navmetro, in the city of Porto metro, is one of the very few examples of systems that made it out of the drawing board, were produced and that are actually in use. Other systems for indoor use were assessed, because of the similar environmental conditions. The possibility of interfering with other systems required for metro operation was appraised in each case.

The list of directives obtained through this study, novel or confirmation of earlier works, compiled into a single text, is an important resource for planners and designers regarding such systems. Also, since it comprises a thorough assessment of the different visual impairment types and their abilities, and, the design of information, and navigation systems, its findings could benefit also other transportation systems and the use of electronic guides for use in public spaces in general.

NavGATe is scheduled for validation tests and homologation within the São Paulo metro. As of March 2017 it is being fully installed on the station where previous testing were done.

Further research should deepen knowledge of navigation aspects, notably: sound and voice communication; determine best phrases and vocabulary used in commands and most appropriate moment to emit a command, as well as best receiver design (e.g. shape and handling), considering the diversity of user needs.

References

1. Instituto Brasileiro de Geografia e Estatística – IBGE: Censo Demográfico 2010. Características Gerais da População, Religião e Pessoas com Deficiência. Compilation 2012, Rio de Janeiro, Brasil (2012). ftp://ibge.gov.br/Censos/Censo_Demografico_2010/Caracteristicas_Gerais_Religiao_Deficiencia/caracteristicas_religiao_deficiencia.pdf. Accessed 13 June 2014
2. Jacobson, R.D.: Talking Tactile Maps and Environmental Audio Beacons: An Orientation and Mobility Development Tool for Visually Impaired People. Institute of Earth Studies, University of Wales Aberystwyth, Ceredigion, SY23 3DB, U.K. (1999)
3. Fratocchi, A.C.G., Guedes, C., Silva, C.T., Cardoso, D.D., Mariani, E., Pontes, F.M.; Ferreira, G., Borges, M. L., Takeshita, N.: Ações de Relacionamento Inclusivas - Experiências do Metrô de São Paulo. Revista Engenharia. São Paulo, n. 626, pp. 162–167 (2015)

4. Brasil: Decreto No 5.296, de 2 de dezembro de 2004. Regulamenta as Leis n. 10.048, de 8 de novembro de 2000, que dá prioridade de atendimento às pessoas que especifica, e 10.098, de 19 de dezembro de 2000, que estabelece normas gerais e critérios básicos para a promoção da acessibilidade das pessoas portadoras de deficiência, ou com mobilidade reduzida, e dá outras providências. Diário Oficial da União, Brasília (2004)
5. Associação Brasileira de Normas Técnicas - ABNT: NBR 9050 - Acessibilidade a Edificações, Mobiliário, Espaços e Equipamentos Urbanos (NBR 9050 - Acessibility to Buildings, Equipment and the Urban Environment). ABNT, Rio de Janeiro (2004)
6. Associação Brasileira de Normas Técnicas - ABNT: NBR 14021 - Transporte - Acessibilidade no sistema de trem urbano ou metropolitano (NBR 14021 - Transport - Accessibility on urban or metropolitan train system). ABNT, Rio de Janeiro (2005)
7. Companhia do Metropolitano de São Paulo - CMSP: Gerência de Operações. Percepção da Pessoa com Deficiência Visual sobre o Deslocamento no Metrô de São Paulo. Research Report GOP/OPR n. 14/2015. CMSP (2015)
8. Mariani, E.: Delineamento de Sistemas Eletrônicos para Guiar Pessoas com Deficiência Visual em Redes de Metrô (Guidelines for Electronic Systems Designed for Aiding the Visually Impaired People in Metro Networks). Master's Dissertation, Faculdade de Arquitetura e Urbanismo, University of São Paulo, São Paulo (2016). http://www.teses.usp.br/teses/disponiveis/16/16132/tde-02092016-151522/. Accessed 28 Feb 2017
9. Ferreira, E., Freitas, D.: Navmetro®: preliminary study applications of usability assessment methods. UDESC Hum. Factors Des. **1**(2) (2012).http://www.revistas.udesc.br/index.php/hfd/article/view/3134/2311. Accessed 5 Mar 2017
10. Silva Filho, J., Mariani, E.: Research About Usability and Intuitiveness of the Synthesized Voice in Electronic Navigation and Guidance Unit for Visually Impaired People. 1º Congresso Brasileiro de Pesquisa & Desenvolvimento em Tecnologia Assistiva – CBTA, Curitiba, Brasil, pp. 50–58 (2016)
11. Flick, U.: An Introduction to Qualitative Research, 4. ed. SAGE Publications, Thousand Oaks (2009)
12. Creswell, J.W.: Research Design: Qualitative, Quantitative and Mixed Methods Approaches. SAGE Publications, Thousand Oaks (2014)
13. Moresi, E. (Org.): Metodologia da pesquisa. Programa de Pós-Graduação Stricto Sensu em Gestão do Conhecimento e Tecnologia da Informação. Universidade Católica de Brasília, Brasília (2003)
14. Angrosino, M.: Doing Ethnographic and Observational Research. SAGE Publications, Thousand Oaks (2007)
15. Cohen, R.: Cidade, Corpo e Deficiência: Percursos e Discursos Possíveis na Experiência Urbana (City, Body and Disability: Possible Routes and Speeches in the Urban Experience). Doctoral Thesis. Instituto de Psicologia, Federal University of Rio de Janeiro, Rio de Janeiro (2006). Retrieved 2017-02-28, from http://pos.eicos.psicologia.ufrj.br/wp-content/uploads/reginacohen.pdf
16. Valentini, S.M.R.: Os Sentidos da Paisagem (The senses of the landscape). Doctoral Thesis, Faculdade de Arquitetura e Urbanismo, University of São Paulo, São Paulo (2012). doi:10.11606/T.16.2012.tde-01022013-143130,www.teses.usp.br. Accessed 27 Feb 2017
17. Queiroz, V.M.: Acessibilidade para Pessoas com Deficiência Visual: uma Análise de Parques Urbanos (Accessibility for Visually Impaired Persons: an Analysis of Urban Parks). Master's Dissertation, Faculdade de Arquitetura e Urbanismo, University of São Paulo, São Paulo (2014). doi:10.11606/D.16.2014.tde-21102014-173356,www.teses.usp.br. Accessed 28 Feb 2017

18. Barbosa, M.B.P.: Wayfinding na Jornada da Pessoa com Deficiência Visual no Sistema Metroferroviário (Wayfinding in the Person's Journey with Visual Impairment in the Subway-Railroad System). Doctoral Thesis, Faculdade de Arquitetura e Urbanismo, University of São Paulo, São Paulo (2015). http://www.teses.usp.br/teses/disponiveis/16/16132/tde-08032016-163338/. Accessed 28 Feb 2017
19. Gerhardt, T.E., Silveira, D.T. (Org.): Métodos de Pesquisa. Editora da UFRGS, Porto Alegre (2009)
20. Rose, D.: Analysis of moving images. In: Bauer, M., Gaskell, G. (Org.): Qualitative Researching with Text, Image and Sound: A Practical Handbook, pp. 246–262. SAGE Publications, Thousand Oaks (2000)
21. Gill, R.: Discourse analysis. In: Bauer, M.; Gaskell, G. (Org.). Qualitative Researching with Text, Image and Sound: A Practical Handbook, pp. 172–190. SAGE Publications, Thousand Oaks (2000)
22. Hatwell, Y.: Psychologie Cognitive de La Cécité Precoce. Dunod, Paris (2003)
23. Bradley, N.A., Dunlop, M.D.: An experimental investigation into wayfinding directions for visually impaired people. Pers. Ubiquit. Comput. **9**(6), 395–403 (2005). doi:10.1007/s00779-005-0350-y
24. Taylor, S.E., Peplau, L.A., Sears, D.O.: Social Psychology. Pearson/Prentice Hall, Upper Saddle River (2006)
25. Bins Ely, V.H.M.: Como Criar Espaços Acessíveis para Pessoas com Deficiência Visual a Partir de Reflexões sobre Nossas Práticas Projetuais. In: Prado, A.R.A., Lopes, M.E., Ornstein, S.W. (Orgs.): Desenho Universal - Caminhos da Acessibilidade no Brasil, pp. 95–104. Annablume, São Paulo (2010)
26. Rogers, Y., Sharp, H., Preece, J.: Interaction Design: Beyond Human–Computer Interaction, 3rd edn. Wiley, New York (2011)
27. World Health Organization – WHO: Visual Impairment and Blindness. Fact Sheet N°282. [Internet - Updated August 2014]. http://www.who.int/mediacentre/factsheets/fs282/en/. Accessed 15 Apr 2015
28. Felippe, J.A.M., Felippe, V.L.R.: Orientação e Mobilidade. Laramara – Associação Brasileira de Assistência ao Deficiente Visual, São Paulo (1997)
29. Associação dos Cegos e Amblíopes de Portugal - ACAPO: Recomendações sobre Atendimento a Pessoas com Deficiência Visual. Núcleo de Estudos e Investigação em Acessibilidade, Nov. (2013). http://www.acapo.pt/. Accessed 18 Aug 2014
30. Harper, S.: Standardising Electronic Travel Aid Interaction for Visually Impaired People. Master's Dissertation. Institute of Science and Technology, University of Manchester, Manchester (1998)
31. Honorato, S.: Percepção de Imagens Através de Frequências Vibratórias Captadas pelas Mãos de Pessoas Cegas. Master's Dissertation, Programa de Pós-graduação em Design e Expressão Gráfica, Federal University of Santa Catarina, Florianópolis (2013). https://repositorio.ufsc.br/bitstream/handle/123456789/107128/318331.pdf?sequence=1&isAllowed=y. Accessed 28 Feb 2017
32. Lora, T.D.P.: Descobrindo o Real Papel das Outras Percepções, Além da Visão, para a Orientação e Mobilidade. In: MOTA, M.G.B. da. Orientação e Mobilidade: Conhecimentos Básicos para a Inclusão da Pessoa com Deficiência Visual. MEC/SEESP, Brasília (2003). http://portal.mec.gov.br/seesp/arquivos/pdf/ori_mobi.pdf. Accessed 14 Dec 2013
33. Espinosa, M.A., Ungar, S., Ochaiata, E., Spencer, M.B.C.: Comparing methods for introducing blind and visually impaired people to unfamiliar urban environments. J. Environ. Psychol. **18**, 277–287 (1998)

34. Almeida, M.F.X.M.: Auxílios à Navegação de Pedestres Cegos Através de Mapa Tátil (Aid to Navigation for Blind Pedestrians Through Tactile Map). Master's Dissertation, Programa de Pós-graduação em Design, Federal University of Pernambuco, Recife (2008). repositoio.ufpe.br/bitstream/handle/123456789/3151/arquivo2185_1.pdf?sequence=1&isAllowed=y. Accessed 28 Feb 2017
35. Crandall, W., Bentzen, B.L., Myers, L., Mitchell, P.: Transit Accessibility Improvement Through Talking Signs - Remote Infrared Signage: A Demonstration and Evaluation. Report from the Federal Transit Act, Cooperative Agreement with the U.S. Department of Transportation, Federal Transit Administration and Project Action of the National Easter Seal Society, San Francisco (1995)
36. du Buf, J.M.H., Barroso, J., Rodrigues, J.M.F., Paredes, H., Farrajota, M., Fernandes, H., José, J., Teixeira, V., Saleiro, M.: The smartvision navigation prototype for blind users. JDCTA Int. J. Digit. Content Technol. Appl. **5**(5), 351–361 (2001)
37. Almeida, A., Orduña, P., Castillejo, E., López-de-Ipiña, D., Sacristan, M.: Imhotep: an approach to user and device conscious mobile applications. Pers. Ubiquitous Comput. **15**(4), 419–429 (2011)
38. López-de-Ipiña, D., Klein, B., Vanhecke, S.: Towards ambient assisted cities and citizens. Deusto Institute of Technology, DeustoTech University of Deusto Bilbao, Spain. European Software Institute, ICT Division, Zamudio (2013)
39. Virtanen, A., Koskinen, S.: NOPPA - navigation and guidance system for the visually impaired. VTT Industrial Systems. Fin-33101 Tampere, Finland (2014)
40. Hersh, M.A., Johnson, M.A. (Org.): Assistive technology for visually impaired and blind people. Springer, London (2008)
41. Furukawa, R.N., Bruno, L.C.: Dispositivo Remoto para Controlar Funcionalidades de Aplicativos em um Computador. XXVI ENEGEP, pp. 1–9 (2006). http://www.abepro.org.br/biblioteca/ENEGEP2006_TR490329_6902.pdf. Accessed 14 Jan 2015
42. Pereira, E.D.: Desenvolvimento de um Sistema de Localização de Fontes Rádio Frequência para aplicações Indoor. Master's Dissertation. Faculdade de Engenharia, University of Porto, Porto (2011). http://paginas.fe.up.pt/∼ee01272/DIS/lib/exe/fetch.php?media=mieec.pdf. Accessed 28 Feb 2017
43. Organización Nacional de Ciegos Españoles - ONCE.: Accesibilidad para Personas con Ceguera y Deficiencia Visual. 1. ed. ONCE, Madrid (2003)
44. Yatani, K., Banovic N., Truong, K.N.: SpaceSense: Representing Geographical Information to Visually Impaired People Using Spatial Tactile Feedback. University of Toronto, Department of Computer Science. Microsoft Research Asia (2012)
45. Gedawy, H.K.: Designing an Interface and Path Translator for a Smart Phone-Based Indoor Navigation System for Visually Impaired Users. Master's Dissertation, School of Computer Science, Carnegie Mellon University, Pittsburgh (2011)
46. Booch, G., Maksimchuck, R.A., Engle, M.W., Young, B.J., Conallen, J., Houston, K.A.: Object Oriented Design with Applications, 3rd edn. Pearson Education Inc., Boston (2007)

Road and Rail: Public Transport

A Comparison of Auditory and Visual in-Vehicle Stop Signals in Philippine Public Road-Based Transportation

Josiel Catherine V. Almeda(✉), Majella Cayley J. Domingo, and Alyannah Bianca L. Saringan

College of Engineering, University of the Philippines Diliman, 1101 Quezon City, Philippines
{jvalmeda, mjdomingo, alsaringan}@up.edu.ph

Abstract. Public road-based transportation vehicles (PRTV's), mainly jeepneys and buses, are popular modes of transportation in the Philippines. While some jeepneys today have added mechanisms, such as switch-controlled buzzers, for passengers to indicate their stop, the most prevalent method remains to be the system of calling out to the driver. This study aims to compare which type of signal (auditory, visual, audiovisual, voice) will be the most effective in catching a PRTV driver's attention. Simulating a driver's environment, participants were tasked to respond whenever they were presented with a signal. Response times were gathered and analyzed using Analysis of Variance. It was determined that there was a significant difference among the different types of signals. Further comparison tests showed that the Light+Buzzer (audiovisual) signal elicited the fastest reaction time while the voice signal caused the slowest reaction.

Keywords: Public transportation · Stop signals · Passenger · Comparison

1 Introduction

1.1 Background of the Study

Public road-based transportation vehicles (PRTVs), such as buses and jeepneys, are an essential part of the average Filipino's commute. Jeepneys, based on military jeeps and popularized in the Philippines after World War II, are designed with seats beside the driver at the front, and perpendicular to the driver at the back. They can accommodate an average of 18–22 passengers. These vehicles are some of the most popular modes of transit in the country, as they are present in all regions and major cities, and offer varying routes and lengths of travel. According to the Metro Manila Development Authority (MMDA), public transport vehicles make up 20% of the vehicles used in the National Capital Region (NCR) in the Philippines. Of the 20%, 3% are made up of buses, and 40% are made up of jeepneys. Additionally, more than 40% of the daily trips in the Metro are carried out by jeepneys [1].

The most common way for a passenger to inform the driver of their exit from the vehicle is by directly telling them to stop once their destination is reached. Filipinos do this by saying the word "*Para*". This can often be difficult for the driver to detect, as the

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_97

traffic conditions of Philippines' major cities tend overload them with other types of information. As a result, vehicles would surpass or miss stops entirely, causing an inconvenience to the passenger. To combat this problem, jeepneys in select routes have added buzzers or light indicators connected to a rope or switch mechanism on their ceilings, which passengers can then pull or press on as a form of request to stop.

1.2 Rationale

The design and process of riding a jeepney has evolved minimally since its conception, while the local buses are underdeveloped and lag behind their international counterparts. This has left room for improvement in terms of efficiency for some aspects of the design, such as the way to signal a driver to stop when passengers wish to alight. There are however little to no studies related to the method of riding a jeepney, as the main focus of its design is on optimizing its capacity and structure. And while studies exist on a driver's response to signals, their main focus is not on public transportation. Determining these most basic aspects, i.e. the effectiveness of calling out to the driver compared to newer signals, is the first step in improving the method.

2 Problem Statement

There is a lack of comparative studies on the effectiveness of stop signals in informing public road-based transportation vehicle drivers of the alignment of their passengers.

3 Review of Related Literature

3.1 Sources of Distraction and Signal Perception

A driver processes information actively from the environment by continuously selecting and transforming them. This is known as the driver's perception. There are several modes of acquiring this information, reaction time being one of the most used methods in measuring the perception of drivers. Its main benefit would be the lack of time lapse errors in experiments making it more convenient [2].

Most experiments and discussions related to the perception of drivers to in-vehicle signals are for the warning and prevention of road accidents or vehicular collisions. They reason that these signals must overcome the innate distraction of driving as well as other possible non-driving activities, such the use of mobile phones and other devices. However, fewer studies exist on the unique distractions that drivers of public transportation vehicles are exposed to. Regan et al. [3] identified seven main sources of distraction in city bus drivers from New South Wales, Australia: technology-related, environmental, infrastructure-related, personal distractions, operational, passenger-related, and bus cabin-related. The latter three are experienced only by PRTVs.

In the Philippines, the physical framework of the two main PRTVs allow for the country's warmer climate and the ambient city traffic sounds to affect the driver. Most jeepneys lack glass windows and doors and are essentially open to additional

distractions from the environment, especially in the presence of rain. Many city buses also keep all their windows and doors open as they do not have air-conditioning systems to ventilate the vehicle. Their structure gives cause for the additional distraction from the signals of passengers who want to leave the vehicle, as their main method is the use of their voice.

The increased distraction PRTV drivers are subject to calls for an improvement in the method of signaling them to stop. Several studies exist on the efficacy of stimulating different senses to call attention to the driver. Though it is commonly assumed that auditory signals are the most effective, since driving is a mainly visual activity, research shows that this is not always true. As summarized by Ho and Spence [4], studies show that driving in the midst of accomplishing separate auditory tasks, such as the use of mobile phones and conversing with passengers, resulted in a diminished driving performance. They instead pointed out the growing research on multimodal signals, or signals that prompt more than one of the senses. It was seen that these signals elicited better response times compared to their unimodal counterparts i.e. faster reactions to audio-tactile signals compared to only auditory or tactile signals.

3.2 Auditory Displays and the Design of an Alarm Signal

One important aspect of the study was constructing a usable auditory display, specifically an in-vehicle auditory alarm signal, whose purpose is to alert the driver of the passenger's request stop. Auditory displays are perceived through hearing the sound it emits. There are three main characteristics attributed to sound: frequency, intensity and duration [5].

Frequency is defined in the Basics of Sound, 2016 [5] as the number of vibrations that travel through a medium per unit time, and can be expressed with the unit Hertz (Hz). The range of a human adult's perception of frequency is between 20 Hz to 20,000 Hz [6]. Meanwhile, intensity is described as the energy of the vibrations of the sound, with higher energy levels recognized as louder magnitudes. They are measured using the unit decibel (dB). The lowest intensity level discernable to the human ear is approximately 0 dB, while the highest intensity level before reaching pain threshold is set at 120 dB. Other important values to take note of for the study are the following: 60–70 dB which is the ambient background noise in metropolitan area, 70–80 dB or the noise on the sidewalk with passing car or bus, and 76 dB, the intensity level of a raised voice [5].

The guidelines set by Edworthy [7] were used as a basis for constructing the alarm signal for the study. Edworthy [7] states that the decision on the kind of tone to be used is dependent on the information to be relayed to the recipient of the message. For example, higher undulation conveys urgency for action, even sometimes with the intention to scare. More consistent tones can be used during less pressing situations, and can merely be used to alert.

The next aspect considered is deciding the frequency of the alarm. It is recommended that the frequency falls within the range of 200–5000 Hz. In addition, display frequency can be used to inform the recipient of the gravity of the situation, with higher frequencies corresponding to higher need for action [6]. Non-speech auditory signals within the range of 100–500 Hz show efficiency in masking simultaneous speech

signals. The last characteristic reviewed is the intensity of the auditory display. Nees and Walker [8] refers to a consensus that levels must be at least 15 dB and at most 25 dB above the ambient background noise.

3.3 Visual Displays and Light

The second mode of signaling a driver to stop is through a light indicator. Light can only be perceived by the human visual system at certain wavelengths below 460 nm, which is the measure used by the lighting industry. Light Emitting Diodes (LEDs) are used for traffic signals and other road lights and generally function as the different light fixtures needed today [9].

There is a study that aims to understand how LED traffic signals of different luminance affect the performance of drivers. A task simulation was performed by steering the vehicle heading towards an intersection with traffic signals. In one of their experiments, they measured the reaction times of 21 subjects. The light signals presented were always from the same source at constant luminance and color. A significant difference was found in the ability to detect red, yellow and green signals. From this, the yellow signal was determined to have the shortest reaction time while the green signal was the longest [10]. Regardless, a red light signal was chosen as the visual signal for the experiment, referencing the current system of red light indicators used by Philippine jeepneys as the benchmark.

Auditory and visual signals have been compared in another study through reaction times. Kumar and Shelton [11] tested 12 participants who were presented with stimuli and asked to press a button once they recognized it. With the use of the Direct RT program, they were able to measure the response time accurately and efficiently. It determined that reaction time was faster when responding to an auditory stimulus compared to a visual stimulus. This test was done purely to determine which had a faster reaction time and did not incorporate driving into it.

4 Methodology

4.1 Determination of Signals

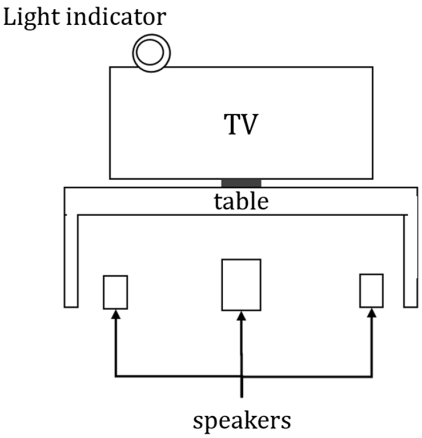
Three distinct types of signals were compared to the current method of using one's voice. Each one is perceived by a different combination of senses: auditory, visual, and audiovisual signals. These signals were replicated using common indicators already found in present jeepneys: a Buzzer, a Red Light and the simultaneous Light+Buzzer signal. A voice recording of a person saying "Para" was used to simulate a passenger telling the driver to stop in order to exit.

4.2 Participants

Data from 30 male participants was gathered. All participants had at least 2 years of legal driving experience. Overall, their average age was 21 and their average number of legal driving years was three.

4.3 Set-up and Calibration

The conditions that PRTV drivers are exposed to were simulated by asking the respondents to engage in a third person driving game amidst traffic noises. The driving game was viewed in a television screen while the ambient traffic noise was played from a set of speakers (Figs. 1, 2).



FRONT VIEW

Fig. 1. Front view of the set-up

a. The temperature of the room was set to room temperature, or about 23 °C.
b. The television screen was set to show the driving game. The in-game audio was muted.
c. A chair was placed 1.5 to 2 m away from the TV screen. Two reference markers, for the subject's right foot, were placed in front of the chair. This served to simulate the position of the brake and gas pedals.
d. Three speakers playing traffic noises were set directly below the screen and faced towards the chair. The intensity level of the noise from the position of the subject was adjusted to about 75 dB [5].

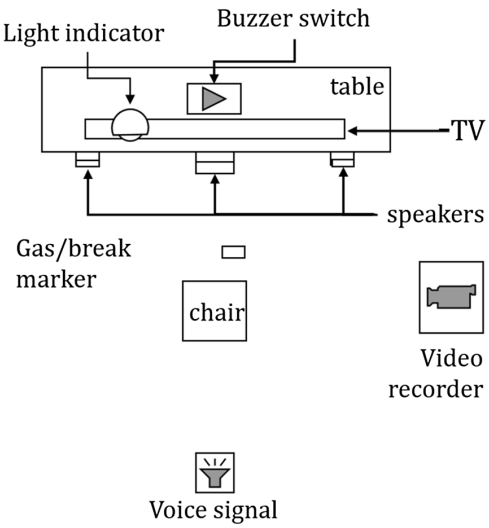


Fig. 2. Top view of the set-up

e. The buzzer was placed directly behind the screen, hidden from the subject's view, with one tester controlling its signal. The buzzer's intensity level, as heard by the subject, was calibrated to about 90 dB [7].
f. The light indicator was placed on the top left corner of the screen, with respect to the subject. The switch of the light source was located behind the screen, hidden from the subject's view. The tester for the buzzer also controlled the signal for the light indicator.
g. The device containing the voice recording was set 2 m behind the chair. Another tester controlled its signal. The intensity level of the recording was adjusted to 76 dB, the approximate level of a raised voice [5].
h. On the right of the chair, a video recorder was situated to capture the light indicator and the movements of the subject's right foot.

4.4 Procedure

a. The subject was seated comfortably on the chair, with their right foot relaxed on the gas marker. The controls of the driving game were explained and the subject was given a short time period to acclimatize to the controls.
b. The four signals were presented to the subject prior to starting the video recording in order to familiarize them with the signal. The subject was instructed to respond to each cue by moving their foot onto the brake marker, then back.

c. After the video started recording, the signals were presented at random intervals while the subject played the game. Each subject was presented with each signal three times.
d. A high accuracy video player was used to determine the subject's reaction time up to the millisecond.

4.5 Summary and Analysis of Data

The data was analyzed using repeated measures Analysis of Variance (ANOVA) due to each reaction time being recorded from each participant in only a single run. This method involved two factors, or sources of variability: the signal type and the individual. The former is a between-group, random factor while the latter is a within-group, fixed factor. The response is the reaction time, defined to be the time between the start of the signal and the moment the respondent's foot is flat on the brake marker. The normality assumption of ANOVA was followed due to the Central Limit Theorem.

The null hypothesis (H_o) of the ANOVA test states that the treatment or factor (type of signal) has no effect on the reaction time. The alternative hypothesis (H_A) states that at least one type of signal resulted in a significantly different reaction time. This is summarized below where τ_i is the effect of treatment i.

$$H_o : \tau_i = 0, \quad for\ all\ i \tag{1}$$

$$H_A : \tau_i \neq 0, \quad for\ atleast\ one\ i \tag{2}$$

However, if the means are not found to be equal to each other, ANOVA does not specify which of the signals prompted the significantly faster or slower reaction times. Succeeding tests, such as Fisher's Least Significant Difference Method (LSD) and Tukey's Test, were conducted to determine the significantly different factors. All of the analysis was conducted at a 0.05 level of significance, or alpha level.

5 Results and Discussion

5.1 Significance of the Type of Signal

A p-value that is less than the level of significance indicates the rejection of the null hypothesis and that the factors tested are significant. Analysis of the ANOVA table, presented in Table 1 shows that type of signal resulted in a p-value of ≤ 0.001. This p-value is less than the 0.05 alpha level and indicates that the mean reaction times among the types of signals are not all equal.

At first glance, the signal that prompted the fastest average reaction time was the Light+Buzzer, while the slowest average reaction time was by the Voice signal. Subjects reacted almost 200 ms slower to the Voice signal compared to the Light+Buzzer. Moreover, both the visual and auditory displays elicited slower average reaction times compared to the audiovisual signal. These results are consistent with the experiments by Ho and Spence [4] where multimodal signals stimulated faster response times compared to unimodal signals.

Table 1. Repeated measures ANOVA table

Source	Degrees of freedom	Sum of squares	Mean squares	F-value	*p*-value
Signal	3	1.941	0.647	8.79	0.000
Individual	29	13.513	0.466	6.33	0.000
Error	327	24.063	0.074	–	–
Total	359	39.516	–	–	–

Further, the adjusted R^2 value is 33.15%. While this value indicates what percentage of the variability in reaction times is taken into account by the model, it does not, however, affect the results of the ANOVA table. The significance of the factor is only reliant on the results of the F-test in ANOVA. Since, the objective is only to determine the significance of the type of signal, and due to the qualitative nature of both factors, the regression equation is not appropriate (Table 2).

Table 2. Mean reaction time per signal

Signal	Mean (ms)
Light+Buzzer	667.2
Buzzer	745.3
Light	817.6
Voice	860.3

5.2 Pairwise Comparisons

Fisher's LSD method and Tukey's Test are similar in that they analyze two types of signals at a time and determine whether the pair is significantly different from each other, i.e. if the resulting *p*-value is less than the level of significance 0.05. Table 3 summarizes each pair of signals and the *p*-value obtained from both tests. It shows that both methods conclude the significant difference between the reaction times of the Light and Light+Buzzer signals, the Voice and Light+Buzzer signals and the Voice and Buzzer signals. These results also denote that while the Light+Buzzer signal had a faster mean reaction time, it is not significantly different from the Buzzer signal.

Table 3. Pairwise comparisons

Signals		Tukey's test *p*-value	Fisher's LSD *p*-value
Buzzer	Light	0.279	0.075
Buzzer	Light+Buzzer	**0.215**	**0.054**
Buzzer	Voice	**0.023**	**0.005**
Light	Light+Buzzer	0.001	0.000
Light	Voice	0.717	0.292
Voice	Light+Buzzer	**0.000**	**0.000**

Figure 3 illustrates the result of the pairwise comparison tests. The signals that are not contained within the same circles are significantly different from each other. It shows that the Light+Buzzer signal is the most effective because it resulted in faster reaction times compared to the Voice and Light signals. Though not significantly different from the Buzzer signal, the latter was not found to be significantly different from the Light signal while the Light+Buzzer was. It can be inferred that the Light+Buzzer signal is most effective.

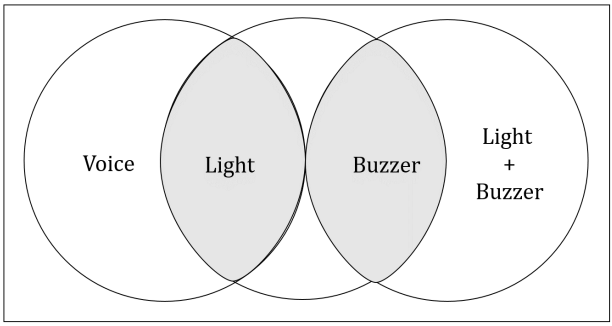


Fig. 3. Diagram of significantly different signals based on Tukey's Test and Fisher's LSD

6 Conclusion

This study simulated the conditions that PRTV's are exposed to in order to determine if there is a more efficient method for passengers to indicate their request to stop. The current system of using one's voice to signal the driver was tested against three other methods: a Light Indicator, a Buzzer and a simultaneous Light & Buzzer signal. The reaction times were analyzed using Repeated Measures Analysis of Variance and Pairwise Comparison Tests. It was found that the Light+Buzzer signal prompted the fastest average reaction time while the Voice signal prompted the slowest mean reaction time. The Light+Buzzer signal was also proved to have significantly faster response times than the Light and Voice Signals. While it was not found to be significantly different to the Buzzer signal, it can still be concluded that the Light+Buzzer signal is the best indicator as the Buzzer signal was also found to be not significantly different to the other signals.

The study suggests that the present method of calling out to the driver is the most ineffective. This could be attributed to two characteristics of the practice: (1) its unimodality, and (2) its tendency to be easily masked by other traffic noises. The current system differs from the Buzzer signal, which was calibrated closely to the optimal intensity and frequency level, and in turn produced better results.

The Light+Buzzer signal produced the best results, which could be accredited to its multimodality. The two modes elicited responses from two senses, instead of the usual one. Hence, the magnitude of the signal has the capacity to be reinforced, given a situation wherein either sense experiences an overload of information arises.

There would still be another sense that could be stimulated in order to confirm the signal. For PRTV drivers, who experience a variety of information everyday, this method is suggested to be more ergonomic than the current system, and among the other systems tested.

7 Areas for Further Study

Further comparative studies can be made by varying the frequency and length of the auditory signal, as well as the color and luminance of the LED light. Having established the optimal method of using an audiovisual signal, there is a wide range of possible combinations that can be made.

The range of signals however is not confined to just auditory and visual. While tactile signals may be harder to implement, there are a lack of studies on their efficacy in the case of PRTV drivers. Further investigations can include ways to adapt this type of signal for the use of passengers, to see if they will prompt better driver response times. The same can be said of multimodal signals involving tactile indicators.

Further research can be conducted on the mechanism that will trigger the signals and their human-machine interaction. Possible factors include the use of a rope or switch and the placement and spacing of the controls in relation to the size and capacity of the vehicle.

References

1. Dabu, F.: Motor Vehicles and Filipino Jeepney Drivers in Metro Manila. Retrieved on November 25, 2016 from http://www.up.edu.ph/motor-vehicles-and-filipino-jeepney-drivers-in-metro-manila/ (2015)
2. Castro, C.: Visual demands in driving. In: Castro, C. (ed.) Human Factors and Visual Cognitive Performance in Driving, pp. 1–29. CRC Press, London (2008)
3. Regan, M.A., Salmon, P.M., Young, K.L.: Distraction and public transport: case study of bus driver distraction. In: Lee, J.D., Regan, M.A., Young, K.L. (eds.) Driver Distraction: Theory, Effects, and Mitigation, pp. 229–245. CRC Press, London (2009)
4. Ho, C., Spence, C.: Crossmodal information processing in driving. In: Castro, C. (ed.) Human factor of Visual and Cognitive Performance in Driving, pp. 187–200. CRC Press, London (2008)
5. Basics of Sound [PDF]. (2016). NoiseQuest
6. Auditory, Tactile and Olfactory Displays [PDF]: University of the Philippines Industrial Engineering and Operations Research Department (n.d.)
7. Edworthy, J.: Urgency mapping in auditory warning signals. In: Stanton, N. (ed.) Human Factors in Alarm Design, pp. 15–30. Taylor & Francis, London (1994)
8. Nees, M.A., Walker, B.N.: Auditory displays for in-vehicle technologies. Rev. Hum. Factors Ergon. **7**(1), 58–99 (2011). doi:10.1177/1557234X11410396
9. Royce, P.R.: Light. In: Royce, P.R. (ed.) Lighting for Driving: Roads, Vehicles, Signs, and Signals, pp. 19–43. CRC Press, London (2008)

10. Bierman, A., Boyce, R., Bullough, J., Conway, M., Huang, K.: Luminous Intensity for Traffic Signals: A Scientific Basis for Performance Specifications, pp. 3–44. Rensselaer Polytechnic Institute, Troy (1999)
11. Kumar, G.P., Shelton, J.: Comparison between auditory and visual simple reaction times. Sci. Res. **1**, 30–32 (2010)
12. Engstrom, J., Harbluk, J.L., Victor, T.W.: Distraction assessment methods based on visual behavior and event detection. In: Lee, J.D., Regan, M.A., Young, K.L. (eds.) Driver Distraction: Theory, Effects, and Mitigation, pp. 135–165. CRC Press, London (2009)
13. Hayami, T., Kemori, S., Morikawa, M., Murata, A.: Basic study on automotive warning presentation to front/rear hazard by vibrotactile stimulation. In: Karwowski, W. (ed.) Advances in Physical Ergonomics and Safety, pp. 519–529. CRC Press, London (2012)

Measuring Mobility and Transport Services: The METPEX Project

Andree Woodcock[1(✉)], Yusak Susilo[2], Marco Diana[3], Roberto Abenoza[2], Miriam Pirra[3], and Michael Tovey[1]

[1] Centre for Mobility and Transport, Coventry University, Priory Street, Coventry CV1 5FB, UK
{A.Woodcock, M.Tovey}@coventry.ac.uk
[2] Kungliga Tekniska Hoegskolan (Royal Institute of Technology), Stockholm, Sweden
{yusak.susilo, roberto.abinoza}@abe.kth.se
[3] Politecnico di Torino, Turin, Italy
{marco.diana, miriam.pirra}@polito.it

Abstract. Public transport is key to access social, economic, civic and cultural life. However, there is still a need to better understand (1) the needs of all transport users and (2) transport provision in cities and regions. The development of an inclusive, validated, passenger experience measurement instrument is the first step in understanding the whole journey, multi-modal journeys. Such a validated tool would enable resources to be focused on areas which travelers felt most important. Such information could be used to create high quality, user centred, integrated, accessible public transport services, capable of attracting and retaining public transport users whilst meeting sustainability targets. This paper describes the METPEX project and the derivation of a set of Key Performance Indicators.

Keywords: User needs · Passenger experience · Measurement instrument · Journey mode · KPIs

1 Introduction

European cities share problems caused by transport and traffic. Transport provision is in many ways unsatisfactory and contributes to deteriorating environments health and wellbeing. Historically, transport has developed in a piecemeal fashion; technological advances have both helped and fueled urban expansion. However, from a transport, health and sustainability perspective many cities are at crisis point. Lack of integration of transport systems has led to fragmented approach, where it is difficult to understand fully the circumstances and needs of transport users. A user-centred approach is needed to understand the experience of all transport users; not just commuters or those using the services of one operator. A method is required, which provides an integrated understanding of whole journey, end-to-end experiences to show where investment is needed. Funded under the EU FP7 program (FP7-SST-2012-RTD-1), between 2012 and 2015, METPEX offers an approach to this, by providing a wide ranging set of

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_98

survey tools to assess the end-to-end journey experience of a wide range of transport users, who may be using a mix of active, private or public forms of transport.

2 Background

Traffic congestion in European towns and cities has an estimated total cost of €80 billion annually, with almost a quarter (23%) of all CO2 emissions from transport coming from urban areas. Additional road fatalities in the EU remains very high (some 28,000 in 2012) with urban areas accounting for 38% of them, and vulnerable users such as pedestrians particularly exposed. Progress in reducing road fatalities has been below average in urban areas [1]. The EU Commission's white paper "Roadmap to a single European transport area – towards a competitive and resource-efficient transport system" called for a 60% reduction in greenhouse gas emissions [2]. This is despite most member states struggling to comply with current EU legal requirements, which seek to protect citizens from harmful exposure to airborne pollutants and particulate matter. To reduce such trends requires a change in thinking and the emergence of new usage patterns.

When METPEX started, 50% of all Europeans used a car every day, compared with 12% who cycled or used public transport (16%). Nearly 40% of travelers encountered problems on journeys within cities, with many reporting the scale and quantity of serious problems as high. A substantial majority believed that air pollution (81%), road congestion (76%), travelling costs (74%), accidents (73%) and noise pollution (72%) were serious problems within cities. The cost of using public transport was regarded as inhibiting its use within cities. More than half believed that that better (56%) and lower priced public transport (59%) would be the best ways to improve urban travel. Less than a quarter of Europeans believed that the urban traffic situation would improve (24%) and most believed it would stay the same (35%) or get worse (37%). Within the EU, just under a third used public transport at least once a week (32%) and fewer than four in ten were satisfied with the price of urban public transport tickets (39%) [2].

Efficient and effective urban transport is seen as being able to significantly contribute to achieving objectives in a wide range of policy domains, for example socio-economic objectives, energy dependency, or climate change. This depends partly on actions taken by national, regional and local authorities. This is evidenced in the number of policies and funding streams which seek to provide more accessible, cleaner pedestrian and cycle ways.

Unfortunately, effective planning requires knowing what the current situation is and being able to predict future scenarios. Although transport telematics is starting to be able to demonstrate and predict movements of goods and people, little progress have been made in measuring the overall quality of end to end, usually multi modal journeys from the travelers perspective. METPEX found significant gaps in knowledge concerning trip chaining, active forms of transport, multimodal journeys and the experience of hard to reach and vulnerable transport users [3]. Transport operators (and other organizations) collect data primarily on the services and situations they are responsible for. Regard for the whole journey experience is lacking. As such, there may be key factors that are discouraging people from making modal shifts that are not understood

or owned by any stakeholder. It may be argued that understanding the nature of the user experience, and addressing issues such as accessibility, would encourage a modal shift towards more sustainable and active forms of transport. This in turn, would lead to reductions in congestion and carbon emissions and improvements to health and the public realm.

The focus of METPEX was on developing Key Performance Indicators (KPIs) to measure the quality of the whole journey (door-to-door) passenger experience (including private or individual forms of transport such as walking, bicycling and car sharing), as a starting point to understanding where improvements need to be made, from the travelers' perspective. Reliable measurement instruments form the basis for robust data collection. These need to be augmented by policy recommendations, commitment and investment by stakeholders to make improvements that will improve the quality of all modes of transport, for all people.

Although many studies have been conducted at both National and EU levels on mobility behaviour, there is no single measurement instrument, which can be used to gather information about the whole journey experience from journey planning to arrival at destination. The lack of such a validated instrument, taking as its unit of analysis the whole journey experience, makes it difficult to understand the barriers faced by active (walking and cycling), public (e.g. bus, metro, tram, light rail, train) or intermodal transport users (e.g. bus-train, cycle-train, bus-walk-train). It is then difficult to make comparisons, to ensure validity and reliability of data, to develop international, multi modal transport benchmarks and standards and to provide a coherent picture of the passenger experience across Europe.

The complexity of urban environments, reduced employment opportunities and the pressure on operators to increase profitability and service efficiency, and the availability of travel information all effect decisions on whether, when and how to travel and the quality of the travel experience. The potential complexity of the travel experiences needs to be acknowledged and measures should consider new mobility patterns, changing transport policy priorities and passenger perceptions and expectations.

The development of an inclusive, validated passenger experience measurement instrument is the first step in creating high quality, user centred, integrated, accessible public transport services which are capable of attracting and retaining public transport users whilst meeting sustainability targets. Such a tool needs to provide reliable data, which can be used by transport providers, policy makers, vehicle designers and municipalities in order to assess where changes need to be made to increase the efficiency, effectiveness and attractiveness of service delivery. Enabling wider access to safe, secure, convenient, comfortable and economical public transport will in turn support the EU's carbon reduction targets by leading to an increase in the number of people who use public or active forms of transport.

Therefore the aim of METPEX was to develop and validate an intelligent measurement instrument, which could be used to measure the quality of the whole journey, multimodal passenger experience across Europe and provide validated KPIs for different modes of transport and user groups. The following sections outline the process of deriving the KPIs.

3 Metpex

Given the stated aim of the project, stakeholder groups for the final outputs of the project were transport operators and authorities, local authorities and other interest groups wanting to measure the quality of the whole journey experience. The ambition was to capture multimodal traveler experiences from journey planning to arrival at destination, allowing for measurement of different journey phases, such as the use of buses, trains, trams, urban waterways, walking, cycling and driving. In terms of inclusivity, the project aimed to look at 'traditionally hard to reach groups' - commuters, women, travelers with children or other dependents, the elderly, young people, rural dwellers, commuters, those on low incomes, with communication difficulties and mobility impairments.

3.1 Method

Measurement instruments were compiled from factors considered in previous research and a pilot study. The project team weighted over 1000 potential factors and mapped the most important 250+ factors against journey stage, mode, traveler type and quality factor (such as safety). Different measurement instruments were developed from these, namely en route surveys, retrospective (paper based and web online), real time (game and navigation apps) and focus groups formats. The question set presented to the traveler depended on their background variables and modes of transport. More information on the detailed design considerations of these instruments can be found in [4, 5], along with preliminary statistical analysis [6, 7].

The survey was conducted for a month simultaneously in Bucharest, Grevena, Coventry, Rome, Vilnius, Dublin, FIA network (Fédération Internationale de l'Automobile), Stockholm and Valencia using all formats. Table 1 shows breakdown of the data against city and user hard to reach group (as indicative of the type of data and breakdowns undertaken. Data were also considered in terms of transport mode, purpose, number of journey legs, number of journey legs and overall satisfaction [8].

Table 1. Responses per city and hard to reach user group [8]

City	Comm. Impaired	Commuters	Low Income	Mobility Restricted	Elderly	Rural dwellers	Travelling with children	Travelling with dependent	Under 24	Visitors	Women
Bucharest	11	76	49	19	21	21	20	2	32	31	83
Coventry	9	67	11	5	16	16	7	1	58	44	74
Dublin	8	209	40	9	10	24	5	3	45	7	78
Grevena	2	8	35	12	38	28	31	0	56	20	35
Rome	8	165	115	22	40	44	4	5	143	32	129
Stockholm	9	110	76	55	31	13	54	3	109	9	151
Valencia	9	62	113	19	49	4	14	2	32	29	127
Vilnius	9	33	31	9	20	3	15	1	22	4	76
FIA networks	150	221	197	57	44	133	29	25	106	157	356
Total	**215**	**951**	**667**	**207**	**269**	**286**	**179**	**42**	**603**	**333**	**1109**

The 5275 valid observations provided immediate information regarding the quality of travelers experience in each of the cities, against mode and user type. The results were combined and and multivariate analyses (mainly principle component) were used to derive Key Performance Indicators [9, 10]. KPIs have emerged as a means of evaluating quality in a transport system. KPIs in public transport can lead to different approaches in transport management. For example, KPIs addressing operating efficiency measuring aspects such as load factor or cost-per-vehicle-kilometre. METPEX KPIs are different because they reflect the user experience (rider comfort, travel speed and reliability, affordability, integration and satisfaction etc.). User-oriented indicators are important for developing public transit systems that respond to user demands and so are able to attract even choice travellers. The KPIs could form the basis of future surveys for use by organizations and authorities to measure the quality of the whole journey, or different parts thereof, for a wide range of traveler groups, including those from traditionally hard to reach groups.

3.2 Higher Level Key Performance Indicators (KPIs)

Only the top-level results can be shown in this section. For more details, the reader is referred to [3]. The initial set of KPIs, and their relationships are shown in Fig. 1.

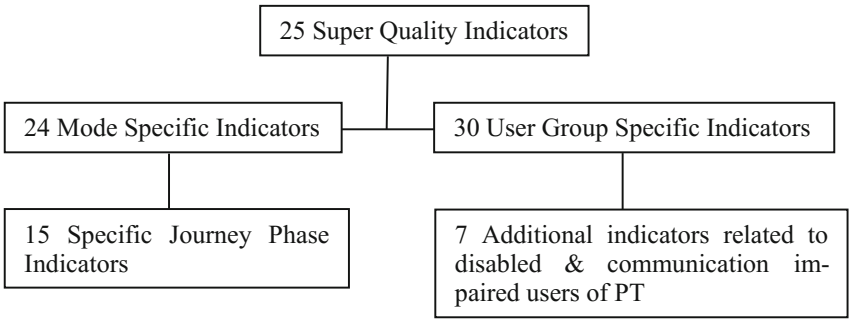


Fig. 1. Relationship of key performance indicators

Figure 2 shows these mapped on to Woodcock's hexagon-spindle model (H-S) [9, 11]. This places the user, in this case the traveler, at the centre of a series of concentric rings, moving out from interaction and workstation design, through to external factors [12, 13]. Under each Service Quality Indicator (SQI) sits a number of factors, which, from the analyses, were shown to be important influences on the quality of the traveler's experience. Critically, Fig. 2 highlights that the travelers' experience needs to be improved across all areas.

Mode specific indicators important to the quality of the traveler experience were grouped as follows [14]:

- Rail services: on trip performance; ticketing and performances before boarding

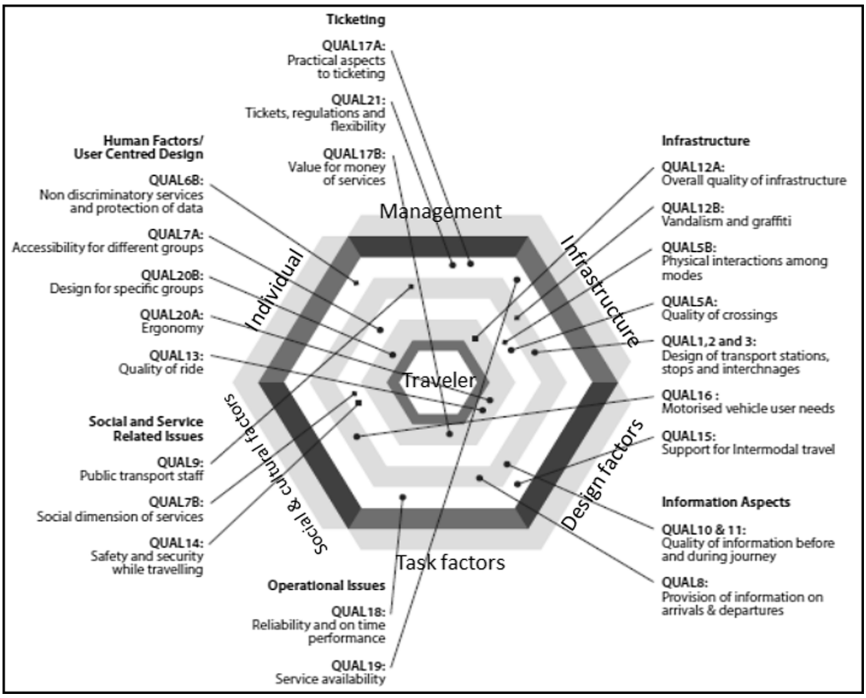


Fig. 2. Quality Indicators mapped on to different aspects of the transport factors. (Adapted from [13, p. 61])

- Metro: on trip performance; ticketing and capillarity
- Tram: service integration and reliability; comfort and staff helpfulness; on-trip quality; tickets and timetabling; fares convenience; information quality
- Bus: reliability; ticketing and other issues; on-board comfort
- Pedestrians: information and safety; environmental factors; sympathy of design; intermodal travel and barrier free access; pavement cleanliness
- Cycling: cycling infrastructure; quality of information, availability and infrastructure; cyclers' interests; public transport and bike trip quality; easiness of carrying bikes
- Private vehicles: traffic calming and parking

Table 2 shows the indicators relating to specific user groups mentioned in Fig. 1. The indicators shown in Fig. 1 were reduced by taking out ambiguities, overlapping categories, repetitions and indicators, which were difficult to measure (given that the aim was to develop a usable measurement instrument) [15]. Figure 3 shows the resulting 14 Super Quality Indicators. These were evaluated in terms of 7 criteria (measurability, specificity, timeliness, uniqueness (in relation to other studies), independence and usefulness) on a 5 point scale. Their relative position in Fig. 3 reflects the score received on the 7 criteria. Those appearing first were rated most highly. Almost all items scored a value above the mean. 2 SQI ("SQI-10: Satisfaction for users of

Table 2. Indicators per traveler group

Group	Indicators
Women	Safety and security, comfort and staff helpfulness; Integrated tickets and range of fares; Reliability
Commuters	Facilities and parking; reliability; ticketing issues; comfort aspects
Travellers over 65	Information and accessibility: performance issues for elders; travel services; infrastructural design
Travellers under 24	Service design and information; reliability; convenience issues
Low income travellers	Low cost services issues; comfort; convenience; information issues
Visitors	Information issues; barriers for foreigners
Rural dwellers	Service coverage and costs; ground aspects; comfort and availability
Travelling with children	Design and quality on-board; ground aspects
Mobility restricted	Quality issues specific for mobility restricted; general quality issues; ability to meet individual needs; ground services
Communication impaired	Quality issues specific for communication impaired; information issues specific for communication impaired; general quality issues; on-board quality and information; comfort; information aspects

motorised private transport means" and "SQI-11: Satisfaction of specific needs for different users groups") were penalised by the use of the indicator before (in previous studies) or the independence, but in both cases their evaluation fell within limits of acceptability. "SQI-6: Reliability of services" was rated most highly because it is the criterion with best measurability level, which highly affects its final value.

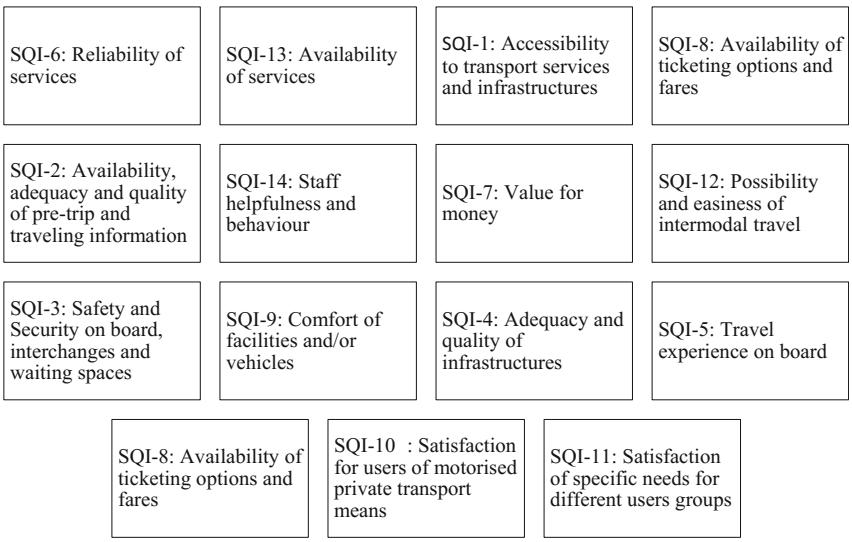


Fig. 3. Final 14 quality indicators [16]

Under each of the SQIs (e.g. as shown in Fig. 3) are sets of indicators. Taking SQ-12 (possibility and easiness of intermodal travel) as an example, this consists of indicators relating to:

- Physical interactions among modes
- Support for intermodal travel
- Public transport plus bike trip quality
- Easiness of carrying bicycles
- Intermodal travel and barrier free access
- Facilities and parking

Each one of these may be regarded as a place holder for a set of statistically related factors. As an example, in terms of the quality of user experience, 'easiness of carrying bicycles' has 4 variables associated with it, which can be used to form the basis of survey tools [15]. These relate to:

- ability to store bikes at stations
- ability to carry bicycle's on public transport vehicles
- level of information available prior to the journey with a bicycle
- adequacy of provision of information on arrivals and departures.

4 Conclusions

One of METPEX's key objectives was to develop and test an integrated approach to the measurement of the whole journey passenger experience to facilitate the harmonization of travel, behaviour research and analysis across European Union Member States. This paper has outlined how a set of Key Performance Indicators were developed from use trials conducted in 8 EU cities, how the material was analyzed and refined into a set of usable KPIs from which future surveys can be developed. Other aspects of the project e.g. focus groups with disadvantaged groups, accessibility and working tools for local authorities are featured in the METPEX book [3]. Key contributions of the project have been in the use social media and real time data collection through bespoke applications.

In developing the KPIs the project has addressed gender and ethical issues in transport research by creating survey instruments which will specifically address the needs of female and vulnerable travelers. The tools developed by the project should enable local and transport authorities to conduct, or at least reduce the cost of commissioning travel surveys – as the project has provided validated instruments. Other aspects of the project will aid decision-making and deployment of technology/solutions to support integrative accessibility. Benchmarking analyses are also possible based on such indicators [17].

In order to enable more integrated, accessible transport systems to develop there has to be a buy in across the transport stakeholder community to investigate and take ownership of travelers' needs. METPEX provided a mechanism for more robust and inclusive measurement. However the tools need to be used, results prioritized and acted upon in order to make multimodal transport options accessible to all.

Acknowledgements. This paper provides a short overview of some of the outputs from the FP7 CP-Collaborative project, METPEX, project number FP7-SCP2-GA-2012-314354, led by Andree Woodcock as PI. As such the authors would like to acknowledge the contributions of the project team: ANGRE (Greece), Coventry University (UK), Eurokleis (Italy), FIA (Belgium), KTH (Sweden), Interactions (Ireland), Integral Consulting (Romania), ITENE (Spain), POLITO (Italy), RSM (Italy), SBOING (Greece), SIGNOSIS (Belgium), SmartContinent (Lithuania), TERO (Greece), VTM (Portugal), ZHAW (Switzerland).

References

1. EC: Special Eurobarometer 406, Attitudes of Europeans towards Urban Mobility Report'. European Commission Public Opinion Survey. http://ec.europa.eu/public_opinion/archives/ebs/ebs_406_en.pdf (2013)
2. COM: Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system. European Commission White Paper, 144 final (2011)
3. Tovey, M., Woodcock, A., Osmond, J. (eds.): Designing Mobility and Transport Services. Taylor and Francis, London (2017)
4. Susilo, Y.O., Cats, O., Diana, M., Hrin, G.R., Woodcock, A.: Implementing a behavioural pilot survey for the stage-based study of the whole journey traveller experience. Transport. Res. Proc. **11**, 172–184 (2015). doi:10.1016/j.trpro.2015.12.015
5. Susilo, Y.O., Woodcock, A., Liotopoulos, F., Duarte, A., Osmond, J., Abenoza, R.F., Anghel, L.E., Herrero, D., Fornari, F., Tolio, V., O'Connell, E., Markucevičiūtė, I., Krithariotі, C., Pirra, M.: Deploying traditional and smartphone app survey methods in measuring door-to-door travel satisfaction in eight European cities. Transp. Res. Proc. (2016, forthcoming)
6. Susilo, Y.O., Cats, O.: Exploring key determinants of travel satisfaction for multi-modal trips by different travellers' groups. Transp. Res. Part A **67**, 366–380 (2014)
7. Susilo, Y.O., Abenoza, R.F., Woodcock, A., Liotopoulos, F., Duarte, A., Osmond, J., Georgiadis, A., Hrin, G.R., Bellver, P., Fornari, F., Tolio, V., O'Connell, E., Markucevičiūtė, I., Diana, M.: Findings from measuring door-to-door travellers' travel satisfaction with traditional and smartphone app survey methods in eight European cities. Eur. J. Transp. Infrastruct. Res. (2017, accepted)
8. Susilo, Y.O., Abenoza, R.: METPEX trial result. In: Tovey, M., Woodcock, A., Osmond, J. (eds.) Designing Mobility and Transport Services, pp. 126–146. Taylor and Francis, London (2017)
9. Diana, M., Pirra, M., Castro, A., Duarte, A.: Quality indicators for the whole journey experience. In: Tovey, M., Woodcock, A., Osmond, J. (eds.) Designing Mobility and Transport Services, pp. 147–157. Taylor and Francis, London (2017)
10. Diana, M., Pirra, M., Castro, A., Duarte, A., Brangeon, V., Di Majo, C., Herrero, D., Hrin, G.R., Woodcock, A.: Development of an integrated set of indicators to measure the quality of the whole traveller experience. Transp. Res. Proc. **14**, 1164–1173 (2016)
11. Diana, M., Duarte, A.,and Pirra, M.: Transport quality profiles of European cities based on a multidimensional set of satisfaction ratings indicators. Transp. Res. Proc. (2017, forthcoming)
12. Benedyk, R., Woodcock, A., Harder, A.: The hexagon-spindle model for educational ergonomics. Work **32**(3), 237–248 (2009)

13. Woodcock, A.: User centred transport design and user needs. In: Tovey, M. (ed.) Design for Transport: A User Centred Approach to Vehicle Design and Travel, pp. 21–69. Gower, London (2012)
14. Woodcock, A.: The whole journey experience through the lens of the hexagon-spindle model. In: Tovey, M., Woodcock, A., Osmond, J. (eds.) Designing Mobility and Transport Services, pp. 57–65. Taylor and Francis, London (2017)
15. Herrero, M.D., Gonzalez, E., Diana, D., Hrin, G.D., Anghel, L.E., Tolio, V., Brangeon, V., Arampatzis, S., Geeorgadis, G., Castro, A., Hoppe, M., Winter, M., Pirra, M., Woodcock, A.: D5.3: Testing the Indicators and Developing Benchmarks. METPEX Project Deliverable (2015)
16. Diana, M., Susilo, Y., Georgiadis, G., Pirra, M., Woodcock, A., Castro, A., Duarte, A., Pimentel, T.: D5.2: A comprehensive set of quality and accessibility indicators for transport services. METPEX Project Deliverable (2015)
17. Pirra, M., Diana, M., Castro, A.: Information provision in public transport: indicators and benchmarking across Europe. In: Dell'Acqua, G., Wegman, F. (eds.) Transport Infrastructure and Systems: Proceedings of AIIT Int. Congress. CRC Press, London, ISBN 9781138030091(2017)
18. Diana, M., Pirra, M., Duarte, A., Herrero, M.D., Tolio, V., Hoppe, M., Abenoza, R., Panea, L., Woodcock, A.: D5.4 A transport services quality and accessibility evaluation manual. METPEX Project Deliverable (2015)

From Task Analysis to Innovation

Denis Miglianico[1(✉)], Luc Moyart[1], and Mathieu Mouchel[2]

[1] Alstom, 48 rue Albert Dhalenne, 93482 Saint-Ouen, France
{Denis.Miglianico,Luc.Moyart}@alstom.com
[2] rHesuS Consulting, 251 Rue Joseph Louis Lagrange, 59300 Famars, France
Mathieu.Mouchel@rhesus-consulting.com

Abstract. An important part of activities of a tramway driver is to keep watching on the environment that includes pedestrians, bicycles, motorcycles or cars. Based on tramway driver's activities review, a preliminary analysis has been done to see how the new eco-driving task could affect existing tasks. The impact on the workload and the risk of errors was investigated. The method and experience presented explain how experiments were conducted in a simulator to assess the risk and to provide different physical means of information transmission based on various levels of abstraction. Different measures deliver assessment of effectiveness of driving. Based on these measures, new solutions are integrated and the results of head-up display implementation are presented.

Keywords: Task analysis · Workload · Head-up display · Eco-drive

1 Introduction

The paper describes the different steps achieved from the tasks analysis of a tramway driver to the integration of innovation inspired and derived from the analysis. Based on the methodology developed at the University of Hull [1] and improved to integrate the sequence of the tasks. The tramway driver's activities have been first analyzed with 60 drivers on three French networks (Valenciennes, Montpellier, Bordeaux) taking into account the internal factors (characteristics of drivers) and external determinants (cab and desk description, organizational contexts).

These observations have highlighted the importance of "macro-goals" within the realization of the real tasks of the drivers.

Some objectives are specific to the drivers (safety, comfort, respect of the times table) and others are directly related to the prescription (respect of the rules). Respecting all these goals requires both a much sustained attention to the external environment and a management of the behaviors of the tramway. Thus, the driver must at the same time pay attention to the front of cabin while supervising the speed of the tram on a display inside (sometimes out of its immediate field of vision).

Different techniques of analysis (verbalization during driving, peer review, video analysis) made it possible to highlight and to model, in the form of Descriptive Analysis Methods (DAM; see. After), real behavioral patterns (both in terms of information and actions on the Master Controller) according to the different situations of driving.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_99

The ergonomic analysis of the activity allows the creation of a knowledge base useful for the implementation of experimentation in a simulator. It's used to assess the impact of the newly developed systems on the activity and the cognitive resources of the driver.

2 Physical and Cognitive Tasks

After the bottom-up analysis based on the in-field analysis, a top-down methodology provides the formalization of driver's tasks. It's organized on three levels: strategic, tactical and operational levels (Fig. 1).

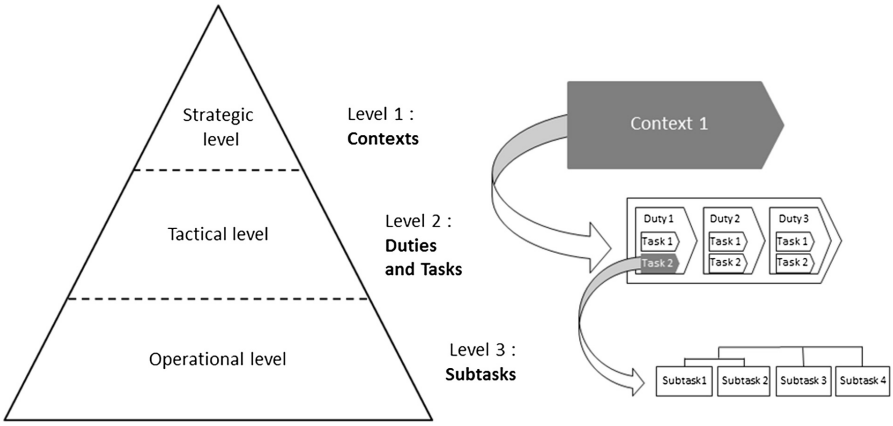


Fig. 1. Top down methodology.

The highest level is the strategic level; it includes the different contexts (nominal and degraded modes). The second level is the tactical level. It includes the duties (for example execute tests before departure) split into different tasks (for example test the brake, test the doors, The operational level lists to the various subtasks performed to reach the objectives of the tasks…).

The Descriptive Analytic Method (DAM) [7] used organizes the sequence of subtasks. The aim of this method is to provide a global view of the tasks realized by each actor in a graphical way. It allows a representation of all sub-tasks on a span, and an identification of those performed in parallel (Fig. 2).

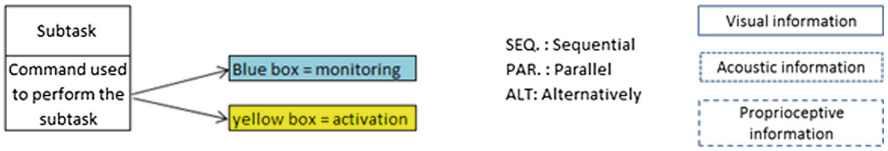


Fig. 2. Presentation of the sub-task

Links represent the relationships between subtasks. Subtask in blue means the driver monitor (an information, a state, …). Subtask in yellow means the driver execute an action.

Three different links between subtasks are possible:

a. Sequential: subtasks performed one after the other
b. Parallel: subtasks performed simultaneously
c. Alternatively: subtasks performed alternatively without logical sequence

To achieve the task presented in the Fig. 3, the driver:

I. Executes the subtask 1 by activating the command 1.
II. Then he performs the Subtask 2, monitoring the screen 1.
III. When Subtask 2 is finished, he manages subtasks 3 and 4 in parallel, activating respectively command 2 and 3,
IV. Then he manages the subtask 5 when subtasks 3 and 4 are finished, activating command 4.
V. To end up, he carries out subtask 6 or 7 depending on the situation, activating either command 5 or 6.

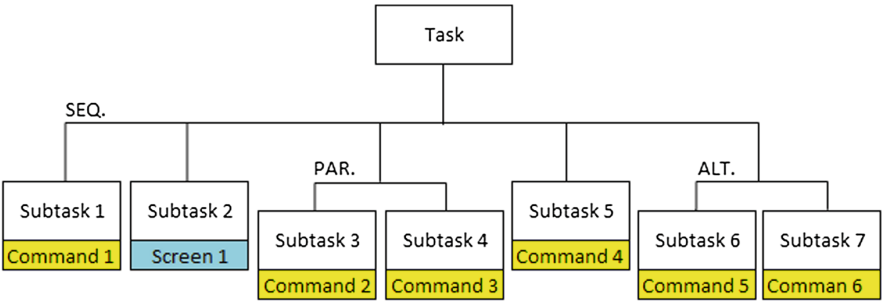


Fig. 3. Generic example of a description

Based on the task analysis (Fig. 4), the compilation of the different physical and cognitive analysis provides an estimation of the workload.

As the first goal of the driver is to maintain a high level of safety for passengers and the external environment, an important part of driver's activities is to keep watching on the environment (to anticipate as much as possible any braking). This is all the more difficult that there are lots of interaction with pedestrians, bicycles, motorcycles or cars and even some time with other trams like in the area of the railway station of Montpellier city.

It appears that during specific stage (leaving a station or arriving in a station), the driver has to manage all the information located in a wide field of view and in the same time he has to adapt the speed, anticipating pedestrian future location based on speed estimation and warn people using gong if needed. Therefore, when he is focused on

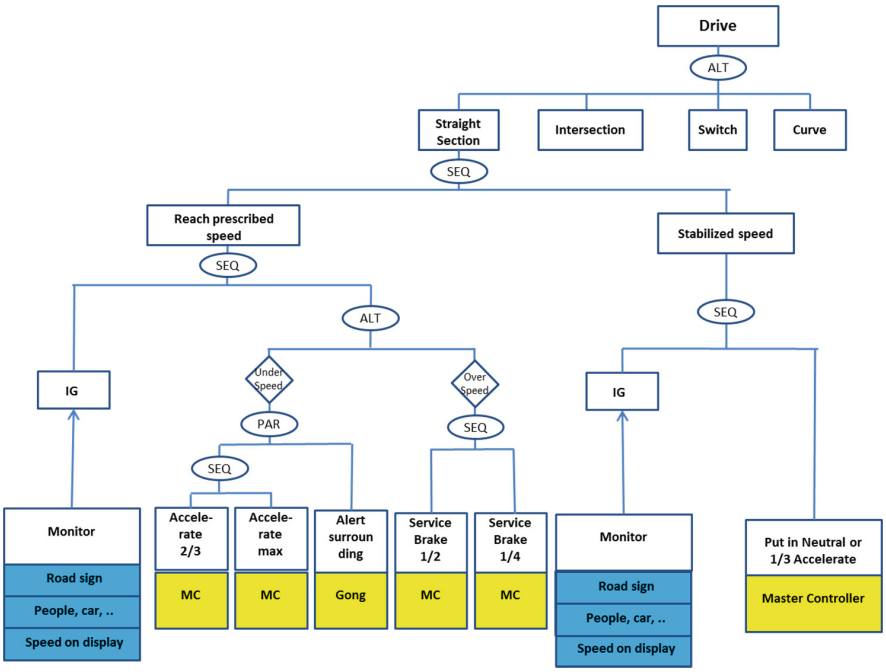


Fig. 4. The driving task analysis

external information he spent less time monitoring the speed, which leads to an increase of master controller movements (increase of the physical workload). On the other side, when he checks more often the speed, he switches more frequently from the inside to the outside increasing the workload and the stress.

3 Technical Solution vs Acceptable Workload

Before introducing a new prescribed driving task such as energy saving, the workload issue has been raised. Will drivers be able to manage physically and cognitively this additional task on top of other? Are some available technologies creating a possible hazard if the organization is requesting tough objective or are they able to reduce this risk?

To answer these questions related to the impact on activity of the addition of a new driving request (eco-drive), a study was carried out to describe what the eco-drive in the tramway environment is. Unlike other types of transportation mean (bus, for example), the tram eco-drive is only managed by the traction/brake master controller. Therefore the driver must follow an optimal traction and braking effort while respecting his other driving objectives (safety, comfort, respect of the rules, …). Indeed, although it is a driving goal, drivers told that energy consumption is not a priority. This secondary objective will be achieved only if the other objectives are met ("I would only do eco-driving if I'm sure my passengers are safe").

In order to save energy, a driver must have the cognitive resource to supervise the situation. He has to consider a lot of different information in the environment depending on the situation. It must also be able to control its traction/brake master controller in order to reach its initial objectives (need for a non-intrusive instruction). As a consequence, the designer has to limit the impact of the new eco-drive task on the cognitive load having it as far as possible non-intrusive (Fig. 5).

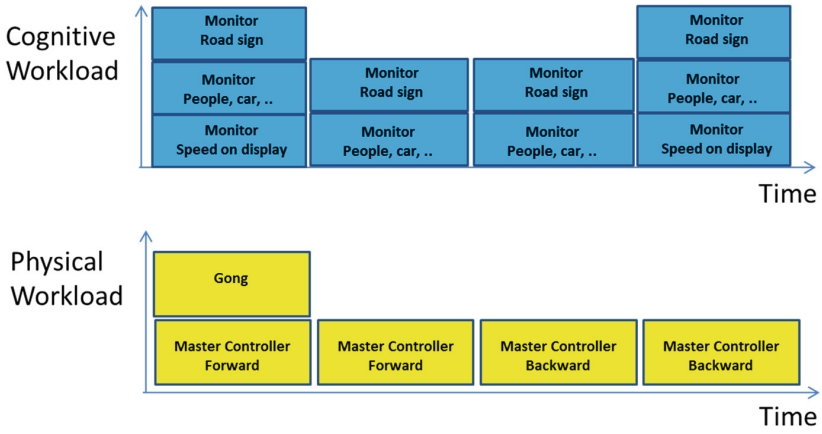


Fig. 5. Cognitive and workload analysis

Taking into account these two main needs, an eco-drive instruction has been developed. It has been designed as simple as possible (impact on the cognitive load) to leave the driver all its ability to control the system (non-intrusive). It presents to the driver a visual interface which simply indicates information about traction or brake performance in order to drive economically. In order to minimize the impact of this new directive on the cognitive load of drivers, the head-up-display technology was used (see below) It has the advantage to present information in the driver's main field of vision [5]. The traction/braking setpoint interface is shown in Fig. 6. When the driver is required to brake, an arrow pointing downwards is presented. The closer the speed is to the set point, the smaller the arrow is. When tram is at the optimum speed, an acronym "ECO" appears to him.

In order to evaluate the impact of this new instruction on the main objectives of the drivers but also on the cognitive load, a simulation experiment was carried out. It allows a test in an environment using a dynamic simulator incorporating a tramway cab and offering realistic 3D driving conditions. For example, if the driver uses the gong, then the pedestrians stops walking in general but sometimes he hesitates. The driver has to stop at each station, opening the doors, checking the passenger exchange process using the video display and then closing the doors. Some particular environment integrates situations where the driver's workload is high due to kids playing close to the track, drunk behind another tram or the surprising crossing of a pedestrian. The simulator integrates the FaceLab tool that measures the level of workload. A first set of

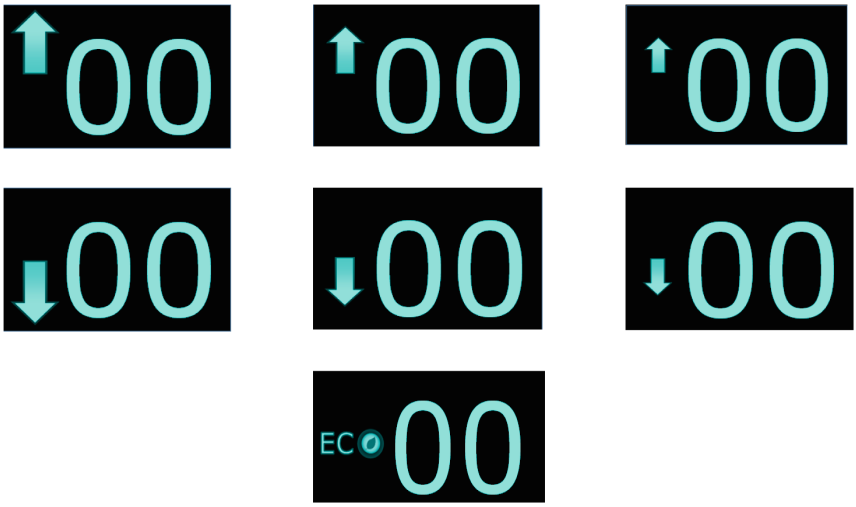


Fig. 6. The visual instruction for eco-drive

tests was conducted with 80 drivers coming from 10 French cities. The environment includes the simulation of a typical route of 15 km with 18 stations and 100 intersections. The surroundings cover different types of driver's interaction: from low-level (country-side) to high-level (downtown). The duration of the travel is 45 min (Fig. 7).

At the end of the experiment, an interview was conducted with the drivers to know their feelings on a set of elements (stress, fatigue, concentration, realism of the situation, …).

To illustrate the different possible types of information taking for managing traction and braking, 4 situations were designed (according to the level of abstraction and the origin of data described by the Cognitive Control Model [2]):

The impact of the new instruction on driving activity was measured using different indicators. Some are related to the driving objectives (safety, comfort, respect of the rules) [6]:

1. Anticipation of braking in meters.
2. Speeds at the beginning, at the middle and the end of the situation are measured in km/h.
3. Braking force: depending on the position of the master controller (from −1 maximum braking to +1 acceleration maximum).
4. Movement of master controller: in occurrence of movements/10 s.
5. Activation of the Deadman control: in occurrence of activations/10 s.
6. Use of the Gong: in occurrence of activations/10 s.

Regarding the impact of the eco-drive instruction on mental workload, an indicator based on FaceReader and FaceLab technology has been used. It evaluates pupillometry and activation. The evolution of the pupillary diameter is correlated with the mental load (which is used, for example, in the ICA indicator), and the activation is related to a degree of concentration which is correlated with attention.

- **Situation A**, the "Ambulance" situation involving the display of external information coming from the environment with a symbolic content. A large sign with an inscription "attention, exit of ambulance" is visible on the left of the cabin.

- **Situation B,** "situation safety mode" is aiming to generate the use of an internal (stored in the memory of the driver) information with symbolic content. To do this, drivers are required to read and retain a particular driving procedure just before the start of the experiment. This procedure does not exist in their activity. After stopping at the station "Marechal Leclerc", the instruction is given by audible announcement: "OCC for train 18, when the next R sign is reached, thank you to go in safety mode". After passing R sign, a new instruction is addressed to the driver in order to interrupt the safety mode operation.

- **Situation C,** "pushchair situation" consists in generating the use of external information with subsymbolic content. In this situation, the driver is traveling in an environment with multiple blind spots next to the tracks. Within this environment, a pushchair is visible suddenly because it is hidden behind garbage.

- **Situation D**, "children", consists in generating subsymbolic internal information. In this situation, the driver is in an unobstructed environment in which it is possible to detect with anticipation different elements: a school near the tracks and children playing around.

Fig. 7. Four situations designed according to the level of abstraction and the origin of data described by the Cognitive Control Model [2]

4 Solution Design

The limitation of the physical workload (movement of the head and of the hand on the master controller) is solved when the driver can maintain the gaze on the tracks. The use of a display in the main field of view is the solution that was requested to be designed. The head-up display technology has already been implemented in the aircraft or automotive industry. It has been tested in some research project (see [3]) but it has never been integrated into a real train driver's desk in commercial services.

A key success factor is, since the beginning, to integrate the standard and regulation requirement into the design. A preliminary visibility study provides a volume free of obstacle as requested by the STRMTG (French government regulatory unit in charge of Rolling Stocks authorization) [4]. The study requests to see a cylinder of 1100 mm height and 300 mm width at a distance of 1000 mm from the tramway. The test shall be conducted with 3 different numerical manikins. The 3 crosses showed hereafter present the location of the driver's eyes (Figs. 8 and 9).

The $\pm 5°$ angle of the combiner movement is calculated to cover all the range of driver's size. The size and the location of the non-transparent part of the head-up display is calculated not to hide a significant part of the cylinder (less than 4% according to the STRMTG document.

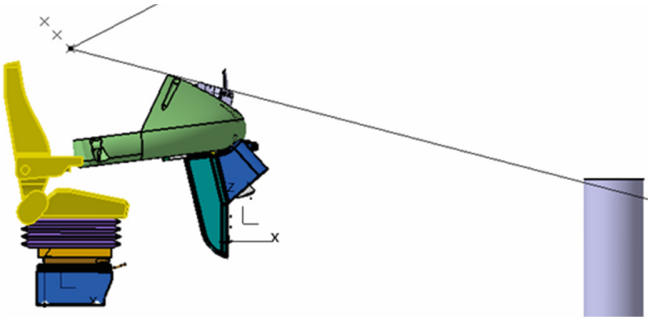


Fig. 8. Visibility for a 5th percentile manikin

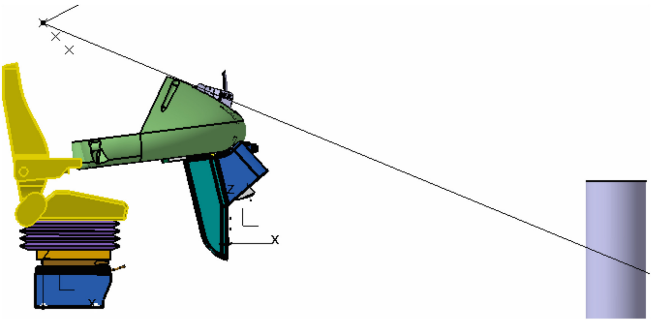


Fig. 9. Visibility for a 95th percentile manikin

5 Tests with Drivers

5.1 Methodology

The experiment carried out on the dynamic simulator with the 80 participants allows us to study the perceived impact of the presentation of the speed on the head up display but also the effect of the eco-drive instruction on the activity of drivers taking into account the objectives. For this assessment, two groups of drivers are defined:

- The control group: they use a standard desk with an integrated head-up Display. The HUD shows only the speed value at a perceived distance of 2 meters from him.
- The test group using the same desk but the HUD shows the speed value and the new eco-drive instruction.

5.2 Head-Up Display with Speed

We have analyzed interviews done after the experiments by using the verbalization of the drivers about speed displayed on the Head Up.

It is interesting to note the great homogeneity of the verbalizations (while the discourse was free). The following remarks are quantified with the proportion of drivers who have formulated them:

- The HUD makes me uncomfortable in my observation of the environment (1.3%)
- The HUD does not seem useful in addition to the speedometer (1.3%)
- HUD location could be improved (higher) (2.8%)
- The HUD allows me to keep my attention towards the external environment (97.2%)
- The HUD allows me to distinguish almost immediately an overspeed (45.8%)
- The HUD allows me to monitor my speed when usually I cannot do it (because my attention is focused outside) (27.8%)
- My eyes are less tired (30.6%)
- I think I have reacted more quickly to the dangerous situation (when the pedestrian crosses for example) than if I had watched my speed inside (13.9%)
- Hopefully we will soon have this technology in our cabins (93.1%)

All of these remarks to the open-ended question: "What do you think about the HUD to display speed?" show a perceived impact on the ability to keep the attention focused on the external environment. The HUD also appears to increase the reactive capacity of drivers while reducing the sensation of ocular fatigue for speed reading. Finally, a large majority of experienced drivers wish to have this type of equipment on board their tram.

The benefits of integrating HUD into the desk are shared by the vast majority of participants (93.1% would like to have it in their cabin).

5.3 Head-Up Display with Speed and Eco-Drive Information

In order to assess the impact of eco-drive information on driver's activities, we use the 6 indicators and the 4 situations (see Chap. 3). Previous research has shown that tramway drivers use particular information depending on the situation encountered. The impact of a visual instruction of eco-drive in the HUD has been considered.

A first set of conclusions shows that for parameters 1, 2, 6, adding eco-drive information has no impact. For parameters 3, 4, 5, results are presented hereafter (Table 1).

The control group brakes more in situation B (+11%; Least Square Difference: $\alpha < .0002$) and in situation D (+12%; LSD: $\alpha < .005$). The stress of the driver is higher and the comfort of passengers may be degraded.

The control group is moving more frequently the master controller (increase of 10 movements during 10 s in A situation, 4.5 movements in B situation, 19.3 movements in situation C and 35 in D situation). The Head-Up provides less physical workload and muscular fatigue.

The control group activates more often the deadman (5 additional actions in 10 s and LSD: $\alpha < .006$ in situation A; 4.3 additional actions in 10 s and LSD: $\alpha < .02$ in situation; 8 additional actions in 10 s and LSD: $\alpha < .008$ in situation C; 3.2

Table 1. Results for parameters 3, 4, 5.

Parameter		3	4	5
Situation		Braking force (% pos)	Movements of the master controller (occurrences)	Activation of deadman (occurrence/10 s)
A	Mean	0.63	43.90	11.75
	SD	0.31	17.44	11
A with HUD	Mean	0.61	33.88	6.66
	SD	0.31	9.76	3.53
B	Mean	0.53	33.85	10.35
	SD	0.29	11.48	6.26
B with HUD	Mean	0.31	29.36	6.01
	SD	0.27	18.37	3.20
C	Mean	0.33	48.74	16.87
	SD	0.38	33.24	15.37
C with HUD	Mean	0.25	29.44	8.89
	SD	0.29	14.14	5.31
D	Mean	0.60	49.28	9.78
	SD	0.27	29.20	5.68
D with HUD	Mean	0.29	14.14	5.31
	SD	0.36	24.74	6.56

additional actions in 10 s and LSD: $\alpha < .005$ in situation D). As for the master controller, the HUD provides less physical workload and muscular fatigue.

The study of the impact of the addition of the eco-drive instruction in the HUD does not reveal any difference between the two experimental groups in an urban driving. Indeed, the average workload of the control group (without an eco-drive instruction) is 0.316 (SD 0.045) and that of the group using the visual instruction is 0.338 (SD 0.043).

6 Conclusions

The tests in simulator confirm the improvement brought to the driving activity (external attention, ocular tiredness, reaction time, …) by the speed shown on a Head Up Display. The use of this technology for eco-drive instruction impacts positively driver and the passengers in term of driver's workload or comfort. Additional measures (not shown in this paper) show that the number of head movement is decreased. The advantage of HUD, which is to read the speed of the tram without having to move the head or refocusing view to observe the surrounding, is confirmed.

During the project, the drivers show a strong involvement in the tests. The HUD is adopted by a majority of drivers. It enables the launch of an in-field test in 2017.

References

1. Annett, J., Duncan, K.D., Stammers, R.B., Gray, M.: Task Analysis. Her Majesty's Stationery Office, London (1971)
2. Hoc, J.-M., Amalberti, R.: Cognitive control dynamics for reaching a satisficing performance in complex dynamic situations. J. Cogn. Eng. Decis. Mak. **1**, 22–55 (2007)
3. Davies, K., Hinton, J., Thorley, P., Buse, D.: A feasibility study of head-up displays in driving cabs. Final Report for Rail Safety & Standards Board: T513 (2007)
4. STRMTG. Sécurité des Postes de Conduite des Tramways. http://www.strmtgdeveloppement-durable.gouv.fr/IMG/pdf/Guide_Securite_des_postes_TW_version_3.pdf (2015)
5. Davies, K., Thorley, P., Thomas, M., Gibson, H.: Feasibility of head up displays in driving cabs. In: Rail Human Factors around the World: Impacts on and of People for Successful Rail Operations, 465 (2012)
6. Mouchel, M., Naveteur, J., Miglianico, D., Anceaux, F.: Contrôle cognitif et prise de décision en environnement dynamique. Cas particulier de la traction et du freinage chez les conducteurs de tramway. In: Bonnardel, N., Pellegrin, L., Chaudet, H. (eds.) Actes du Huitième colloque de psychologie ergonomique EPIQUE 2015, Arpege Science Publishing, Aix en Provence, pp. 201–210, juillet (2015)
7. Sebillotte, S.: Methodologie pratique d'analyse de la tache en vue de l'extraction de caracteristiques pertinentes pour la conception d'interfaces (Doctoral dissertation, INRIA) (1994)

Road and Rail: Driving Automation and Human Factors Issues

Three Driver and Operator Behaviour Models in the Context of Automated Driving – Identification of Issues from Human Actor Perspective

Pirkko Rämä(✉) and Hanna Koskinen

VTT Technical Research Centre of Finland Ltd,
Vuorimiehentie 3, P.O. Box 1000, 02044 VTT, Finland
{Pirkko.Rama, Hanna.Koskinen}@vtt.fi

Abstract. The study aimed to systematically define human issues related to automated driving. Three theoretical approaches were applied into two automation scenarios, low and high, in urban and motorway driving. The analysis encompasses five use cases in both environments, and several contextual circumstances. The list of approximately 13 human issues was provided. All theoretical approaches showed to be useful in defining the issues, and the views complemented each other. The core task analysis, not previously showed to be promising model for analyzing automated driving, and should be further developed. The other approaches applied were the traditional driver behavior theories called as motivational theories and the model for human actor operator as information processor.

Keywords: Automated driving · Driver behavior · Theoretical models of human behavior

1 Introduction

Increasing automation in road transport has revealed several challenges from drivers' and human actors' point of view. Just to mention few, it has been expressed concern about, how to cope with mixed transport where automated and non-automated cars or automated cars and vulnerable road users (VRUs) should interact with each other; How to handle the shift from automated to manual driving; how to inform the driver about the status of automated vehicle or traffic situation? What would be the implications for fluency of traffic flow? Automation has been reasoned to rule out the most risk-bringing component – human error – and thereby increase traffic safety dramatically. However, it is known, that automation has not solved all safety problems in the transport modes where it already has been applied widely. Thus, despite recent advances in automatic driving, the interaction between the drivers and the automated vehicles is still relatively poor understood. The present study aimed to analyse and list human issues related to the shift towards automated driving. In analysing the human behaviour in the context of automated driving, some well-known driver behaviour theories were utilized. To extend the analysis, also human behaviour models developed for the analysis of operator work in industrial work environments was inquired into.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_100

For our analysis, we broke down automated driving journey into use cases in two environments, (1) urban or city area and (2) motorway. In this paper we use term 'use case' to indicate part of the journey in motorway driving or driving in urban environment, such as 'merging the motorway', 'driving in link section', or 'driving in intersections'. In addition, we defined several contextual circumstances for driving such as traffic situation (e.g. free flow and jam), traffic composition, weather and lightning conditions and additional details in the environment.

Driver behaviour was analysed in two different level of automation scenarios, 'low' and 'high'. Already now, some cars with level 2 (refers to SAE classification [1]) functionality and applications are seeing in the market. Furthermore, rapid advances in the field makes it likely that in the future the automotive industry is going to directly launch systems operating in level 4 (i.e., automotive system monitors and controls) rather than level 3 systems (shared responsibility of human and vehicle in monitoring and control) that represent many challenges of human-machine interaction. In our analyses, 'Low scenario' refers to the situation in which level 2 automation (refers to SAE 2014 classification [1]), and 'High scenario' to the level 4 automation. Both automation scenarios were described in detail for each use case in each driving environment (i.e., urban and motorway driving).

The three theoretical approaches of human behaviour explored were: (1) the motivational driver behaviour theories, (2) the human operator as information processor, and (3) the practice theory approach and core task model. Next, short reviews of the selected approaches are given. After presenting other aspects of the scenario design and the procedure, the results are shown separately for the urban and motorway environments. In the end of the result section a list of identified human issues presented. Finally, we discuss the role of theories, make conclusions about relevant human issues, and present some ideas for the future work.

2 Human Behavioral Models

We selected three theoretical approaches which provided us three different frameworks to study and discuss human issues of automated driving. The first group of theories comes from the traditional traffic psychology, we called it 'motivational theories' (as e.g. [2, 3]). The second model [4] sums up traditions from ergonomics and cognitive psychology and analyses the human information processing. The model and its concepts have been applied both in work environment and in transportation including aviation with experience of automation.

The third model, practice theory approach and core task model [5] has before this not applied in transport. Its origin is in studying work in diverse industry and work contexts such as nuclear power plants, medical operating rooms, precision farming.

The classification of theories is approximate, but tries to catch some of the most relevant features of the models. We acknowledge that the selection may leave out several approaches and theories, which would have been worth including.

2.1 Motivational Theories

The 'motivational models' of driver behaviour focus on how drivers perceive risk of a crash while driving. Koornstra [6] suggested that the three models (below) would be special cases of the risk-adaptation theory. The origin of the theories is usually traffic safety, and the intention was to understand human behaviour and define measures to improve safety.

The zero-risk theory [7] stresses the large variety of motives drivers may have. The theory suggests that drivers try to keep their perceived risk level at zero, and when the threshold is exceeded some compensatory actions are taken to bring the risk level back to zero. With experience, driving becomes a habitual, largely automatised activity in which risk control is based on maintaining safety margins [8].

The "threat avoidance model" by Fuller [9] posits that driver behaviour is determined by two conflicting motivations, making progress towards destination and avoiding hazards. Fuller also emphasises the role of learning—with exposure, drivers learn to recognise emerging risks and thereby make anticipatory avoidance responses. Later, Fuller [10] has specified risk homeostasis as a special case of task difficulty homeostasis.

According to Wilde's risk homeostasis theory [11], drivers aim driving at a constant, tolerated level of risk. Wilde [11] stated that when a safety measure is applied, the driver population adapts its behaviour so that the risk will not change, known as risk compensation. Introduction of risk compensation stresses the importance of studying the long-term impacts of any measure and the concept of behavioural adaptation.

2.2 The Human as the Information Processor

The second group comes from ergonomics; it is more general and deals with the human being as the information processor. The theoretical models in this group focus on perception and information processing of a human actor. Based on earlier findings and literature, Wickens and Hollands [4] summed up and provided a general model of human information processing stages.

The model presents the stages—sensory processing, perception, response selection, and response execution—all utilising attention resources. Furthermore, the whole chain is interacting with memory (working memory and long-term memory). Feedback is provided from the response execution back to the sensory.

The essential feature in the model is that the attention resources are limited, evident especially in dynamic and complex tasks. In some driving contexts, there may be too many elements in the environment and parallel processing of the information, which increases driver workload. People seem to form an internal representation, a mental model of the properties of the events when the situation or elements of a situation are repeated. The mental models are guiding perception.

Endsley [12, 13] suggested that good situation awareness is important in selecting the most appropriate mental models for a task. She presented the concept of situation awareness (SA) as a crucial construct in operators' decision-making in dynamic and complex environments. It incorporates an operator's understanding of the situation as a

whole and forms the basis for decision-making. Endsley [12] defines SA as three successive levels: (1) perception of elements in the environment, (2) comprehension of their meaning, and (3) projection of their status in the near future—as "knowing what is going on."

With experience and consistent practice, the automaticity of car driving increases. Automaticity is characterised as fast, effortless processing, whereas controlled processing refers to slow serial and effortful processing [14].

2.3 Practice Theory Approach and Core Task Model

Various theoretical models have been introduced to explain the human behaviour in complex safety-critical work domains such as aeronautics or nuclear power production. The practice approach focuses on examining how and according to what logic the human actor and the surrounding environment, including organisational structures and tools, get organised into arrays of action with regard to the purposes of the work. [5, 15] These arrays of action may express and be explained by their meanings, and they reveal the value of the situation and environment for the human actor.

The practice theory approach revises the unit of analysis and aims at dissolving two persistent dichotomies that exists in the human behavioural and social sciences, i.e. the division between the individual and societal and between the organism and the environmental determination of conduct [5]. The first dichotomy, that is, determination of human conduct either by societal and organisational structures or by individual intention prevents reaching an understanding of the full resolution and nature of the current problems in work and everyday life. Moreover, viewing the two separately does not capture all aspects of what is called a socio-technical system. Instead, it is necessary to step from one level of a system (i.e., the individual) to the other levels (i.e., those of the group/team and organisation) and view the elements and their interactions from a more systematic point of view [16]. The second dichotomy, which has a major influence even today, is the division between the organism and the environment, which are considered two separate elements. This dualistic view examines the interaction mainly as an information flow between the organism and the environment and may explain the internal structures of the two elements. However, it cannot explain the content of the relationship between them and their meaning. Alleviating the above division yields a better understanding of how certain elements of the human organism and the environment accomplish functions together for reaching a shared goal [17]. An ecological psychologist, James Gibson [18] created a theory of direct perception suggesting that meaning is inherent in the organism-environment system and the environment affords certain action possibilities to the human organism that has certain set of capabilities to perceive and take these possibilities into use. This theory also, challenges the previous dualistic conceptions of the relation between a human organism and the environment as distinct categories by turning the attention to the total ecosystem of the two. The concept of practice may provide a comprehensive way to define work and, it provides a basis with which to analyse the core task of complex work activity [5].

The core task design approach aims at developing complex work systems, i.e. work practices and tools [5]. Concretely, in core task modelling the practice is understood as

activity as defined in the cultural historical activity theory, and exploited, e.g., in the HCI tradition [19]. Activity theory emphasizes the contextual understanding of work, and therefore place the object of activity in a central role. The object of activity is that part of the environment that has become the focus of human actor's attention and thereby a part of the human-environment system. Thus, the object of activity evidently affects how acting is structured. The core task [5] provides a functional perspective to work (i.e., object of activity) and refers to "the generic developing content of the work and expresses itself as joint functions emerging from the meeting of the human organism's resources with the possibilities and constraints of the environment for reaching certain global objectives of work activity" [5]. In analysing the core task, the work activity is generalised beyond the individual work tasks. Thus, the more generic aims and possibilities of certain work are comprehended and, are expressed as core tasks of that specific work activity. On a general level, the core task defines how the human actor and the environment can become organized to collaborate in a certain work context.

3 Method

3.1 Procedure

The method was qualitative analyses of the automation use cases, utilising the three theoretical models described above.

After identifying the theoretical approaches of interest, we defined five use cases for both studied environments, that is urban and motorway (use cases are listed in tables in the Sect. 4). For each use case, we provided the descriptions from the three different theoretical views. The focus was on describing what would be the change that automation would bring with it, compared to a situation without automation.

In the analysis we first focused on a 'basic case', which was seen as not a very complex driving task – driving in car following situation without any additional difficulty due to the contextual circumstances. Following the basic case, the degree of difficulty in the driving task was increased by adding contextual variance (e.g., challenging weather condition) in the situation. As we had three theories, we ended up with a $3 \times 6 \times 12$ matrix with 216 cells for the descriptions in urban driving. The data for motorway driving was documented in the matrix of 162 cells ($3 \times 6 \times 9$). In addition, many contextual circumstances had additional specifications to be considered, such as traffic lights (green; yellow; red; blinking; off).

Each use case was described using the terms, main concepts, postulates and propositions of the specific theoretical approach. Features of behaviour not addressed in the current theory were not included.

3.2 Definition of Automation Scenarios and Use Cases

Two different automation scenarios, SAE-level 2 'low' and SAE-level 4 'high' were considered. In low scenarios, the driver is responsible at all times for monitoring and overall control of the vehicle. In high scenarios, the vehicle is controlled automatically

but in some special situations, manual takeover is demanded (Table 1). All scenarios assumed 50% penetration in vehicles on the road.[1] Connectivity between vehicles (v2v) and to the infrastructure (v2i) was assumed only in the high scenario.

Table 1. Specification of level of automation in the two focused driving environments

Level of automation	Urban	Motorway
Low	• Automatic Cruise Control (ACC) Stop and the Go app based on following the car ahead • Automatic Emergency Braking system (AEB) • P-assist in which driver uses the gas pedals and brakes and the assist takes care of the lateral control	• ACC • Lane Keeping (LK) • AEB
High	• Automated driving system 'Urban Autopilot' (AP) based on 3D maps, high quality sensors and laser radar • Automatic P-assist, which takes the car into and out of the parking space, even the driver would be outside the vehicle • AEB	• Highway AP, driver takes into use on the ramp. Countdown base on driver set destination to prepare for the manual control e.g. 3 min before exiting • AEB

In the urban low scenario, it was roughly estimated that the automation support would be available 25% of the time and in the urban high scenario approx. 60% of time. In low scenario the support would not be available when turning, overtaking, partially when no vehicle to follow, sudden unexpected situations, fog, snow and when not willing to use the automation. The nature of non-supported situations was quite similar in the high scenario. In the motorway low scenario, automated systems would be in use approx. 70% of the time. Whereas, in the motorway high scenario the automated driving support would be most continuous, the automation would handle all driving tasks also in winter conditions 95% of time. The automated vehicle would drive automatically to the side of the road in very adverse conditions or in an unexpected situation.

We focused on car driving in urban and motorway environments. The urban environment was defined as a more suburban area that was clearly city rather than countryside. The motorways were 3 + 3 lane highways. Driving journey in these two environments was broken up into several use cases (e.g., leaving the parking place). The use cases are listed in the tables in the next chapter.

[1] The analysis did not make any specific assumptions regarding the non-equipped vehicles regarding ITS support available. In some degree, however, it was assumed that in the high scenario, all vehicles could have some additional ITS or automation support.

4 Findings of Human Issues in Automated Driving

The list of human issues was created based on statements as indicated in the Tables 2 and 3.

Table 2. Examples of results in urban driving. (L2 = low level of automation, L4 = high level of automation. Numbers refer to the theories where the statement coms from 1 for motivational theories, 2 for information processing and 3 for core task model)

Urban use case	Examples of human issues
(a) Planning of the trip	L2: Introduction of low-level automated systems would slightly increase the share of trips made by own cars (1). The workload would be experienced as smaller (2) and the uncertainty less (3)
	L4: The share of travelling by own car could decrease (1). Still, due to increased comfort, low workload and the possibility to do parallel secondary tasks (2)
(b) Parking or leaving the parking place (or yard)	L2: Favoring a fluent driving style might lead to a more rapid performance when leaving the parking place (1). Visual performance is not supported by P-assist. Consequently, an increase in crash risk could be observed (2). P-assist might support the driver in performing sequential rather than parallel actions (3)
	L4: The driver can be involved in other activities and, situation awareness for car driving declines (2). With time, the driver's ability to park manually declines, and leaving the slot and parking manually can be experienced difficult (1, 2)
(c) Driving on a link section	L2: An increase in the mean speed of the traffic flow is expected (1). Doing secondary tasks is questionable, because visual demand remains, and the execution of most secondary tasks may demand the same visual capacity and load the same channel as the driving task (2). The complexity of the task may increase with automation, as the driver may need to shift from supported to non-supported mode quite frequently (3)
	L4: Speeding no longer occurs, as the maximum speed is set by the automated system (1). With only few opportunities for manual control, the manual driving task becomes more difficult and the workload heavier (2)

(*continued*)

Table 2. (*continued*)

Urban use case	Examples of human issues
(d) Driving at intersections	L2: The driver prefers to follow the vehicle in front but may need to give way, stop or turn and interrupt following (1). Situation awareness is good, as a transition is expected at an intersection (2)
	L4: Intersections are places where the support is most typically interrupted, which may increase subjective risk, especially if encountering pedestrians and cyclists. Also, other circumstances may evoke feelings of risk, such as hazardous weather conditions, darkness, or information on an incident. As a compensatory action, the driver may shift to manual control (1)
(e) Navigation (during the trip)	Fluent roads are favored, but also main roads due to AD support if experienced as increasing fluency (1). The fastest route is preferred (1)

Table 3. Examples of results in motorway driving. (L2 = low level of automation, L4 = high level of automation. Numbers refer to the theories where the statement coms from 1 for motivational theories, 2 for information processing and 3 for core task analysis.)

Motorway use case	Examples of human issues
(A) Planning of the trip	L2: The workload would be experienced as smaller (2) L4: Knowledge about traffic situation on the journey may have big potential in optimizing achieving the overall goal of the journey (3)
(B) Merging	L2: If not clear benefic perceived and still a lot interaction with the AD demanded, may lead to the use of AD systems more restricted parts of the journey and more polarized manner (3)
	L4: Over time, or in specific contextual conditions, the automation setups may be adjusted, e.g. reduced following distance or the speed increased (1)
(C) Driving on a link section	L2: The driver can divert more attention to secondary tasks. Doing so may also cause a reduction in SA. Forming new kinds of mental models is required in order to switch between supported and non-supported driving (2). The interaction between the human driver and the AD systems becomes more central (3). Monitoring the driving task may be challenged by the fact that the physical activity related to the driving (i.e. steering, accelerating) is handled by the AD systems and the driver is no longer so holistically involved in the driving task (3)
	L4: When the AP can be fully utilized, the driver becomes a passenger; i.e. the core task changes completely (3)

(*continued*)

Table 3. (*continued*)

Motorway use case	Examples of human issues
(D) Leaving the motorway section	L2: The driver may try to delay switching to manual driving as long as possible due to limited AD support in the leaving (1). In the leaving-the-motorway task, all models highlight transformation from one driving mode to another (1,2,3)
	L4: The shift from secondary back to primary task is critical. The countdown is central to defining how quickly and well the driver comprehends the situation and may take the control. (2) How well the driver knows the functionalities and the limits of the AD systems and how well they are learned to use in the driving task is critical (3). Slowing down to the appropriate speed may be challenging after driving at highway (1,2,3)
(E) Navigation (during the trip)	L2: Navigators have become somewhat ubiquitous in the car-driving context, and people have learnt to trust them quite a lot (3) L4: The level of navigation support should consider the driver's knowledge of the route selected and the prior driving experiences (3)

4.1 Summary – A List of Issues Identified

Based on the analyses, we provide the following summary of human related issues:

1. The share of trips made by personal cars can be increased with L2 automation. The offering of public transport services and quality of service is critical with L4 automation, if the intention is to decrease the amount of car driving if needed for environmental or efficiency reasons. Over time, and when used for automation support (in L4), drivers may consider other modes than cars in adverse weather conditions assuming manual driving.
2. Fastest routes are favoured by drivers when setting the route. P-assist systems may encourage people to park their cars in busy city centres. Use of L2 P-assist leads to more fluent parking. In case of any lack or error in the detection system this could increase the risk.
3. The role of the car as fulfilling the extra motives decreases. The time allocated for traveling gets new content and can be used for new activities. The effect on the use of own cars is not known.
4. With automation there will be more space for information, and the risk of distraction is not as relevant as before. Several types of information are needed: real time information to support detection; real time information to anticipate; general guidance to introduce new automation systems; information to support creation of new appropriate mental models in parking and driving; information or education pertaining to manual driving skills.

5. Communication when one wishes to enter the traffic flow would be challenging, especially when trying to enter a queue. Connectivity of vehicles would facilitate entering but seems far in the future.
6. Communication with passing cyclists when leaving a parking space calls for specific attention. New arrangements of city space and roads may be needed.
7. Learned automacity skills may no longer be valid, and the driver needs to develop new mental models and atomicities. It is assumed that crash risk may increase in partial automation in L2 scenarios in this learning phase.
8. A lower workload is expected, leading to complete secondary tasks already with a low level of automation when the driver has the responsibility to monitor the situation and control the vehicle. This may lead to an increase in crash risk, specifically in exceptional, unexpected situations.
9. Following distances are becoming shorter. Mean speed may increase with low automation. Factory installations can influence this development.
10. Use of automation calls for improved communication between parties, and this should be seen as a specific design task.
11. Limited support in low automation may be confusing and influence use.
12. Trust in automation develops with experience. Even a few serious crashes in an automated system would be enough to reduce trust.
13. With low automation, drivers may not be willing to give way to other participants (VRUs, trams, cars) if this implies giving up familiar automated modes of driving (in a car-following situation).

5 Discussions

In the analysis, we followed the journey from planning to leaving the parking place, driving through an urban area, merging and driving on motorway, until leaving the motorway. Introduction of automation seems to influence and change every phases in car driving. We defined the list of human issue. There were some individual topics in the list but four main categories are (1) the strategic decisions, (2) role of information, (3) interaction and communication and (4) traffic safety.

Strategic Decisions. Already when planning the trip, very important strategic decisions [20] regarding the whole journey are made, influencing the driving tasks in the coming trip. Especially, in the low-level scenario of automation it seems that the effects may be negative: more car traffic in general, and specifically in the city centres. This could result reduced traffic fluency and safety and health problems if it also means reduction in physical exercise of people.

In the high scenario, the trend may be similar but there some factors pointing also to the other direction. E.g. in adverse weather conditions where one could encounter the situation in which manual driving is needed the drivers could favour more sustainable bus transportation. The quality of the public or shared transport service will be critical in influencing peoples' mode choices and sustainability.

Regarding the routes, most fluent routes would be preferred according to the analysis. However, in the high scenario, route choice would at least partly be part of the

planning and this could make more possible to offer alternative routes, based on other criteria such as safety or environmental benefits.

Role of Information. Information was identified to be important. Automation is assumed to give space for a more efficient informing of the driver when driving the car, as distraction would not be similar issue as currently in manual driving. Information would be needed to introduce the automation in cars – drivers should be at least in some degree being aware of the automated system, its functionalities and current state. The driver education will be renewed, but probably some updated training can be needed to support travellers to create new mental models for handling with the automated vehicles, and later to pertain the driving skills.

It also seems that the role of information may become important in the actual driving situations – real time information may be needed to support anticipating of exceptional situations, detection and interpretation of elements in the environment.

Communication. Communication and interaction is a particular process going on in road traffic. An example is entering a roundabout which can be seen as an end result of driver learning process and regards fine adjustment of how to react to the movements of other parties. Electronical connectivity of the vehicles (v2v) and automatic optimization could be tools for organizing communication needed between vehicles. Especially, in low-scenario connectivity was not assumed. However, also in high-scenario communication cannot be built on connectivity until the penetration of automation would be very high. Therefore, this topic is going to be actual for a long time both between cars and between automated cars and vulnerable road users not connected.

Interaction between parties is typical for road traffic. For this reason, e.g. examples from aviation are not able to address the issue. Also the traditional driver behaviour theories are quite individualistic. Complementing views have been looked from social psychology [21]. The core task analysis also explicitly discusses collaboration aspects, which can be a topic for further research.

Traffic Safety. The new phase of driver learning may involve increase in risk before the system is adapted to high automation. The analysis suggests that increase in risk is especially expected in low-scenario of automation. The increase of risk will be significant if at the same time when the driving task changes, it possibly becomes more complex (for a while) and the driving speeds are increased. The overcompensation of advantages of the systems can also lead to accomplishing secondary tasks, declining the situation awareness and increase risk.

Role of Theories. All models were useful in defining the human issues. The amount of issues was somewhat more from the information processing and core task analyses. However all theories had its' own value, and the different models exposed different aspects. The motivational theories has the advantage that they are directly linked to measures such as speed and driving distances which creates direct link to the basic concepts of traffic flow and traffic technology. The relationship with speed to traffic safety and other goals in road transport than e.g. situation awareness. The concept is much more fuzzy in that sense.

Issues Not Covered. We compared the findings with the framework of nine impact mechanisms of intelligent transport systems on driver safety [22]. The approach has been recently applied in automated driving [23]. It seems that the analyses covers best the direct effects for traffic safety and personal mobility, and indirect effects on traffic safety. The less covered areas were effects on exposure and consequences. The analyses covered in some degree non users and mode choices. Regarding impact areas efficiency and environmental effects were less well covered. All these create need for further analyses.

6 Conclusions

A list of human issues was provided. Many of the issues have been taken up and discussed in the literature before. However, the theoretical models provide a context for the individual issues, providing new insight into the topics. We see this as valuable when finding measures how to react to the topics raised.

Further development is needed to better cover some topics such as communication between the parties in traffic, education and information delivery to the drivers and other participants in road traffic.

All approaches were useful in defining the problems, and the views complemented each other. The core task analysis, not previously applied in road transport showed to be a promising model and should be further analysed and applied to this context.

Acknowledgements. Appreciation is extended to the research programme entitled Traffic Safety 2025. The current members of Traffic Safety 2025 include Finnish Transport Agency, Finnish Transport Safety Agency, Nokian Tyres plc., VTT Ltd.

References

1. SAE: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Standard J3016, issued 2014-01-16. Society of Automotive Engineers (2014)
2. Ranney, T.A.: Models of driving behavior: a review of their evolution. Accid. Anal. Prev. **26** (6), 733–750 (1994)
3. Vogel, K.: Modelling Driver Behavior – A Control Theory Based Approach. Institute of Technology, Linköping University, Linköping (2004)
4. Wickens, C.D., Hollands, J.G.: Engineering Psychology and Human Performance, 3rd edn. Prentice Hall, Upper Saddle River (1999)
5. Norros, L., Savioja, P., Koskinen, H.: Core-Task Design. A Practice-Theory Approach to Human Factors. Morgan & Claypool Publishers, San Rafael (2015)
6. Koornstra, M.J.: Risk-adaptation theory. Transp. Res. Part F **12**, 77–90 (2009)
7. Näätänen, R., Summala, H.: Road-User Behaviour and Traffic Accidents. North-Holland, Amsterdam (1976)
8. Summala, H.: Risk control is not risk adjustment: the zero-risk theory of driver behavior and its implications. Ergonomics **31**, 491–506 (1988)

9. Fuller, R.: A conceptualisation of driving behaviour as threat avoidance. Ergonomics **27**, 1139–1155 (1984)
10. Fuller, R.: Towards a general theory of driver behavior. Accid. Anal. Prev. **37**, 461–472 (2005)
11. Wilde, G.: Target Risk. PDE Publications, Toronto (1994)
12. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Fact. **37** (1), 32–36 (1995)
13. Wickensin, C.D.: Situation awareness: review of Mica Endsley's 1995 articles on situation awareness theory and measurement. Hum. Fact. **50**(3), 397–403 (2008)
14. Schneider, W., Schiffrin, R.M.: Controlled and automatic human information processing: I. Detection search and attention. Psychol. Rev. **84**, 1–54 (1977)
15. Schatzki, T.R., Knorr-Cetina, K., Savigny, E.V.: The Practice Turn in Contemporary Theory. Routledge, London (2011)
16. Engeström, Y.: Expansive Learning. Orienta, Jyväskylä (1987)
17. Järvilehto, T.: The theory of the organism - environment system (I): Description of the theory. Integr. Physiol. Behav. Sci. **33**(4), 317–330 (1998)
18. Gibson, J.J.: The Ecological Approach to Visual Perception. Hougton Mifflin, Boston (1979)
19. Kaptelinin, V., Nardi, B.: Activity theory in HCI: Fundamentals and Reflections. In: C.J.M. (Series Ed.) Synthesis lectures on Human-Centered Informatics, vol. 5(1), pp. 1–105 (2012)
20. Michon, J.A.: A critical view of driver behaviour models. What do we know, what should we do? In: Evans, L., Schwing, R. (eds.) Human Behaviour and Traffic Safety. Plenum Press, New York (1985)
21. Ajzen, I., Fishbein, M.: Understanding Attitudes and Predicting Social Behavior. Prentice-Hall, Englewood Cliffs (1980)
22. Kulmala, R.: Ex-ante assessment of the safety effects of intelligent transport systems. Accid. Anal. Prev. **42**(4), 1359–1369 (2010)
23. Innamaa, S., Smith, S., Barnard, Y., Rainville, L., Rakoff, H., Horiguchi, R., Gellerman, H.: Trilateral impact assessment framework for automation in road transportation. Draft, 4 January (2017)

The Right Moment for Braking as Informal Communication Signal Between Automated Vehicles and Pedestrians in Crossing Situations

Matthias Beggiato(✉), Claudia Witzlack, Sabine Springer, and Josef Krems

Department of Cognitive and Engineering Psychology, Chemnitz University of Technology, Wilhelm-Raabe-Straße 43, 09120 Chemnitz, Germany
{matthias.beggiato, claudia.witzlack, sabine.springer, josef.krems}@psychologie.tu-chemnitz.de

Abstract. Automated vehicles must be able to handle situations requiring cooperation with other road users in mixed traffic scenarios. Braking is an important informal vehicle signal for cooperation, indicating e.g. the intention to let a pedestrian cross. The present experimental study assessed the effects of daytime, approaching vehicle speed and participant's age on the last moment of "gentle" braking initiation from a pedestrian's perspective. Using a Labview-based simulation environment, pre-recorded real videos of approaching cars on a parking area were presented to 42 participants. Independent within-subject-variables were daytime (midday/dusk) and approaching vehicle speed ranging from 10 to 40 km/h in steps of 5 km/h. The between-subjects-factor consisted of two age groups ranging from 20 to 30 years and 50+ years. Results of the mixed ANOVA showed a main effect of daytime, vehicle speed and age as well as an interaction between age and vehicle speed. More conservative time gaps were chosen in the dusk condition. In line with previous studies, accepted time gaps decreased with increasing vehicle speed, indicating more risky crossing decisions. Older participants took more conservative decisions, especially on lower speed levels. Results show that applying one simple time gap for automated cooperative braking does not fit human/pedestrians perception and expectations. A function considering vehicle speed and daylight condition is recommended instead.

Keywords: Pedestrian-vehicle interaction · Informal communication · Gap acceptance · Automated vehicles · Crossing situation · Vulnerable road users

1 Introduction

Automated driving is supposed to bring several mobility-benefits such as improved safety, reduced emissions and congestions, higher driving comfort as well as economic competitiveness and enhanced mobility in the context of demographic changes [1]. These expected benefits are often claimed on the basis of a technology-centered perspective of vehicle automation, emphasizing technical advances. However, to exploit the potential of vehicle automation, human-machine interaction (HMI) issues are

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_101

considered a key question [2–6]. Automated vehicles will be part of the existing transport system; thus, communication/interaction[1] between automated vehicles and other road users in mixed traffic becomes an increasingly important issue. As part of the German DFG Priority Program 1835 "Cooperative Interacting Cars", the project KIVI at Chemnitz University of Technology focusses on the interaction between automated vehicles and pedestrians in parking scenarios. Due to the shared space situation, parking areas pose high demands on cooperation between vehicles and pedestrians including formal and informal communication. Formal communication (also referred as explicit or standard communication) comprises defined/regulated communication procedures and means such as turn signal, emergency lights, brake lights, the horn or warning lights [7, 8]. Informal communication (also referred as implicit communication) covers non-regulated signals such as eye contact, gestures, facial expressions, body movements and "anticipatory behavior" defined as small actions that make intentions predictable for others [7, 9]. Examples of anticipatory pedestrian behavior can be the changes in walking speed, head movements or placing a foot on the street. Car drivers can as well communicate informally by e.g. slowing down as signal to let pedestrians cross. The interpretation of informal signals is more difficult than formal signals, as informal communication is often unconscious and depends on culture, learning/socialization processes as well as the situational context [8]. However, informal communication plays a crucial role for negotiation processes in transport [7] and thus, need to be implemented in automated vehicles.

But how to "teach" automated vehicles informal communication? A practicable way is to analyze human informal communication patterns, e.g. by investigating interaction situations in non-automated traffic [10, 11]. However, for implementing these patterns into automated vehicles, exact parameters and functions for e.g. the moment of automated braking need to be derived. This can be achieved by experimental research. The focus of this experimental study was to determine the (last) moments of "gentle" car braking initiation from a pedestrian's perspective under varying speed and daytime conditions. The aim was to assess the range and function of accepted time gaps to finally give recommendations for implementing the functionality in automated vehicles. The topic of gap acceptance in the transport domain has a longer research tradition (for an overview see [12–14]). Gap acceptance in the pedestrian-vehicle interaction is usually defined and measured by assessing the decision of persons on the last moment at which they would safely cross a street before an oncoming vehicle. However, the focus of previous research is almost exclusively on safety issues in terms of accident prevention. This study transfers established gap acceptance research methods into the domain of vehicle-pedestrian communication with a focus on smooth and comfortable interaction for all partners rather than prevention of critical situations. Nonetheless, previous findings can give hints on which factors are supposed to influence gap acceptance and therefore the interaction process. In summary, main influence factors comprise vehicle speed, light/weather conditions, size of vehicles (size arrival effect/size speed bias), road characteristics (type and width), observation duration, observation angle and perspective as well as persons

[1] With reference to [8], the terms "communication" and "interaction" are used as synonyms.

characteristics such as age, gender, risk behavior, cognitive and motoric abilities [13, 14]. In this experimental study, light conditions and vehicle speed were manipulated as independent variables and age was investigated by sampling a younger and older age group. All other factors such as vehicle type, observation duration and perspective, road type and person's characteristics were controlled by keeping them constant. The latter was achieved by using exactly the same videos and presenting all conditions to all participants in a within-subjects experimental design. Regarding light conditions, [15] found more conservative gap decisions at dusk in a study on higher speed levels (40–80 km/h). Similar results were reported from [16] for rainy weather conditions in a field experiment in naturalistic traffic environment. Age related differences in gap acceptance are reported by several researchers [17–19]. In almost all studies, older participants chose longer, i.e. more conservative time gaps. This rather stable effect is attributed to either compensatory behavior due to cognitive, visual and motoric impairments at higher age, but also to a generally less risky behavior compared to young people. Vehicle speed showed as well a stable influence pattern on gap acceptance: The higher the speed, the more risky time gaps were chosen [15, 19, 20].

Based on these findings, the present study aimed at assessing the effect of daylight, vehicle speed and age on accepted comfortable time gaps, with a new focus on shared-space parking scenarios at low speed levels from 10 to 40 km/h. In addition, the effect of speed is analyzed on a higher level of detail in order to derive exact time gap values for "gentle" automated car braking initiation. In line with previous findings it is hypothesized that (1) the higher the oncoming vehicle speed, the smaller (more risky) time gaps are chosen, (2) at dusk, higher (more conservative) time gaps are chosen and (3) older people chose higher (more conservative) time gaps.

2 Method

Research Design. The experimental study was conducted using a $2 \times 2 \times 7$ mixed design. The independent between-subject factor was age of the participants in two sampled groups of 20 to 30 years and 50+ years. The two independent within-subjects variables were 1) daytime with two categories "day" (video recorded at 11:13 AM in April) and "dusk" (video recorded at 19:25 PM in April) and 2) speed of the oncoming vehicle from 10 to 40 km/h in seven steps of 5 km/h. To cover each condition for each participant, $2 \times 7 = 14$ trials were required at least. In order to stabilize the results and avoid missing data due to e.g. distraction or single missed trials, every condition was repeated three times. Thus, a total of 42 trials were presented in randomized order to each participant. Dependent variable was the last accepted time gap in seconds of the oncoming vehicle, i.e. the time difference from the moment when the participants pressed the button until the vehicle reached the position of the participant as pedestrian. After an outlier analysis (see section participants for details), mean values of the three repetitions were calculated for every condition, resulting in 14 mean time gap estimations per participant.

Video Material. The experiment included prerecorded real-world videos on a public parking area at Chemnitz University of Technology. All videos were recorded at

exactly the same place and one day, Saturday 02 of April. The recording device was a GoPro Hero 4 camera with Full HD resolution of 1920 × 1080 pixels and 120 frames per second. The camera was placed at a height of 1.60 m and 50 cm left of the roadside. This position resembles the perspective of a pedestrian intending to cross the street (Fig. 1). The oncoming vehicle was a white smart electric drive, driven by an investigator at a constant speed of 20 km/h. To ensure precise vehicle speed in the videos, unobtrusive markers were placed at the pavement. Using the markers and a video editing software, all videos were accelerated or decelerated to get an exact and constant speed of 20 km/h at the straight part of the road. For every video recording session it was made sure that no other moving objects such as other vehicles, pedestrians, bicycles, animals etc. passed the scene. Due to these precautionary measures, vehicle speed could exactly be manipulated without artificial side-effects in the simulation software by regulating the playback speed of the videos.

Fig. 1. Screenshots of video recordings at midday (11:13 AM, left) and dusk (19:25 PM, right).

Simulation Software and Apparatus. To ensure exact play rate control of the videos, precise logging of participant's reactions as well as full experimental control of instructions, randomized trials, messages etc., a dedicated simulation was programmed in LabVIEW 2015 (Fig. 2). LabVIEW's timed loop structures allow for exact control of all timings, image processing algorithms enable smooth real time presentation of video material and extensive logging functions are provided. Instructions, messages, as well as the order of experimental trials could be controlled by a separate configuration file. The simulation was presented to the participants on Laptops with 13″ Full HD Screen, the resolution of the videos was reduced to 1280 × 720 to fit into the LabVIEW simulation environment (Fig. 2).

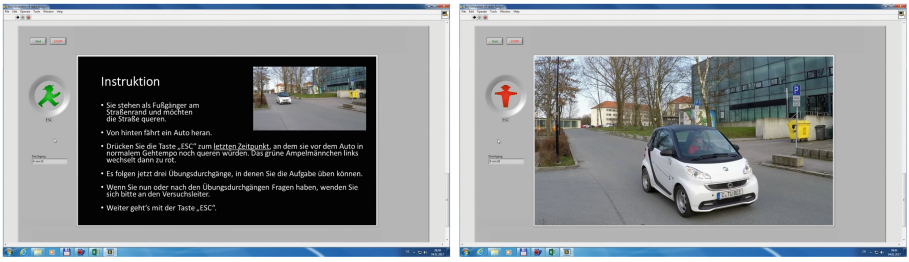


Fig. 2. Screenshots of the simulation software programmed in LabVIEW with instructions (left) and after pressing the time gap button (right).

Instructions and Procedure. Before the start of the simulation, all participants were informed about the scope and procedure of the study and anonymity was guaranteed. In addition, participants were asked beforehand if they suffer on impairments of the visual or musculoskeletal system – this would have been an exclusion criterion. All instructions regarding the simulation were presented as standardized text slides in the simulation software (Fig. 2). The instruction was to press a defined key at the last moment, when they would cross the street comfortably (without running) before the vehicle. After the instruction, three test trials were presented to let the participant practice the task and ask question in case something would be unclear. After that, the 42 experimental trials with data logging were presented in randomized order. After each trial, a dialog message with an OK-Button was shown to give participants control over the start of the next trial whenever they were ready. Participants were paid 10 Euro for participation.

Participants. A total of 43 persons participated in the study. Before calculating results, an outlier-analysis was carried out. Unrealistic trial values with suspect of missed or too fast reactions were excluded (44 of the 1,806 single values). The criteria applied was to eliminate time gaps of more than 10 s and less than 1 s. In addition, Boxplots were created for every single participant over all his/her conditions to identify individual outliers. Criterion was the exclusion of extreme values, defined as more than 3 × IQR (Interquartile range) over the third quartile or under the first quartile (8 values). As a result of the outlier analysis, two participants had to be excluded from ANOVA calculations because all three values of one condition did not satisfy the inclusion criteria and therefore no within-ANOVA could be calculated. The remaining 41 subjects consisted of 17 women and 24 men ranging from 21 to 76 years (M = 39 years, SD = 18.32). A stratified sampling was applied to get two distinct age groups under 30 years (N = 23, M = 24 years, SD = 2.16) and over 50 years (N = 18, M = 59 years, SD = 7.85). The sample was composed of 41% students, 46% persons working and 13% with other occupational status such as retired.

3 Results

Results of the mixed ANOVA showed a statistical significant main effect of speed, daytime and age as well as a significant interaction between speed and age (Table 1). Mauchly's test indicated that the assumption for sphericity had been violated for speed, therefore Greenhouse-Geisser-corrected degrees of freedom are reported. The main effect of speed can be seen in both Figs. 3 and 4. The higher the speed, the shorter minimum time gaps were chosen (= more risky decisions). The decrease of time gaps resulted as stable and continuous trend over all presented speeds.

The main effect of daytime is depicted in Fig. 3. Participants decided for significantly longer time gaps at dusk (red/dashed line) compared to midday (blue/solid line). This means that at all speed levels more conservative time gaps were chosen at dusk.

Figure 4 shows the time gaps selected by the two age groups under 30 years (blue/solid line) and over 50 years (red/dashed line). The main effects of age means that older participants chose generally greater (i.e. more conservative) time gaps than

Table 1. ANOVA table for main effects and interaction of daytime, speed and age. Statistically significant results marked in bold, N = 41.

Main effects and interaction	ANOVA		
	F	p	η_p^2
Speed	$\mathbf{F(1.59, 61.84) = 67.22}$	**<.001**	**.633**
Daytime	$\mathbf{F(1, 39) = 29.28}$	**<.001**	**.429**
Age	$\mathbf{F(1, 39) = 4.46}$	**.041**	**.103**
Speed × age	$\mathbf{F(1.62, 63.32) = 7.95}$	**.002**	**.169**
Speed × daytime	$F(4.86, 189.55) = 1.63$	.155	.040
Daytime × age	$F(1, 39) = 0.30$	.590	.008

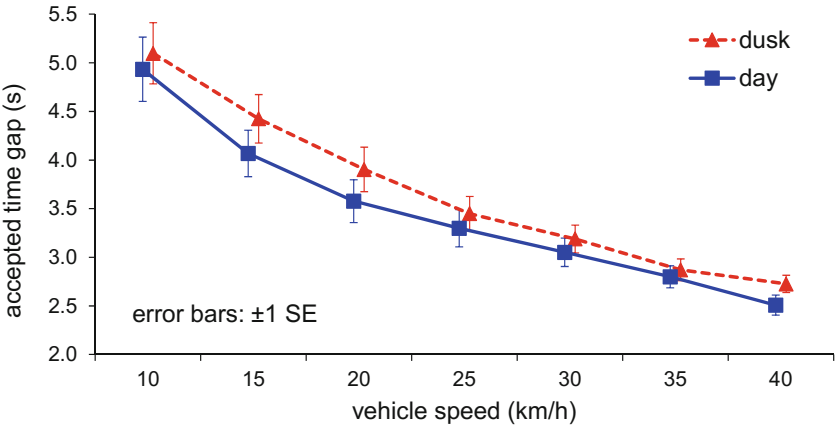


Fig. 3. Accepted time gaps for dusk/day at different speed levels.

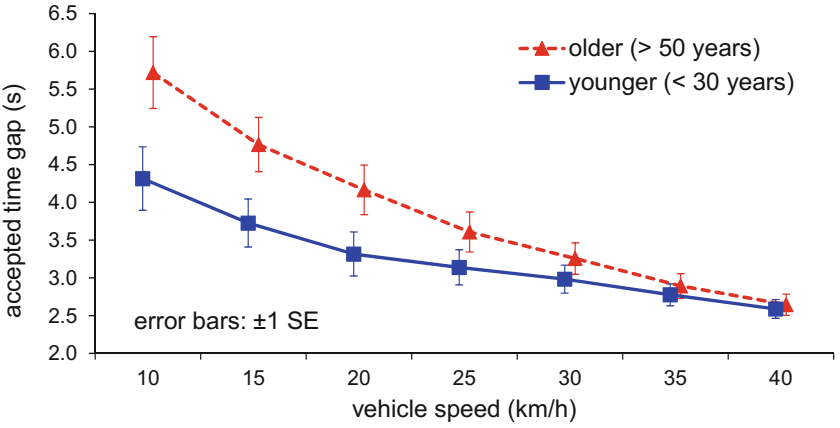


Fig. 4. Accepted time gaps for younger/older age group at different speed levels.

younger people, regardless of speed. The interaction effect between speed and age (Table 1) indicates that the trend of time gap selections at different speed levels differed between the age groups. Older participants showed a steeper decline of time gap decisions with higher speed than the younger group. The older group took more conservative decisions at lower speed levels than the younger group, however, ending up at almost the same ratings at higher speeds.

Finally, Fig. 5 shows minimum, maximum as well as 25, 50 and 75 percentile values of accepted time gap values for all participants and all speeds. It allows for estimating the number of pedestrians that would be satisfied with certain time gaps at different speed levels. The 50% line (blue/thicker) indicates that half of the pedestrians chose time gaps between 4.5 s at 10 km/h and 2.58 s at 40 km/h. Therefore, it can be inferred that half of the pedestrian would consider cooperation signals at these time gaps as (latest) moment for signaling the intention to let the pedestrian cross.

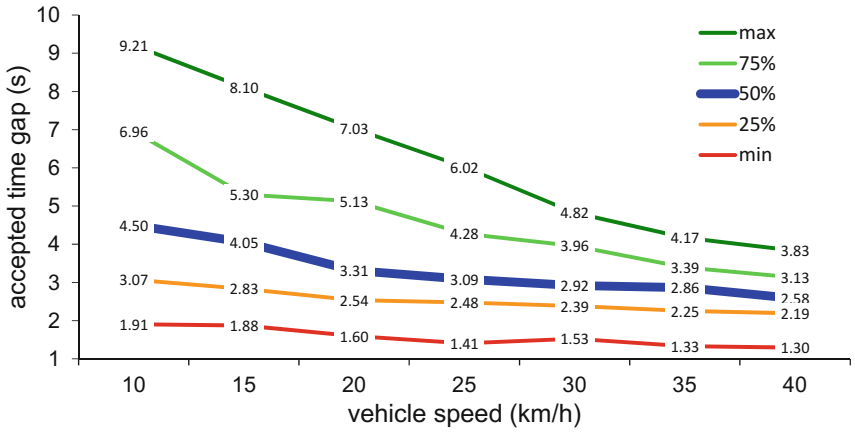


Fig. 5. Minimum, maximum as well as 25, 50 and 75 percentile values of accepted time gap values at different speed levels.

4 Discussion and Conclusions

As automated vehicles will be part of the existing transport system, communication with other road users in mixed traffic becomes an increasingly important issue. The present experimental study focused on the interaction between automated vehicles and pedestrians in parking scenarios. The aim was to determine the moment of "gentle" car braking initiation from a pedestrian's perspective under varying speed, pedestrian age and daytime conditions. These timings shall be used to implement braking initiation of automated vehicles as informal communication signal for the intention to let a pedestrian cross.

Overall, results showed that there is no simple one-fits-all solution for accepted time gaps. Hypothesis 1 on speed-effects could be proven: The higher the speed of the oncoming vehicle, the shorter time gaps were chosen, i.e. the riskier decisions were

taken. The decrease of time gaps was stable and continuous over all presented speeds. These findings are in line with previous results on safety-effects [15, 19, 20] and give a clear indication that pedestrians do not fully take the speed of the oncoming vehicle into account when deciding to cross. An explanation for this effect could be the argumentation of [18], that time gap decisions are mainly based on vehicle distance rather than time to arrival. Another explanation could be learned informal communication [7, 8]: pedestrians expect vehicles at a certain distance to break in any case if a pedestrian enters the street – almost independent of the vehicle speed at that particular moment. These results lead to the practical recommendation, that braking initiation for automated vehicles cannot be based on a single time-to-contact-value for all speeds. In order to be "naturally" perceived as informal signal by pedestrians, brake initiation needs to be adjusted according to the speed-dependent time gap curves presented in Figs. 3, 4 and 5. Taking into account technical sensor limitations of e.g. camera-based pedestrian detection systems, not all "ideal" time gaps can probably be realized due to long distances and few time for image processing. The percentile values presented in Fig. 5 allow for an estimation of how many pedestrians would be satisfied with a certain moment of braking initiation.

However, speed is not the only factor influencing time gap decisions. Hypothesis 2 on the effect of light conditions could be proven too, indicating consistently higher (i.e. more conservative) time gap decisions at dusk. This effect is also in line with similar findings at higher speed levels and in naturalistic settings [15, 16]. Thus, automated vehicles should consider light conditions as well for comfortable and natural-looking brake signal to pedestrians.

Vehicle speed and light conditions are parameters that can rather easily be captured by sensors of automated vehicles. However, for most comfortable braking moment, also the age of the pedestrian should be taken into account. The statistically significant effect of age according to Hypothesis 3 indicates that older pedestrians prefer longer time gaps than younger adults. This tendency towards more conservative and less risky decisions corresponds with previous findings [17–19] and can be related to either compensatory behavior due to cognitive, visual and motoric impairments at higher age, but also to a generally less risky behavior compared to young people. However, the relatively high number of speed levels in this study allows for deeper insights into this relation. The age effect is especially pronounced at lower speed levels (interaction speed $\times$ age), which could be related to the importance of distance for time gap decisions [18]. The closer a vehicle in terms of distance, the higher the effect of conservative/less risky decision taking on time gap acceptance. Even though assessing the age of a pedestrian will be a rather complex task for automated vehicle sensors (such as video cameras), these results give indication that age recognition would especially be important at lower speed levels. According to Figs. 3, 4 and 5, lower speed levels go along with shorter pedestrian-vehicle distance to cover for video cameras and more time for image processing. A second argument for the importance of assessing pedestrian characteristics is the fact that manual driver perform this evaluation too and adapt their behavior. Considering the process of "schema formation" [7], other road user's behavior is anticipated based on specific characteristics such as being an elderly person or a child. Therefore it can be assumed that manual driver also adapt braking initiation as sign for cooperation in accordance with perceived characteristics of the pedestrian.

Limitations of the study are mainly related to the presentation of video material instead of having real world interactions. The different field of view, ambient sound as well as embodied cognition could change time gap decision on real roads. However, (1) field tests do not allow for such exact manipulation of dependent variables and (2) previous studies found evidence that time-arrival estimates do not differ between video and real world situations [21]. Even though, on-road tests are required to validate these results. Within the KIVI project, an automated vehicle is currently built up, instrumented with a complementary two-camera system for precise pedestrian detection. Using this vehicle, on-road tests will be carried out by implementing the time gap functions found in this study. Final aim of KIVI is to demonstrate "gentle and naturally perceived" automated braking as informal communication signal to pedestrians in crossing situations.

Acknowledgements. This project was funded within the Priority Program 1835 "Cooperative Interacting Cars" of the German Science Foundation DFG.

References

1. Meyer, G., Deix, S.: Research and innovation for automated driving in Germany and Europe. In: Meyer, G., Beiker, S. (eds.) Road Vehicle Automation. Lecture Notes in Mobility, pp. 71–84. Springer, Cham (2014)
2. Gasser, T.M., Arzt, C., Ayoubi, M., Bartels, A., Eier, J., Flemisch, F., Häcker, D., Hesse, T., Huber, W., Lotz, C., Maurer, M., Ruth-Schumacher, S., Schwarz, J., Vogt, W.: Rechtsfolgen zunehmender Fahrzeugautomatisierung. Wirtschaftsverl. NW Verl. für neue Wissenschaft, Bremerhaven (2012)
3. Banks, V.A., Stanton, N.A.: Keep the driver in control: Automating automobiles of the future. Appl. Ergon. **53**, 389–395 (2015)
4. Riener, A., Boll S., Kun A.L.: Automotive user interfaces in the age of automation (Dagstuhl Seminar 16262). Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik. Dagstuhl Reports, Volume 6, Issue 6 (2016)
5. Beggiato, M.: Changes in motivational and higher level cognitive processes when interacting with in-vehicle automation. Dissertation, TU Chemnitz. http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa-167333 (2014)
6. Beggiato, M., Hartwich, F., Schleinitz, K., Krems, J.F., Othersen, I., Petermann-Stock, I.: What would drivers like to know during automated driving? Information needs at different levels of automation. 7. Tagung Fahrerassistenz, Munich, 25–26 November 2015 (2015). doi:10.13140/RG.2.1.2462.6007
7. Färber, B.: Communication and communication problems between autonomous vehicles and human drivers. In: Maurer, M., Gerdes, J.C., Lenz, B., Winner, H. (eds.) Autonomous Driving, pp. 125–144. Springer, Berlin (2016)
8. Hölzel, A.: Unterscheidung von formeller und informeller Kommunikation im Straßenverkehr. Diplomarbeit, Wien (2008)
9. Lagström, T., Lundgren, V.M.: AVIP - Autonomous vehicles' interaction with pedestrians. Master of Science Thesis, Chalmers University of Technology (2015)

10. Witzlack, C., Beggiato, M., Krems, J.: Interaktionssequenzen zwischen Fahrzeugen und Fußgängern im Parkplatzszenario als Grundlage für kooperativ interagierende Automatisierung. In: VDI (ed.) Fahrerassistenz und automatisiertes. Fahren VDI-Berichte, vol. 2288, pp. 323–336. VDI-Verlag, Düsseldorf (2016)
11. Gohl, I., Schneemann, F.: Analyse der Fahrer-Fußgänger Interaktion an Fußgängerüberwegen. In: VDI (ed.) Fahrer, Fahrerunterstützung und Bedienbarkei. 8. VDI-Tagung Der Fahrer im 21. Jahrhundert: Braunschweig, 10. und 11. November 2015. VDI-Berichtet, vol. 2264, pp. 41–52. VDI Verlag GmbH, Düsseldorf (2015)
12. Petzoldt, T.: On the relationship between pedestrian gap acceptance and time to arrival estimates. Accid. Anal. Prev. **72**, 127–133 (2014). doi:10.1016/j.aap.2014.06.019
13. Schleinitz, K., Petzoldt, T., Krems, J., Kühn, M., Gehlert, T.: Geschwindigkeitswahrnehmung von einspurigen Fahrzeugen. Forschungsbericht, vol. 33. Gesamtverband der Deutschen Versicherungswirtschaft, Berlin (2015)
14. Hagen, K., Schulze, C., Schlag, B.: Verkehrssicherheit von schwächeren Verkehrsteilnehmern im Zusammenhang mit dem geringen Geräuschniveau von Fahrzeugen mit alternativen Antrieben. FAT-Schriftenreihe, vol. 245. VDA; Technische Informationsbibliothek u. Universitätsbibliothek, Berlin (2012)
15. Liu, Y.-C., Tung, Y.-C.: Risk analysis of pedestrians' road-crossing decisions: Effects of age, time gap, time of day, and vehicle speed. Saf. Sci. **63**, 77–82 (2014). doi:10.1016/j.ssci.2013.11.002
16. Sun, R., Zhuang, X., Wu, C., Zhao, G., Zhang, K.: The estimation of vehicle speed and stopping distance by pedestrians crossing streets in a naturalistic traffic environment. Transp. Res. Part F Traffic Psychol. Behav. **30**, 97–106 (2015). doi:10.1016/j.trf.2015.02.002
17. Lobjois, R., Benguigui, N., Cavallo, V.: The effects of age and traffic density on street-crossing behavior. Accid. Anal. Prev. **53**, 166–175 (2013). doi:10.1016/j.aap.2012.12.028
18. Oxley, J.A., Ihsen, E., Fildes, B.N., Charlton, J.L., Day, R.H.: Crossing roads safely: an experimental study of age differences in gap selection by pedestrians. Accid. Anal. Prev. **37** (5), 962–971 (2005). doi:10.1016/j.aap.2005.04.017
19. Alexander, J., Barham, P., Black, I.: Factors influencing the probability of an incident at a junction: results from an interactive driving simulator. Accid. Anal. Prev. **34**(6), 779–792 (2002). doi:10.1016/S0001-4575(01)00078-1
20. Yan, X., Radwan, E., Guo, D.: Effects of major-road vehicle speed and driver age and gender on left-turn gap acceptance. Accid. Anal. Prev. **39**(4), 843–852 (2007). doi:10.1016/j.aap.2006.12.006
21. Recarte, M.Á., Conchillo, Á., Nunes, L.M.: Estimation of arrival time in vehicle and video. Psicothema **17**(1), 112–117 (2005)

Older Drivers and Driving Automation

Anabela Simoes(✉)

Universidade Lusofona, 376, Campo Grande, 1749-024 Lisbon, Portugal
anabelasimoes@ulusofona.pt

Abstract. Older drivers, representing the most growing segment of the driving population, are generally considered safe and cautious drivers and it is known that they self-regulate their driving behavior to minimize the risk of crashing. However, their age-related perceptual and cognitive declines lead them to change their driving habits as a way of compensating their limitations and difficulties in driving. Being older drivers particularly vulnerable, it is considered that in-vehicle technologies represent a potential to increase safety and possibly save lives. Automation in the road transport system is coming faster than expected but most of the potential users have unrealistic representations of autonomous vehicles, the driver's role in automation or the impacts of automation on the road transport system. Thus, the adoption of partial or full automation to help older drivers in the performance of the driving task is still requiring in-depth research, as well as the design and application of appropriate training programs.

Keywords: Older drivers · Self-regulation · In-vehicle technologies · Human-centered design · Driving automation · Function allocation · Human-automation interaction · Older drivers' training

1 Introduction

The road transport system is a very dynamic and complex system involving human-technology interaction and human-technology cooperation (in-vehicle assistive technologies), together with numerous social interactions in the traffic. Furthermore, the road transport system is totally open, in opposition to other modes of transport (railway, maritime, aviation) in which highly trained and competent professionals drive or pilot different vehicles. Within the road transport system, there are an enormous variety of road users (drivers, riders and pedestrians; novice, experienced, young, old and mobility-impaired, together with professional drivers, etc.) evolving on the road environment. Such variety, just covered by traffic laws and the corresponding supervision, as well as the diversity of behaviors in the road environment, lead to an increased dynamics, complexity and uncertainty on the road transport system. Being safety a major concern in transport, the new approach based on a human-centered design approach integrating users, technology and organization [1] seems to be the way to achieve automation aiming at harmonizing behaviors and improving safety in such complex system.

Older drivers have the highest growing representation in the population in most industrialized countries. The proportion of older drivers is increasing in society and is

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_102

expected to enlarge even more [1–3]. In most of OECD countries, it is expected that by the year 2030 one person out of four will be 65 years old and over. This is explained by the maturation of the "baby boom" generation, combined with a greater longevity and the decline of birth rates. Furthermore, the number of older drivers is also increasing since the twentieth century 60s, when much more young men and women got their driving license. Driving has nowadays an important role in society and is essential to determine the quality of life of older people. Driving enables easy access to activities and services, to fulfill social needs and because of that it is a good indicator of mobility, independence, good health, quality of life and well-being. The car represents to the elderly the possibility of maintaining their autonomy and self-esteem being also a symbol and a way of freedom, independence and self-reliance [4, 5]. A car is also the safest and most comfortable transport mode for older people. Therefore, two actions are being undertaken to improve or maintain older drivers' safe mobility: (1) specific training programs to contain age-related declines and develop self-regulation ability; (2) automated driving.

Automated driving will change our lives and travel patterns but, so far, few is known about safety and security issues, as well as the impacts of human-automation interaction. Automation represents a major technological advancement influencing and shaping the future of mobility. However, automation won't replace human activity but instead it will impose new coordination demands to the human driver or user. So far and for long years, it is expected that autonomous vehicles will share the road environment with conventional ones, which are driven by human drivers. This requires continuous research on human factors issues, as well as specific actions towards public awareness on new risks introduced by autonomous vehicles on the shared road, together with training towards the appropriate use of automation, the prevention of risky behaviors and avoidance of misuse and disuse.

2 The Road Transport System and the Older Driver

2.1 The Road Transport System as a Safety Critical System

Modern organizations and the related technological development require fast, efficient and precise actions from human beings. Aiming at ensuring the success of those actions, as they ultimately rely on an adequate cooperation between humans and technologies, the different factors leading to the variability of human performance must be identified, understood and managed. These factors, resulting from the diversity of human characteristics and functioning, as well as their short-term and life span variability, contradict the assumption of the stability of human activity over time that often presides to the design and management of numerous systems.

The more developed and technology-based working systems are, the more safety critical they become, thus imposing high performance levels towards safety and efficiency. Modern companies are increasingly shaped by technology (automation, computer-supported information and cooperation networks, multimedia applications in production planning and control, etc.), integrating social and organizational elements (Technology, Organization and People), which are fundamental to understand the system

behavior and should be viewed as embedded in the system [6]. As much as a system is complex and safety-critical the more human operators' skills and abilities are important for the system efficiency and safety. Human operators represent the most flexible, adaptable and valuable elements of any system. They learn with experience as responsible, collaborative, tool creating/wielding agents towards success. Even under resource and performance pressure at all levels of the system, they learn and adapt to different situations and multiple task goals. However, they are also the most vulnerable elements of a system, due to human variability and instability, as well as human behavior.

The road transport system is very dynamic and complex involving human-technology interaction and cooperation, together with numerous social interactions in the traffic. In opposition to other transport modes (e.g. railway, maritime, aviation), in which highly trained and competent professionals drive or pilot different vehicles, the road transport system is totally open, accommodating a huge variety of users (drivers, riders and pedestrians; novice, experienced, young, old and mobility-impaired, together with professional drivers, etc.) evolving on the road environment. Such variety of users, just covered by traffic laws and the corresponding supervision, as well as the diversity of behaviors in the road environment, lead to an increased dynamics, complexity and uncertainty on the road transport system. Considering the whole system as a joint cognitive system, the human-technology cooperation should be viewed as the state of being in action within a system in a perfect coupling between the human and technology, which is the basis for the system efficiency and safety.

The introduction of new technologies on the road transport system with purposes of improving safety, mobility and system efficiency, as well as protecting the environment, require appropriate design and use to avoid bad outcomes from very good intentions. This requires a good acceptance of the innovation, which depends on: (1) the system design, accommodating the needs and expectations of all potential users; (2) the respect for their rights (privacy and security); and (3) the provision of adequate training [7]. These are key issues that will ensure the person's trust on the technology and its appropriate and safe use. The same principles are required for any technological innovation in the road transport system, such as, automation, which is coming faster than expected but most road users have an unrealistic representation of autonomous vehicles. Very few is known about the driver's role in automation or the impacts of full automation on the road transport system. There is still much to investigate and learn about human-automation interaction during the long run towards a fully connected and automated road transport system.

2.2 Older Drivers

Older drivers are generally considered safe and cautious drivers and it is known that they self-regulate their driving behavior to minimize the risk of crashing [4]. The age-related perceptual and cognitive declines in sensory, motor and cognitive domains are often subtle and gradual. It is difficult for older drivers to draw the relationship between these changes and the impact on their driving ability but they start changing their driving habits in a natural way and using their experience to overcome some perceived difficulties. It is common to hear from older drivers that they avoid driving

conditions imposing higher demands, such as driving at night, with poor weather conditions or during rush hours. However, they continue driving safely compensating their failures by using their remaining abilities in a deeper way. In a healthy ageing process, age-related decrements don't have significant impact on task performance if previous experience can be used.

Despite an increasing individual diversity with ageing, it seems that, for the same task performance, older people develop the same self-regulatory behaviors for functional declines. This results in some common behavioral patterns to compensate the age-related impaired perceptual-motor functioning. Such self-regulation behaviors should include the insight to one's own limitations and related compensatory behaviors [8], which reduce the stress and anxiety felt by older drivers in some driving situations, as well as the risk related to driving in such situations [3, 9]. Avoidance behavior is a self-regulatory behavior that depends on high-level strategic choices requiring efficient metacognitive abilities. Thus, in a prevention perspective, it is important to determine which instruments best reflect metacognitive abilities and are most predictive of avoidance. As pre-conditions for an effective self-regulatory behavior, the use of experience, as well as a more stable and user-friendly road environment are required. However, the nowadays roadway system is increasingly dynamic and complex. Moreover, the diversity of in-vehicle technologies could increase the task complexity, reducing the possibility of using previous experience, which they don't have as such technologies are recent. Therefore, two actions are being undertaken to improve or maintain older drivers' safe mobility: (1) specific training programs to contain age-related declines and improve self-regulation ability; (2) technology-based support to the driver, including automated driving.

Safe driving requires an accurate adjustment of the perceived task demands to the perceived individual capabilities, which defines self-regulation. The accuracy of self-regulation strategies (e.g., reducing driving speed, increasing headway or avoiding night driving) depends on the drivers' accurate perception of their ability to assess their own driving and cognitive performances (accurate self-awareness) and to correctly assess the task demands (accurate risk awareness). Both accurate self-awareness and risk-awareness are important pre-conditions to correctly adapt the driving behavior to age-related declines. Drivers who can accurately assess their own capabilities will be able to self-regulate and prolong their safe mobility. However, drivers who under- or overestimate their capabilities or under- or overestimate task demands won't be able to self-regulate adapting their driving behavior to their own age-related declines. This knowledge about the quality of older drivers' self-assessment, hazard perception and self-regulation, is essential for any intervention towards the safety improvement of older drivers, together with the identification of drivers with potential to benefit from such intervention and their specific needs. Besides some interventions in the infrastructure and equipment aiming at easing the mobility-related tasks for all users, direct training interventions with older drivers will also offer good opportunities for informing them about (1) their age-related physical and cognitive declines, (2) their driving difficulties arising from these declines, and (3) how to modify driving strategies to avoid these difficulties [10]. Furthermore, when driving a car is no longer considered a mobility option, the intervention program provides counselling for the transition to other transport modes. With the same purpose, education campaigns should inform

other road users about the difficulties older drivers experience in traffic, which may help them to anticipate older drivers' behavior, and help to prevent future crashes.

Older Drivers and In-vehicle Technologies. Being older drivers particularly vulnerable, it is considered that in-vehicle technologies represent a potential to increase safety and possibly save lives, highlighting the six major objectives to be achieved by such technologies [11]: (1) to improve safety; (2) to improve mobility; (3) to improve system efficiency; (4) to increase productivity; (5) to conserve energy and protect the environment; (6) to satisfy customers.

Technological advances regarding in-vehicle technologies are designed and deployed with the aim of assisting drivers in the driving task performance, but they require some behavioral adaptation and may overload them by imposing additional tasks to driving. These in-vehicle technologies are split into two categories [12]: In-Vehicle Information Systems (IVIS) and Advanced Driver Assistance Systems (ADAS). The first ones provide drivers with information and communication during the vehicle motion, being some of them useful for driving, such as, a navigation system or a collision avoidance system. However, providing the driver with information not related to driving (e.g. phone conversation or consulting emails) represents a factor of distraction, overloading the driver and compromising his/her ability to react promptly to an unexpected and critical situation. ADAS (Advanced Driver Assistance Systems) provide drivers with different types of assistance supporting the driving task in the vehicle maneuvering by informing, warning and/or actively performing an aspect of the driving task. Such assistance is generated by the diagnosis of the driving situation, which is based on the information provided by sensors in the car and configurations introduced by the driver (e.g. adaptive cruise control, lateral and longitudinal control, etc.). ADAS represent a step towards automation giving rise to human-machine cooperation as the technology detects events, processes information and acts according to the diagnosis and the configurations previously introduced by the driver, supporting him/her on the task performance. IVIS just provide information and communication, even when alerting the driver (e.g. collision alert, blind spot information system).

As innovative technologies, they require adequate design to avoid bad outcomes of well-intentioned but poorly designed and tested technology. Even easy-to-use and somehow intuitive, their use might require some training, helping the driver to make a safe use of the system or understand how to cope with the limits of the technology. Synthesizing several studies on older drivers and in-vehicle technologies that have been carried out, Table 1 presents the main benefits from the use of the available technologies by older drivers [11, 13–16].

Another perspective in driver assistance technology is the design and development of specifically addressed systems based on research results and existing knowledge on older drivers' activity and age-related difficulties. Another study [17] aimed at specifying and designing a future Senior-Adapted Driving Assistance System (S-ADAS) based on an in-depth analysis of drivers' activity. The final objective was to design an advanced co-piloting system based on monitoring functions that have already been proved to "enable a joint analysis of the driver's activity according to the data enabling to characterize the driving situation" [18].

Table 1. In-vehicle technologies with potential to assist older drivers with selected impairments

Age-related limitations	Driving-related problem	Technology
Increased reaction time: inability to divide attention between tasks	Difficulty driving in unfamiliar or congested areas	Navigation/route guidance; traffic information; variable message signs
Deteriorating vision, particularly at night	Difficulty seeing pedestrians and other objects at night and difficulty reading signs	Night vision enhancement systems; in-vehicle signing
Difficulty judging speed and distance	Failure to perceive oncoming vehicles; crashes at intersections	Collision warning, automated lane changing and merging
Difficulty perceiving and analyzing situations	Failure to comply with yield signs and traffic signals; slow to appreciate hazards ahead	In-vehicle signing and warnings; intelligent cruise control
Difficulty turning head, reduced peripheral vision	Failure to notice obstacles while maneuvering; worries over merging and lane changing	Blind spot/obstacle detection; automated lane changing and merging
More prone to fatigue	Getting tired on long journeys	Intelligent cruise control; automated lane following
General effects of aging	Worries over inability to cope with a breakdown; worries about driving to unfamiliar places, driving at night, and driving in heavy traffic	Emergency callout (Mayday) systems; vehicle condition monitoring; navigation and route guidance systems; traffic information via on-board display of signs and warnings of hazards (e.g., HUD), and use of variable message signs
Impairments varying in severity from day to day	Concern over fitness to drive	Driver condition monitoring

This represents a human-centered design approach addressing older drivers with the aim of designing, developing, testing and implementing S-ADAS based on such monitoring functions specifically adapted to older drivers' needs. The increasing number of older drivers justifies the availability of such useful functions in ADAS to be activated when desired.

3 Autonomous Vehicles and Driving Automation

3.1 The Long Run Towards Driving Automation

Automation in the road transport system is coming faster than expected but most of the potential users have unrealistic representations of autonomous vehicles, the driver's role in automation or the impacts of full automation on the road transport system. There is still much to research and a long run towards a fully connected and automated road system. From a technological point of view, autonomous vehicles are practically ready to be used. Cooperative systems have been developed and deployed so that the road infrastructure is already connected in most developed countries. Required roadway

equipment and traffic control systems are ready to operate under a fully connected mode. However, is the human operator ready to manage and operate in such dynamic and complex system? Comparing to the technological development in aviation, where the automation of several components aims at assisting the pilot or assuming a certain level of control over the aircraft, automation introduces interesting perspectives in terms of human-automation interaction. Briefly, one of the lessons learned from aviation is that the systematic use of automation can lead to degraded performance, resulting from a lack of reinforcement of pilot skills. This emphasizes the needs for high demands regarding pilots' training and performance, which means that automation does not represent a way of replacing the pilot but instead it imposes high skills, functional abilities, knowledge and competency to manage human-automation interactions.

As the road transport system involves human-machine interaction and cooperation, as well as multiple social interactions on the road environment, automation requires a human-centered design approach with the aim of avoiding any poor outcomes from well-intentioned but poorly designed and tested technology. From a human-centered design perspective, the role of the human in driving automation and his/her interaction with such system should lead research and development in this field as "automation is not a matter of all or nothing" [20]. The extent of automation should be decided based on the identification of the cognitive functions to be transferred from the human to the technology.

Different from aviation and other transport modes, it should be reminded that the road transport system is totally open, involving a diversity of users who are walking, driving or riding, sharing the road space and being just regulated by traffic laws, frequently under a poor supervision. This, together with the permanent social interactions in the shared space, increases the system dynamics, complexity and uncertainty. The evolution towards a fully connected and automated road system is being fast in what concerns the technological availability but it will be much slower in terms of adaptation of all users to a fully regulated and automatic system. Similarly, in the context of aviation, a well-trained, competent, skilled and in good fitness driver is required behind the wheel as, at any moment, he/she will be called to intervene. Furthermore, a fully connected and automated road transport system doesn't seem compatible with the complexity of the nowadays shared space on the road environment.

3.2 Levels of Automation

In automation design, it should be highlighted the decision about the extent of automation. Since the last decades of the 20th century several studies proposed different levels of automation resulting from theoretical classifications of the different functions to be allocated to the system. A new taxonomy has been proposed in the context of automation in air traffic management (ATM) [19]: the LOAT (Level of Automation Taxonomy). The taxonomy is organized according to the generic functions

defined in [21, 22] to be allocated to the system: information acquisition, information analysis, decision and action selection, and action implementation. The taxonomy defines, for each function, different levels of automation (from 5 to 8), starting with a default level zero corresponding to the manual task and a level 1 corresponding to an external support to the task performance, which is far from automation. The next levels are increasing the extent of automation until full automation, as following: (1) Both information acquisition and information analysis functions are fully automated at level 5; (2) The allocation of the decision and action selection function requires 6 levels until the full automation; (3) The allocation of the action implementation function is fully automated at level 8.

Although being developed with ATM automation in mind, this taxonomy has been considered applicable to other contexts under automation processes as it represents a human-centered approach to automation based on the definition of generic human functions to be allocated to the system. In the context of road transport, with the aim of providing a common terminology for automated driving, the Society of Automotive Engineers (SAE) developed a new standard (J3016) providing a Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles [23]. This taxonomy provides definitions for terms related to automated driving, clarifying the role of the human during the engagement of the driving automation system. The provided definitions simplify communication and facilitate collaboration within technical and policy domains as it addresses the practice related to users and stakeholders. The taxonomy identifies the following six levels of automation from no automation to full automation: 0. No automation; 1. Driver assistance; 2. Partial automation; 3. Conditional automation; 4. High automation; and 5. Full automation. Each level is defined based on functional aspects of the technology. It is also provided a description of the categorical distinctions for a step-wise progression through the automation levels. The classification system is consistent with the current industrial development and is considered useful in educating a wider community of drivers and stakeholders making clear the driver's role in the performance of the driving task at any level of automation (Table 2).

In the SAE taxonomy, the automation levels are differentiated according to the following functional characteristics: (1) the subtasks of the dynamic driving task (the lateral and longitudinal control and the object or event detection and response); and (2) the functional capabilities (driving mode and fallback capability). Based on these functional characteristics, higher levels of automation exceed and include lower automation capabilities, as following [24]:

- At level 0, the driver performs the entire dynamic driving task providing the appropriate responses to all driving conditions;
- At level 1, the driving mode is assisted by a driver assistance system for either lateral or longitudinal control using information collected from the driving environment, but the driver is expected to perform the remainder of the DDT;

- At level 2, the system performs part of the OEDR subtask in conjunction with the performance of both lateral and longitudinal control subtasks, being the driver expected to perform the remainder of the DDT;
- At level 3, higher driving automation systems automate the entire DDT under limited conditions outside of which the driver performs the entire DDT; the system provides appropriate responses to relevant objects and events with limits in functional capabilities and fallback performance capability;
- At level 4, the driving automation system performs the entire DDT, providing appropriate responses to relevant objects and events, but still being situationally-limited in operational capabilities in terms of driving modes, circumstances, and/or locations; if the system is nearing the end of its operating capabilities, the automation system will warn the driver of the need to resume performance of the DDT far enough in advance to permit an orderly transfer;
- At level 5, the driving automation system performs the entire DDT, providing appropriate responses to relevant objects and events without situational limits in functional capabilities.

Table 2. Levels of automation and driver's role (SAE taxonomy)

Automation level	Dynamic driving sub-tasks		Functional capability	
	Sustained execution of lateral and/or longitudinal control	Object and event detection and response (OEDR)	Fallback performance of dynamic driving task (DDT)	Driving mode circumstance, location capabilities
The human driver monitors the driving environment				
0 – No automation	Driver	Driver	Driver	None of the DDT is automated
1 – Driver Assistance	Driver and system	Driver	Driver	Some driving modes
2 – Partial Automation	System	Driver	Driver	Some driving modes
The automated driving system monitors the driving environment and performs the entire DDT while engaged				
3 – Conditional automation	System	System	Driver	Some driving modes
4 – High automation	System	System	System	Some driving modes
5 – Full automation	System	System	System	All driving modes

3.3 The Older Driver and Driving Automation

Very few is known about the impact of automatic driving on traffic and how drivers from different groups will accept, use, trust and interact in traffic when driving in automatic mode. Considering the older drivers' characteristics and their low familiarity with the most advanced technology, it will take some years to introduce autonomous vehicles in the market and make them attractive to older drivers. From the early nineties of the past century several research projects in the frame of ITS (Intelligent Transport Systems) research programs addressed older drivers with the aim of identifying the most useful driving assistance technologies for this group of drivers. A characteristic of older ages is the difficulty in self-learning, particularly in what concerns advanced technology. Therefore, even proved that a driving assistance technology has a great potential for helping older drivers in their age-related difficulties in driving, they need to be taught and get some tutorial practice in such use. Otherwise, they will give up at any difficulty related to the use of the technology. That's why, after all these years of ITS development, there are not so many older drivers using the very helpful assisting technologies. This means that older drivers need specific training addressing the use of advanced in-vehicle technologies. The training activity will make them understand the system functioning, adapting their driving behavior to the system and developing trust on the technology and the available automation.

A recent study [25] was based on a survey aiming at identifying the drivers' preferences among three levels of automation, including preferences for interacting with and overall concern about riding in self-driving vehicles: no self-driving, partially self-driving and completely self-driving. A total of 618 licensed drivers in US (23.6% being 60 years old and over) replied to the questionnaire. The most frequent preference for vehicle automation is for *no self-driving* (45.8%), followed by *partially self-driving* (38.7%), and *completely self-driving* as the last preference with 15.5%. It was also noted that the preferences for self-driving generally decreased with increasing age of respondents with the older group (60 and over) responding as following: 56.2% for *no self-driving*; 34.2% for *partially self-driving*; and 9.6% for *completed self-driving*. This is expected once common citizens are not so much aware of innovation in the car industry and they can't imagine how close we are from such reality. Anyway, it will take decades until having a fully connected and automated road environment. It will take time as it will be necessary that users will appropriate the current innovations in this field, develop trust on their use and the required willing to use and pay for it.

It can be said that automation has a potential to be very useful for older drivers to support their mobility needs in safety, comfort and efficient conditions. However, the systematic use of a technology requires learning and practice allowing for the construction of trust, which plays a leading role in determining the willingness of potential users to rely on automated systems in situations characterized by uncertainty. It is a main request for a safe and effective use of automation but it requires learning and practice, which should be enforced and even provided by regulatory authorities, particularly when safety and security issues are on stake. In the context of driving

automation, accidents related to automation misuse by over-trusting or automation disuse by under-trusting, can occur. Thus, developing trust in automation use is the key to safety improvement and should be promoted by means of specific training for a safe driving. This is particularly important for older drivers to ensure the development of their adaptation to the technology-related changes in the road transport and their safe use based on appropriate trust. Furthermore, the maturity of technology will also play an important role in developing trust and its safe and effective use.

4 Final Remarks

A new reality requires immediate and appropriate actions towards its success. Due to both safety and security issues featuring the road transport system, such new reality involving autonomous vehicles and automated driving requires the acceptance of the automated technology by users together with the formation of the required trust on such technology. For all road users, including those who will continue driving a conventional vehicle for long years, both public awareness and understanding of the actual changes are also being necessary and should be mandatory. The new reality will generate behavioral adaptations, which should be facilitated by the road environment design, new traffic laws and the provision of appropriate and in due time information. However, there is still a long way towards a fully connected and automated road transport system. If the technological development is fast, the use of the related innovation by common people takes time. In the present case and due to the safety and security constraints, the knowledge gap between the industrial development and the society is huge and requires in-depth and urgent research on human-factors with the aim of providing answers to questions, such as: (1) drivers' preferences for the automation levels across gender, age, education level, user group, etc.; (2) limits of the technology; (3) changes in traffic laws, as well as licensing and training; (4) driver's promptness to take the vehicle control following a long period of automated driving; etc. Many questions without answers are arising and answers are urging.

References

1. Boy, G.: Cognitive Function Analysis. Ablex Publishing Corporation, Stamford (1998)
2. Holland, C.A.: Older drivers: a literature review. Department for Transportation. Technical Report (no. 25). London, UK (2001)
3. OECD: Ageing and transport. Mobility needs and safety issues (2001). www.oecd.org/dataoecd/40/63/2675189.pdf
4. Charlton, J., Oxley, J., Scully, J., Koppel, S., Congiu, M., Muir, C., Fildes, B.: Self-Regulatory Driving Practices of Older Drivers in the Australian Capital Territory and New South Wales. Accident Research Centre, Monash University, Clayton (2006)
5. Oxley, J., Charlton, J., Fildes, B., Koppel, S., Scully, J., Congiu, M., Moore, K.: Crash Risk of Older Female Drivers. Accident Research Centre, Monash University, Clayton (2005)

6. Tschiersch, Schael: Concepts of human-centred systems. In Brandt, D. (ed.) Human-Centered System Design - First: People; Second: Organization; Third: Technology (2003). http://www.ima-zlw-ifu.rwth-aachen.de/fileadmin/user_upload/INSTITUTSCLUSTER/Alumni/human-centred_system_design.pdf
7. Meyer, J.: Personal vehicle transportation. In: Pew, R.W., Van Hemel, S.B. (eds.) Technology for Adaptive Aging. The National Academies Press, Washington (2003)
8. ERSO: Older drivers. European road safety observatory (2006). www.erso.eu/knowledge/content/06_drivers/20_old/olderdrivers.htm
9. Kostyniuk, L.P., Trombley, D.A., Shope, J.T.: The process of reduction and cessation of driving among older drivers: a review of the literature. Michigan University, Transportation Research Institute. Report No. UMTRI-98-23. UMTRI-91238 (1998)
10. Davidse, R.J., Hagenzieker, M.P.: Assisting the older driver: Strategies to improve and prolong the safe mobility of older people. Tijdschrift voor Ergonomie **33**(4), 10–17 (2008)
11. Okola, A.R., Walton, C.M.: Intelligent Transportation Systems to Improve Elderly Persons' Mobility and Decision Making within Departure Time Choice Framework. Centre for Transportation Research, University of Texas at Austin, Austin (2003)
12. Brusque, C., Bruyas, M-P., Carvalhais, J., Cozzolino, M., Gelau, C., Kaufman, C., Macku, I., Pereira, M., Rehnová, V., Risser, R., Schmeider, K., Simoes, A., Simonova, Z., Turetscheck, C., Vaa, Vasek, J., Zehnalová, V.: The Influence of In-Vehicle Information Systems on Driver Behaviour and Road Safety. Synthesis of Existing Knowledge, COST 352. INRETS, Arcueil, France (2007)
13. Molnar, L., Eby, D.W., Miller, L.: Promising Approaches for Enhancing Elderly Mobility. UMTRI, Ann Arbor (2003)
14. Staplin, L., Lococo, K.H., Martell, C., Stutts, J.: Taxonomy of older driver behaviors and crash risk. NHTSA. DOT HS 811 468C Report (2012)
15. Davidse, R.: Assisting the older driver: intersection design and in-car devices to improve the safety of the older driver. SWOV-Dissertatiereeks, Leidschendam, Nederland (2006)
16. Staplin, L., Lococo, K. H., Martell, C. and Stutts, J.: Taxonomy of older driver behaviors and crash risk. (Report No. DOT HS 811 468C). National Highway Traffic Safety Administration, Washington, DC (2012)
17. Paris, J-C., Bellet, T., Cour, M., Marin-Lamellet, C., Deleurence, P., Moreau, F., Boverie, S., Andre, J-M., Claverie, B.: Driving assistances for senior drivers: a human centered design approach. In: Transport Research Arena - 5th Conference, Paris (2014). https://hal.archives-ouvertes.fr/hal-00985111
18. Bellet, T., Hoc, J.-M., Boverie, S., Boy, G.A.: From human-machine interaction to cooperation: towards the integrated copilot. In: Kolski, C. (ed.) Human-Computer Interaction in Transport, pp. 129–156. Ashgate, London (2011)
19. Save, L., Feuerberg, B.: Designing human-automation interaction: a new level of automation taxonomy. In: De Waard, D., Brookhuis, K., Dehais, F., Weikert, C., Röttger, S., Manzey, D., Biede, S., Reuzeau, F., Terrier, P. (eds.) Human Factors: A View from an Integrative Perspective. Proceedings HFES Europe Chapter Conference Toulouse (2012). http://hfes-europe.org. ISBN 978-0-945289-44-9
20. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **30**(3), 286–297 (2000)
21. Endsley, M.R., Kaber, D.B.: Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics **42**, 462–492 (1999)
22. SAE International: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Standard J3016-201609. Downloaded from SAE International by Michiaki Sekine, Sunday (2016)

23. Christensen, A., Cunningham, A., Engelman, J., Green, C., Kawashima, C.: Key considerations in the development of driving automation systems. In: Proceedings of the 24th International Technical Conference on the Enhanced Safety of Vehicles (ESV) (2015)
24. Schoettle, B., Sivak, M.: Motorists' preferences for different levels of vehicle automation: 2016 (Report No. SWT-2016-8). The University of Michigan Sustainable Worldwide Transportation, Ann Arbor (2016). http://www.umich.edu/~umtriswt/PDF/SWT-2016-8.pdf
25. Schoettle, B., Sivak, M.: Motorist's Preferences for Different Levels of Vehicle Automation. The University of Michigan Sustainable Worldwide Transportation Ann Arbor, Michigan 48109-2150 U.S.A. (2016)

Road and Rail: Rail Safety and Design

A Comparison of Three Systemic Accident Analysis Methods Using 46 SPAD (Signals Passed at Danger) Incidents

Kate Dixon[1(✉)], Patrick Waterson[1], and Jo Barnes[2]

[1] Human Factor and Complex Systems Group, Loughborough Design School, Loughborough University, Epinal Way, Loughborough LE11 3TU, UK
{K.L.Dixon, P.Waterson}@lboro.co.uk
[2] Behavioural Safety and Injury Prevention Research Group, Loughborough Design School, Loughborough University, Epinal Way, Loughborough LE11 3TU, UK
J.Barnes@lboro.co.uk

Abstract. During the period 1996–2003 there were five fatal accidents on the UK railway network, three of which were Signals Passed at Danger (SPAD) events (Watford Junction, 1996; Southall, 1997; Ladbroke Grove, 1999). SPAD events vary in severity and whilst most are not fatal there is the potential to cause serious injuries to passengers and train staff and damage to railway infrastructure. This paper investigates how the current system accident analysis tool used within the railway, the Incident Factor Classification System (IFCS) identifies and analyses causal factors of SPAD events. To evaluate the effectiveness IFCS was used to analysis SPAD incident reports (n = 46) and the outputs were compared with two systemic accident analysis methods and relevant outputs (the Human Factors Analysis and Classification System – HFACS and AcciMaps). The initial reporting process proved to hinder all systemic accident analysis methods in the extraction of causal factors. However, once extracted, all system accident analysis methods were successful in categorizing causal factors and demonstrated various outputs to illustrate the findings.

Keywords: Human factors · Complex systems · Railway · Signals passed at danger · SPAD · Incident analysis · Reporting

1 Introduction

A signal passed at danger (SPAD) event is defined as a train passing a red 'stop' signal into a section of track without authorisation to do so [1]. The consequences of a SPAD event can vary depending on the situation, they are categorised using the probability and severity of potential outcomes and the distance the train travelled past the signal. In the highest category (A), SPADs can lead to:

- Injuries and deaths of both passengers and train staff;
- Delays within the network costing the train/freight operating company (TOC/FOC) money and damaging their reputation;

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_103

- Damage to the train, track or signal costing the TOC/FOC/Network Rail money and reputation.

Lowe and Nock [2] estimated that each category A SPAD costs £22,000, therefore, in the year 2007 alone, the annual cost of category A SPADs was £7 million. The potentially severe consequences and regularity provide good reasons for investigating SPADs.

Findings from a literature review identified that research has been carried out to improve compliance of drivers and railway staff with reporting regulations and practices but there is very little research on how the reporting system is structured and its effectiveness. Several researchers have highlighted the need to learn from past incidents and near misses [3, 4]. Therefore, this paper will focus on the current SPAD reporting system using different analytical approaches.

1.1 Objectives

The objectives of this study are to:

1. Analyse a set of SPAD incident reports (n = 46) to determine common contributory findings
2. Apply and compare the outcomes from using three systematic accident analysis models to extract and identify common causal factors leading to SPAD incidents.

2 Method

Network Rail provided access to forty-six SPAD incident reports from eight different train operating companies (TOCs; n = 24) and six different freight operating companies (FOCs; n = 22). All the incident reports provided had already been assessed by Network Rail to calculate SPAD risk, all reports were reported as category A.

Currently the railway industry uses the Incident Factor Classification System (IFCS) to analysis incident reports, therefore this method of analysis was chosen along with AcciMap and the human factors analysis and classification system (HFACS). AcciMap and HFACS methods were chosen based on the findings of a review on system accident analysis methods and were deemed to be most suitable for the data given. Therefore, all three system accident analysis methods used the same forty-six incident reports to allow comparison of the three methods and their subsequent results.

IFCS was developed by a team at RSSB [5] and is used to classify human error and identify underlying factors relating to the incident. The aim of IFCS is to support and promote key human factor issues in relation to accidents, to allow cross industry learning via causal trends. The intended users are incident investigators, to allow incidents with similar causes to be easily accessible so previous work and recommendations are not repeated or contradicted. Validation of the IFCS was undertaken using 300 pre-evaluated incident reports and applying the IFCS before being implemented within the railway incident investigation process.

The IFCS analysis involves analysing an incident report, highlighting any factors which are human error factors or incident factors. These two categories have further sub categories. Any factor which can be learnt from and describes a system is an incident factor; any factor which is due to a human action is a human error factor.

RSSB have developed specialist computer software to aid the analysis and presentation of the IFCS results. Due to not having access to this software, the 46 SPAD incident reports were analysed by two investigators then combined to minimise the effects of subjectivity and bias. The relevant factors were provided to the researcher on a Microsoft Excel spreadsheet and imported into NVivo for further analysis. The results were then published in table format to allow visual representation of the sub-categories of the human error category and error type of the incident factor category.

The AcciMap method was developed by Svedung and Rasmussen [6] to graphically represent the 'causal flow of accidents, analyse hazardous work systems and…. interactions'. The AcciMap provides a high level of flexibility within designing the levels and categories/conditions, allowing it to suit a wide range of uses and domains and making it a very popular analysis tool to use [7].

For the AcciMap analysis factors were broken down differently to the IFCS analysis. Normally the bottom layer on an AcciMap is a timeline with an 'individual' layer above, however, due to the AcciMap needing to illustrate the findings of forty-six incident reports, a timeline level could not be used. Instead this level was replaced with an 'action layer' with an 'individual' layer above. The 'individual' layer shows all factors which relate to the individual, be that a driver, signaler or conductor. Above the 'individual' layer is the 'team/group' layer with 'management', 'organisation' and 'environment' representatively higher. After associating each factor with a layer, relationships were then illustrated via lines linking factors together. This stage was done on A3 paper before being imported into Microsoft Visio.

HFACS was developed by Shappell and Wiegmann and is heavily influenced by Reason's Swiss Cheese model [8]. Reason's Swiss Cheese Model [9] is described as four levels of human failure – unsafe acts (active failures), preconditions for unsafe acts (latent failures), unsafe supervision (latent failures) and organisational influences (latent failures). Each layer influences the next, either positively or negatively and when a failure occurs at each layer an accident occurs [10]. The aim of HFACS is not to place blame on one factor but to investigate causal factors.

The HFACS has four identical levels to the Swiss Cheese Model that can be further defined using categories similar to the IFCS. Unsafe act is level 1 and is further separated into errors and violations, using the terms skill-based errors, decision errors, perceptual errors, routine violations and exceptional violations. Level 2 is preconditions for unsafe act which is separated into three categories; environmental factors, condition of operators and personal factors. The next level, level 3, is unsafe supervision; this is broken into inadequate supervision, plan inappropriate operation, fail to correct known problem and supervisory violation. The final level to the HFACS is organisational influences, with resource management, organisational climate and operational process sub-categories.

The result of HFACS analysis is a graphic representation of the layers in which if one factor malfunctions or fails on each level then an incident can occur. However, by highlighting prior to an incident the risk each factor has on the chance of an incident

occurring, the risk can be reduced. Alternatively, retrospectively conducted HFACS will highlight the failed factors so work can be conducted to ensure it cannot happen again.

The forty-six incident reports were analysed and causal factors were categorised under the HFAC levels. To allow for editing of factors this was conducted first on a whiteboard before computerised using Microsoft Visio.

3 Results

3.1 Incident Factor Classification System

There was a wide distribution of the occurrence of factors throughout the six human factor categories and ten incident factors. As can be seen in the graph below (Fig. 1) the three human error categories which occurred the most throughout the reports are perception slip, decision error, route violation and action slip (n = 61, 35, 19 and 19 respectively). The most common incident factors were personal, communication and knowledge, skills and experience factors (n = 49, 38, 37 respectively). Factors related to the human error categories appeared 168 times through the incident reports, whilst 228 factors were associated with the ten incident factors.

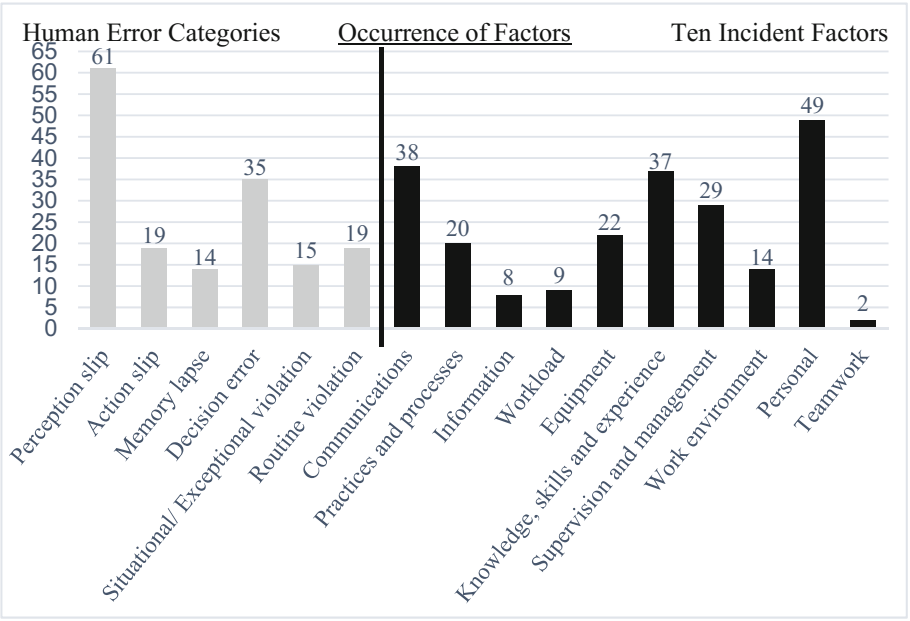


Fig. 1. Distribution of occurrence of causal factors

Figure 1 also highlights the factors which were not commonly reported within the incident reports. Teamwork appeared in less than 1% of the overall number of factors

for the incident reports. Information and workload combined covered a mere 7.5% of the total numbers. Whilst in the human error categories memory lapse was the least mentioned as a causal factor and only covered 8.3% of the total. Each of the categories were broken down further within the analysis to establish factors which were contributory to the incident.

Perception slip (n = 61) is split into 'error in interpretation of visual stimuli and 'error in searching for visual stimuli' (n = 35, 26 respectively). Most errors in interpretation of visual stimulus were due to the driver responding correctly to the incorrect signal (n = 23). Of these eleven cases were due to misreading or reading across the signal whilst twelve were due to the driver incorrectly predicting what the signal would be. Whilst the most common errors in searching for visual stimulus were more spread out across a variety of causes: failure to locate relevant signal/AWS magnet, speed restrictions and distraction by other lights were all the most popular causes (n = 5 for each group). In four occasions the driver claimed their focus was on the speed restriction rather than the signal, whilst one driver failed to locate the speed restriction for the section that the signal was in. Driver distraction due to light stimulus was also highlighted as an issue with three occasions of the signal being in the same line of sight as another signal and twice a driver became distracted by another light stimulus.

Decision errors were split into four different groups: error in decision, error in strategy, error in judgement and signal predicting (n = 8, 6, 7 and 14 respectively). Signal predicting is the highest occurring error and arises from a driver reacting to what they expect to see rather than what they see. For example, a driver may remember the normal pattern of signals within an area, so they react to what they are expecting to see, not what is there. Thus, they may not slow the train down appropriately, meaning it's impossible to stop before passing the signal at danger. A common theme within the decision error category is that errors are often based on the correct knowledge and information and sound recollection but the knowledge and information is applied wrong due to the situation.

The ten incident factors focus on management and system errors and the spread of occurrence amongst these factors is more variable than the human error factors. Personal factors were the most common and are those which can affect an individual's ability to maintain attention and focus resulting in fatigue and/or poor physical and mental wellbeing. Fatigue and distraction are the most common causal causes (n = 10 and 12 respectively) due to a range of issues covering complacency, overtime and family issues (n = 1, 2, 4 respectively). Interestingly, less breakdown was available of causes for this factor in comparison to the human error categories. Most reports which mentioned fatigue as a factor did not mention the cause of the fatigue (n = 6).

Issues due to verbal communication were separated during analysis to before and after the occurrence of the SPAD event. Errors which occurred before the incident predominately lie in communication between the driver and driver manager (n = 7), with only one case of communication error between driver and signaller. The most common errors were within communication of information and knowledge (n = 6).

However, after the incident all but one error was due to the communication between the driver and signaller (n = 13). The reports vary in detail but 77% of errors in communication between driver and signaller after a SPAD incident is due to poor or lack of communication. Miscommunication was also considered to be a causal factor

(n = 4) but it was not clear from the reports who was involved, just that it occurred before the SPAD event. Both before and after the SPAD occurrence there are cases where communication guidelines are not followed and either one or multiple people within the line of communication fail to repeat back what they have heard before continuing (before n = 7, after n = 11). This safe guard is in place to aid communication and prevent misunderstanding and poor communication levels.

3.2 AcciMap

The complexity of the railway system is reflected in the AcciMap analysis, with 150 causal human error factors and relationships depicted in the Human Error Accimap and 287 causal incident factors and relationships depicted in the Incident Factors AcciMap. The two factor groups were categorised into levels and then subsequently grouped into common factor groups, where possible any link to an IFCS factor was included to help in comparison of the two techniques (Fig. 2).

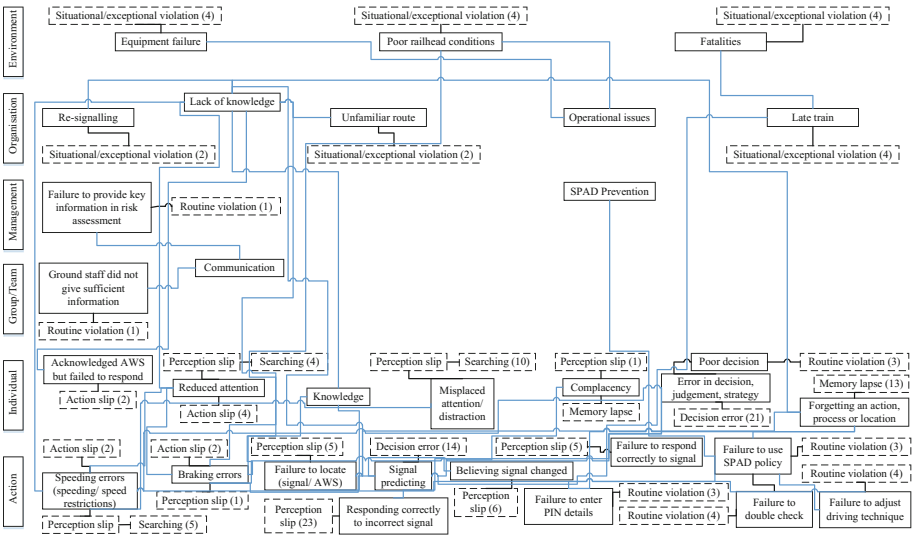


Fig. 2. Human error AcciMap.

The human error AcciMap illustrates a skewed distribution of causal factors amongst the levels, with more factors present in the action and individual levels. The action level contains eleven common factor groups with a subsequent fourteen related IFCS factors groups. In total seventy-seven contributory factors were categorised to this level. Responding correctly to an incorrect signal due to a perception slip and signal predicting caused by decision errors are the most common contributory factors at this level (n = 23, 14 respectively). Factors at this level are all direct causation contributory factors e.g. the final factor that caused the incident to happen. The individual level contains less common factor groups (n = 7) and less IFCS factor groups (n = 10)

and covered fifty-eight contributory factors. Decision errors within decision, judgement or strategy were the most common (n = 21) but within this error three factors were also routine violations. Errors in decision, judgement and strategy were linked to signal predicting, speeding and braking errors within the action level.

The remaining fifteen contributory factors were spread amongst the remaining levels – group/team, management, organisation and environment. The most common IFCS factor group was situation/exceptional violation – violations of procedures and rules which are out of the driver control such as issues with late trains, equipment failure and poor railhead conditions. This skew in the presentation of contributory factors illustrates that many of the contributory factors are related to the individual or in this case, the driver and can be seen in Fig. 2.

The AcciMap also highlights the relationships between common factor groups. Although interestingly there are limited relationships on an intra level (common factor group on one level with another common factor group on same level) throughout the AcciMap. Many relationships are on an inter level (common factor group on one level with another common factor group on a different level) illustrating that a factor can influence factors on other levels (Fig. 2).

The incident factor AcciMap is a lot more complex diagram which at a glance shows no skewing towards any level. In comparison to the human error AcciMap there is an absence of factor groups within the action and group/team level (n = 4, 6 respectively), however these two levels still include 58 contributory factors across 14 IFCS factor groups (Table 1).

Table 1. Distribution of factors in Incident Factor AcciMap

Level	Common factor group	IFCS factors	Contributory factors
Environment	6	8	13
Organisation	14	20	42
Management	14	13	60
Group/team	6	8	14
Individual	19	28	114
Action	4	6	44

The individual level contains the highest amount of common factor groups, IFCS groups and causal factors (n = 19, 28, 114 respectively). No common factor group is duplicated from the human factor AcciMap and the most common IFCS groups are personal and communication. Personal factors include needing a cold drink, fatigue and health and commonly were related to communication due to these personal factors often not being communicated until an incident has occurred. The abundance of relationship lines to and from this level illustrates the influence factors on this level have on the factors of other levels.

Unlike the human factors AcciMap, the management, organisational and environment levels are more heavily populated. Again, the IFCS factor of communication is common within the management level, often being attributed to issues with information. Communication issues were most likely to be between driver and signaller

(n = 15) than driver to another railway colleague - driver and driver manager (n = 7), driver and conductor (n = 1). The management level highlights issues with training, non-technical skills and scheduling/work load.

Issues with reporting and equipment faults within SPAD protection equipment, trains, railhead, brakes and signals are organisational issues due to the financial and organisational aspects behind these factors. Maintenance of current infrastructure and finance to upgrade that infrastructure requires planning from a higher level than local management. Therefore, unsurprisingly the most common IFCS factors on this level are equipment factors and supervision and management factors.

The environment level demonstrates factors outside the drivers control which can contribute to an incident. All common factor groups within this level shared the same IFCS factor of work environment, only thirteen examples of the work environment being a causal factor were found within the analysis. However, this level has relationship lines to all the other levels, demonstrating how potentially powerful this level is. Weather conditions and cab design were common factor groups which are interlinked due to temperature, noise and design being common causal factors. It appears from the reports that at times of extreme weather the cab design is less than favourable and can cause distraction at an individual level. Lighting, both internal and external could also be causal due to the interference with signal sighting.

The incident factor AcciMap illustrates a variety of intra and inter relationships, demonstrating that a contributory factor can affect factors both within the same level and different level. Therefore, indicating that the incident factors have a higher ability to influence more contributory factors.

3.3 Human Factors Analysis and Classification System

A graphical representation of the HFACS is presented in Fig. 3, demonstrating how incidents occur due to the influence of multiple factors across multiple levels.

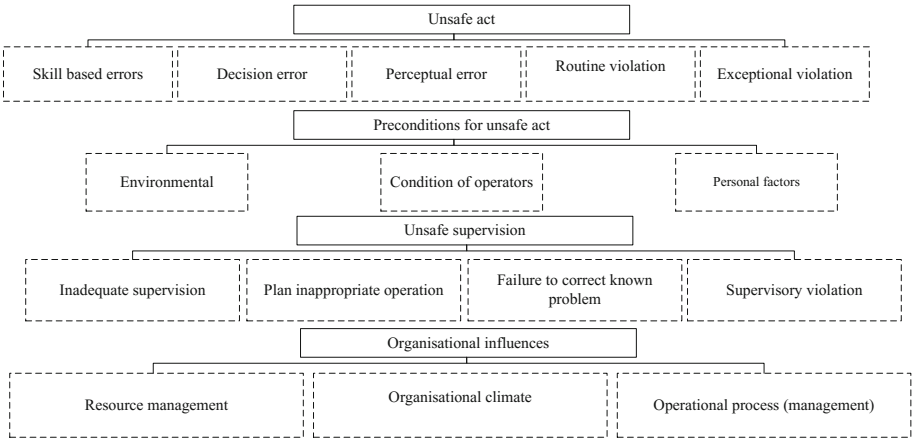


Fig. 3. HFACS

The 'unsafe act' is the final defence layer meaning these are often described as the main causation for an incident as it's the most obvious and what people can see. For example, with the use of on-board train recorders it is simple to see if a braking or speed error has occurred. Also, drivers within interviews may try to explain why the incident happened, often resulting in the cause being stated due to a specific driver error. The HFACs analysis categorises potential errors into five main categories which coincidentally coincide with the IFCS factors; skill based errors, decision errors, perceptual error, routine violation and exceptional violation.

For there to be an error to occur within the 'unsafe act' level there must have been weakness or failings within the previous three levels – preconditions for unsafe act, unsafe supervision and organisational influences. Preconditions for unsafe act in relation to signal passed at danger events include both factors outside the driver's control (environmental factors, condition of operators) as well as personal factors.

Yet prior to this an error must occur at both the 'unsafe supervision' and 'organisational influences' levels. At the highest level (organisational influences) errors include planning, budgets and safety culture/climate and can occur on a national scale. Whilst 'unsafe supervision' errors can occur on a local scale due to inadequate supervision or supervisory violations, failure to correct a known problem, or planning an inappropriate operation.

The HFACs model highlights that it is possible for the driver to be doing everything correctly and a SPAD can still occur. For example, issues with signal sighting, (preconditions for unsafe act), can lead to the driver not being able to locate the signal (unsafe act) as the view is being obstructed by foliage or another signal. This stems from the 'unsafe supervision' layer of 'failing to correct a known problem'. But can be further linked to the first defence layer of 'organisational influences' by poor planning that the foliage has not been trimmed back (Fig. 3).

4 Discussion

The incident reports were found to vary in all aspects from detail and content to the reporting structure, and also within the same train or freight operating company (TOC/FOC). This proved a limiting factor for all analysis methods as analysis could only take place on what the railway investigator had written.

The style of writing within the incident reports varied from a 'keepsakes' or note style to detailed, cohesive writing. Whilst both styles are valid reporting styles, the former aids those who have prior knowledge of the incident, and hinders those who do not, as information or description of the incident is limited. Reports in this style were often found to state reasons rather than investigate reasons, yet this could be due to this information being missed from the report. This 'unwritten' content may be implicit for those 'in the know' but difficult for an unconnected reader to understand. For example, in one report under 'immediate cause' the following was written 'the immediate cause of this incident was the driver passing…at danger without authority'. Whilst providing a definition of a SPAD event it failed to inform the reader of the immediate cause to the SPAD.

The structure varied amongst TOCs and FOCs and sometimes varied between the same TOC or FOC. Non-technical skills and IFCS analysis were included in some reports but was not a common feature throughout the reports, in some cases the sections were left blank or the writer confuses the terminology between the non-technical skill terms and IFCS terms.

Therefore, due to the reports being written by experienced employees within the railway network, experienced readers will be able to understand the unwritten assumptions, however readers with limited experience could link factors which, in reality, cannot be linked. However, despite these drawbacks within data extraction, all three analysis methods were fruitful in data extraction and categorisation of the subsequent causal factors. The IFCS extracted 396 causal factors from the 46 incident reports, whilst the AcciMap analysis extracted 437 causal factors. The HFACs analysis does not quantify any factors but was efficient in categorising the causal factors.

The IFCS and HFACs analysis provided criteria for categorising factors whilst the AcciMap analysis provided a framework to place causal factors. Due to the subjective nature of extracting and categorising the factors from the incident reports, there were two coders to ensure that each extraction was associated with the correct factor for all the methods of analysis. The criteria provided by ICFS and HFACs varied in terms of detail, with ICFS providing the most detailed criteria for categorising. Whilst this initially aided analysis, there were cases where a causal factor did not fit 'neatly' into a predetermined term or fitted between two or more terms, groups or levels. When similar cases arose in the AcciMap analysis, the causal factor was placed between levels to illustrate it should be considered on both levels. The predetermined groups within the HFACS analysis are the easiest to use as it allows for duplicate factors to appear on the same or other levels. This duplication of factors or presence of very similar factors on any level suggests that this factor is a critical causal factor worthy of further investigation.

It was evident from the results that the methods varied in terms of presentation and degree of detail. The AcciMap was the only method of analysis to illustrate relationships on a multiple scale which allows an expanded view of which factors can affect the other, albeit not directly showing how they can. Whilst this is positive on an AcciMap analysis containing few factors, it can complicate and hinder understanding when analysing large numbers of factors as seen in the incident factor AcciMap (Fig. 2). IFCS and HFACs do not illustrate relationships but in the case of IFCS this could be caused by not having access, therefore not using, the official IFCS analysis software and database. HFACs analysis does not illustrate relationships with lines, which may be considered an advantage as the user/reader will not be prohibited by visual lines that may enable a more thorough investigation of the incident.

Whilst the HFACS analysis, arguably presents the degree a contributory factor can influence other contributory factors, all system accident analysis methods showed this to some degree. The HFACS analysis illustrates that whilst a factor may be the factor that causes an incident, it is not the only contributory factor. For failing to occur on the 'unsafe act' level there must be issues and failures on the other three layers. The AcciMap illustrates this with the relationship lines and the IFCS demonstrates it by the inclusion of the same factors under different categories between the human error factors and incident factor groups.

5 Conclusion

The incident factor classification system (IFCS), human factor analysis classification system (HFACS) and AcciMap methods are successful in highlighting causal factors to recent signal passed at danger events on the UK railway network. However, as Gibson et al. [5] states when introducing the incident factor classification system; the system has been designed so it can be altered to be adaptable to the changes within the network. Presuming that no changes to the incident factor classification system have been made then although it's a modern system the addition of further, less ambiguous guidelines and definitions could remove some of the subjectivity.

All the methods have different advantages and disadvantages. A weakness within the incident factor classification system is the inability to analyse and record interaction between the factors. This is an advantage of the AcciMap analysis but also a weakness, as with complex systems such as the multifactorial nature of the railway the diagram can get very complicated and confusing. The HFACS analysis does not depict the relationships either but provides the knowledge there must be failures on the previous defense layers. A weakness however is that no background or context is provided.

What must be considered when investigating incident reports and methods of analysis of incident reports is that what may be the best way of practice for someone looking into the system may not suit those that are actually working in the system. The findings highlight that there is scope to improve the railway reporting system which will in turn aid system analysis. Furthermore, the incident factor classification system could be combined with elements from the AcciMap HFACS analyses to strengthen this system analysis model. Further studies are underway exploring the causal factors of SPAD events and the differences within reporting styles amongst TOCs using interviews with key stakeholders.

References

1. Naweed, A.: Psychological factors for driver distraction and inattention in the Australian and New Zealand rail industry. Accid. Anal. Prev. **60**, 193–204 (2013)
2. Lowe, E., Nock, P.: Changing safety critical communications behaviour. In: Norris, B., Clarke, T., Mills, A. (eds.) People and Rail Systems, pp. 399–407. Ashgate Publications, Aldershot (2007)
3. Speirs, F., Johnston, C.: Signals passed at danger: a case study in the application of visualisation techniques. http://www.dcs.gla.ac.uk/~johnson/papers/Speirs/spad1.pdf
4. Wright, L., van der Schaat, T.: Accident versus near miss causation: a critical review of the literature, an empirical test in the UK railways domain, and their implications for other sections. J. Hazard. Mater. **111**, 105–110 (2004)
5. Gibson, W.H., Smith, S., Lowe, E., Mills, A.M., Morse, G., Carpenter, S.: Incident factor classification system. In: Dadashi, N., Scott, A., Wilson, J.R., Mills, A. (eds.) A Rail Human Factors - Supporting Reliability, Safety and Cost Reduction, pp. 653–658. Taylor and Francis, London (2013)
6. Svedung, I., Rasmussen, J.: Graphic representation of accident scenarios: mapping system structure and the causation of accidents. Saf. Sci. **40**(5), 397–417 (2002)

7. Waterson, P.E., Jenkins, D.P., Salmon, P.M., Underwood, P.: 'Remixing Rasmussen': the evolution of Accimaps within systemic accident analysis. Appl. Ergon. **59**(Part B), 483–503 (2017)
8. HFACS. https://hfacs.com/hfacs-framework.html
9. Reason, J.: Human Error. Cambridge University Press, Cambridge (1990)
10. Shappell, S.A.: The human factors analysis and classification system – HFACS, pp. 1–18. Office of Aviation Medicine, Washington (2000)

Understanding Railway Employees' Perceptions of Senior Managers' Safety Commitment

Kate C. Bowers(✉) and Mark Fleming

Psychology Department,
Saint Mary's University, 923 Robie Street, Halifax, NS, Canada
{kate.bowers,mark.fleming}@smu.ca

Abstract. Numerous empirical studies have concluded that employees' perceptions of managers' commitment to safety is related to important safety outcomes, including injuries. Specifically, employees who perceive that managers prioritize safety over other competing demands are more likely to comply with rules and participate in safety. This is especially important in the railway industry as many employees (e.g. train engineers) are not directly supervised. The present study utilized qualitative methodology to investigate how employees perceive senior managers' commitment to safety. Results suggest that a variety of sender, message, and receiver characteristics interact to impact perception formation. Participants' descriptions of senior manager safety commitment included representations of engaged safety leadership, consistent safety leadership, the allocation of finances to safety, and implementation of policies and procedures that reflect a value for safety. Suggestions for practice are discussed.

Keywords: Manager safety commitment · Safety culture · Signaling theory · Employee perceptions · High-Risk industry

1 Introduction

Driven by the significant human and financial cost, the priority for safety in the railway industry is paramount. Statistics from the Association of Workers' Compensation Boards of Canada indicated that across all industries, a total of 232,629 lost time claims and 852 fatalities were reported in 2015 with 15,538 lost time claims and 74 fatalities specific to the transportation and storage industry [1]. Understandably, significant research has been conducted to understand the factors that influence safety outcomes.

Findings from major disasters highlight the importance of management in shaping safety outcomes [e.g., 2–4]. More specifically, employee perceptions of managers' commitment to safety is related to important safety outcomes, including injuries [5]. Employees who perceive that managers prioritize safety over other competing demands are more likely to comply with rules and participate in safety [6].

The influence of management safety commitment on safety outcomes has largely focused on the influence of direct reports; however, recognition that different levels of management uniquely impact employee perceptions has been supported [7].

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_104

Interestingly, although recognized for their substantial contribution to organizational functioning (e.g., financial outcomes), senior managers have been largely neglected from empirical research regarding safety outcomes [8–11]. The role of perceptions of senior managers' safety commitment is perhaps most critical in high-risk organizations where direct supervision of employees is limited (e.g., train engineers). Consequently, this study takes a first step to understand how senior managers' safety commitment is perceived.

2 Unique Influence of Senior Managers

Numerous research findings support the notion that levels of management uniquely impact employee perceptions. Clarke's study found that frontline employees (e.g., train drivers) perceived local managers and supervisors as more concerned about safety than senior managers [12]. Additionally, Zohar and Luria's work supports that variation in climate exists among different departments of an organization, prompting intervention models to be shaped for the targeted level [13]. More recently, Zohar identified that employees develop level-specific climate perceptions, and thus, there is a need to develop level-specific subscales for measures of climate [14].

The importance of senior managers in shaping culture has grown in response to the Deep-Water Horizon accident in 2010. A report produced in response to the accident identified that an organization's Chief Executive Officer (C.E.O.) and board of directors must create a safe culture whereby all employees feel responsible and motivated to prevent accidents [4]. In response to this document, publications have begun to investigate senior manager commitment behaviour, including how this position shapes safety culture, and how managers promote safety commitment.

In the few studies that have investigated senior manager safety commitment, results suggested that problem solving, specifically the number of issues considered, information sources used, generation of ideas, social competence, personal involvement in safety matters, influence on organizational attributes, and communication about safety reflected senior manager commitment to safety [10]. Additionally, perceiving and understanding employees intentions and emotions were deemed a reflection of commitment [10].

A similar study identified six attributes demonstrating commitment to safety: (1) social competence (understanding others emotions and ability to persuade others), (2) safety knowledge (theoretical and practical understanding of safety issues, knowledge of facts and information), (3) motivation (goal motivation via context appropriate promotion or prevention focus; see Crowe and Higgins [15]), (4) problem-solving (understanding the problem, generating ideas and planning the implementation of ideas), (5) personality, and (6) interpersonal leadership skills (transactional, transformational, and authentic leadership characteristics) [11]. These attributes were defined as senior managers' safety intelligence, representing their understanding of safety issues and relevant policies regarding safety [11].

The aforementioned research highlights a variety of senior manager behaviours representing safety commitment, however, because these findings were derived from a samples of senior managers, it is unclear if employees perceive these behaviours to be a reflection of safety commitment. Using the aforementioned examples and considering

senior managers in a large transportation organization, Table 1 represents the channels of influence through which employees receive information from senior managers. This table provides a preliminary framework to represent the types of information relayed to employees reflecting senior manger safety commitment.

Table 1. Channels of senior manager influence

Channel	Description	Examples
Manager decisions	Senior manager decisions that produce observable outcomes reflecting safety commitment	Allocation of resources (time, money, people), response to safety events, how safety practice is rewarded, written statements about safety
Manager behaviours	Senior manager actions that are directly observable and reflect safety commitment	Wearing safety equipment, abiding by safety rules/policies, statements about safety
Other information	Senior manager decision making or behaviour reflecting safety commitment that is relayed by an indirect source	Stories from colleagues, friends, supervisors, article in a magazine or newspaper, union discussion, recognition by industry peers (e.g., safety award), news story

3 Signaling Theory

Expanding on the framework represented in Table 1, the present study adopts signaling theory to illustrate how perceptions of senior managers' safety commitment are formed. Signaling theory was first described in economics literature to describe information processing between two parties [16]. Signals are typically used when one party is privy to information that another party is not, commonly referred to as information asymmetry. When this happens, the informed party can choose to send a signal in order to communicate the unknown information and the receiving party must choose how to interpret this signal. For example, Spence described that job applicants will use costly degrees of higher education to signal their superior quality over other candidates [17]. Popular, more simplistic examples of signaling, include using an expensive sports car to signal wealth or wearing business attire to signal a degree of professionalism. The complexity in signaling lies in the challenge of ensuring that the receiving party will identify the signal as credible.

Recently, signaling theory has been applied in management literature to frame information processing between leaders and their reports [18]. Adopting this framework, the principle of signaling theory can be applied to manager safety commitment whereby managers use signals to demonstrate their value for safety. Employees must infer managers' safety commitment by observing signals, typically through observations of managers' behaviours and decisions. In organizations where employees have limited opportunity to directly observe senior manager behaviours and decision-making (e.g., senior managers in large organizations), managers must be strategic to ensure the signals used to relay their value for safety are appropriately received. Connelly and colleagues

suggest four characteristics that can impact signal interpretation: (1) signaler characteristics, (2) signal characteristics, (3) receiver characteristics, and (4) feedback/environment [18]. Applying this framework in the context of information processing between senior managers and frontline employees, the success of employees perceiving authentic senior manager safety commitment (i.e., credible signals) will depend on:

- perceived characteristics about the senior manager (e.g., personality characteristics);
- characteristics of the signal (e.g., medium used to relay the signal, frequency, cost);
- the frontline employee's ability or willingness to receive the message (e.g., competency, attention); and
- the environment or opportunity for feedback (e.g., existing culture; opportunity to clarify understanding by asking questions).

This framework and the aforementioned channels of senior manager influence generate the research question of this study: how do employees perceive senior manager safety commitment?

4 Method

In order to utilize a methodology sensitive to underlying meaning in gathering and interpreting data, the present study adopted qualitative methodology using one-on-one interviews with frontline employees to understand perception formation of senior manager safety commitment.

4.1 Participants

Fourteen participants from two worksites at an international transit organization were recruited for this study using convenience sampling. Inclusion criteria included being English speaking and employed with the organization for at least 6 months.

4.2 Materials and Measures

A semi-structured interview was used to guide participant interviews as this method provides flexibility to evaluate participant's responses and probe appropriately. The semi-structured interview guide was created using a modified version of Flanagan's critical incident technique [19]. Examples of interview questions included: "think of a manager who is committed to safety. How do you know this senior manager is committed to safety?" Prompts were used to stimulate discussion regarding perceptions of senior manager commitment behaviour. A demographic survey was also used to collect participant's age, gender, and current employment role. Interviews were recorded using an IPhone 4s and stored on a password-protected laptop computer. Microsoft Word© and Microsoft Excel© software was used for data analysis and word processing.

4.3 Procedure

Participants were recruited by approaching management of an international transit organization and requesting interview sessions at two worksites. Interviews were conducted within one week. Participants were given a consent form detailing voluntary participation, study content, and confidentiality of the data. Verbal consent was confirmed prior to audio recording. Interview content included 15–20 min of questions related to a pre-existing safety initiative at the organization, followed by 40–45 min of manager safety commitment discussion. Interview guides were used during interview sessions to ensure consistency in interviewee experience. At the end of each interview session participants were given a feedback form thanking participants for their participation and providing researcher contact information. Data saturation was ensured by following principles outlined by Francis et al. [20]. The proposed study involved no direct threat or risks to participants. To mitigate risk of employees feeling uncomfortable when discussing manager safety commitment or lack of safety commitment, identifying information was omitted in all research documents produced from the data.

4.4 Qualitative Analysis

This researcher most closely aligns with the postpositivist paradigm in conducting qualitative research. At a high level, this perspective maintains a scientific approach to research and uses a logical, step-by-step process in analysis whereby qualitative studies are typically written in the form of scientific reports. Thematic analysis followed five steps outlined by Braun and Clarke, including (1) data familiarization, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, and (5) defining and naming themes [21].

5 Results

The purpose of this study was to understand how employees perceive senior managers' safety commitment. Participant ages ranged from 21 to 66 years with a mean age of 40. Eight participants (57%) indicated having met a member of senior management in their present organization. Experience in the transportation industry varied greatly among participants, averaging 8 years and ranging from 1 year to 30 years. Participants were all English speaking and average interview time across the 14 interviews was 46 min.

5.1 Themes

Engaged Safety Leadership (versus Unengaged). Participants described senior manager safety commitment as behaviours demonstrating engagement in worksite safety management, an open and approachable demeanour, and respect for employees. This included senior managers visiting the worksite and employees, demonstrating a concern for safety by inquiring about process, and employees feeling able to approach

and converse with this level of management. For example, one participant stated senior manager safety commitment meant:

> "walking, walking around, walking on the property. […] going out to the job sites, maybe spending an hour or two with the crew, section crew and their vehicle, whether they're going to a work place, just get on the work site, do a quick job review with them. How did this go, how did that go, what's the process?"

Participants also stated that commitment was a reflection of senior managers who dedicated time to visit the worksite and ask questions in order to make informed decisions:

> "…get involved, like and I don't know, I've never met them so I can't say they probably have lots of stuff they're doing, I understand that, they're busy guys but if you're making decisions that big, they need to know what they're deciding on, not just look at a couple of pieces of paper that have a budget on it and say no that's not going to work sorry. Like find out do these guys need this to be safe?"

Senior managers' openness to accepting questions and concerns was also a consistent description among participants. In describing how senior managers could improve their commitment to safety, a participant described providing means for following up on a raised safety concern:

> "…if we could see some of that stuff and maybe given a phone number even above or whoever is in charge of making the decision, like I'm sure he's probably a ways up the management pole so he doesn't want a bunch of phone calls from all these people but if we had some way of getting in touch with somebody that's involved with the decision and see like what's their reasons…."

Aligned with this notion was mention of feeling respected and valued by senior managers. A participant described that committed senior managers will practice and ensure that lower level management "…won't ask somebody to do something that they personally wouldn't do themselves." As well, one participant described his opinion of the organization's C.E.O., noting the man's value and respect for individuals:

> "every month, yah every month we get a note, [the C.E.O.] on email, and it's his own personal writings, he and everyone I've talked to that has met him, couldn't ask for a nicer person. And he wants to talk about the person, not talking about the job, not talking about what happened yesterday or today, he will maybe at the tail end, find out about you, what makes you happy, what's your life story."

Consistency in Safety Leadership (versus Inconsistency). Participants described that senior managers who practice what they preach and follow through on statements made in response to raised safety concerns demonstrate a commitment to safety. One participant stated:

> "well like when you're, say you're somewhere that needs personal and protective equipment, they're the ones first to have their hard hats on, their safety glasses or vests, they're not getting out of the truck with their crocks on you know and their sweaters, or their coffee in their hand saying hey you don't have your safety glasses on right. Someone who's basically saying you know this is the way it's supposed to be done, this is the example that I'm setting and now you follow, that's what I'd see."

Consistently prioritizing safety over production was also mentioned as a demonstration of safety commitment:

> "I guess they don't compromise like the situation gets tough, they don't compromise and say O.K. well let's just do it this one time. It's always, they put their foot down and they say no this isn't safe we're not going to do it. […] and I think that you need somebody who has the ability to put their foot down and say no like this isn't right, we're not doing it."

Many respondents described that senior managers' alignment in statements and actions is the strongest indicator of safety commitment.

Allocating Money to Safety (versus Neglecting to Allocate). Senior managers were often linked with financial responsibility, "well a C.E.O., they have a lot to do with the financial aspect… and income, and so, that's their main priority I guess." Further to this finding, allocating money to maintenance and purchasing of equipment, safety programs, and other safety causes were described as reflections of senior manager's safety commitment. For example, one employee stated:

> "a company that was truly committed to safety I think they'd be a little more willing to spend the money needed on equipment or tools or training. Training especially actually now that I think of it, on how to be safe, on proper procedures for things, stuff like that."

Policies and Procedures Reflect a Value for Safety (versus Disregard for Safety). Last, senior managers' commitment to safety was described as a demonstration of organizational policies and procedures that reflect a value for safety. This included formal plans implemented by senior managers in order to ensure safety related processes in the organization, such as the existence and culture that promotes the use of engaging in a stop-work policy:

> "…guys can reject work for sure and they're aware of that. And if they don't feel safe, they don't do it and there's no manager here that will make you do it if you don't feel safe. That's on a guy who can look at you face to face, you know what I mean, he'll be face to face and say listen I'm not doing that today I don't feel safe…"

Implementation of safety rules in response to accidents and injury was also described as an indication of senior managers' safety commitment:

> "what it seems like if anything happens anywhere in the system, well we have a new safety rule. Well somebody got hurt […] and O.K., well, we're going to change the safety, this is the way it should be, which I understand."

Participants described that requiring a formal approval process at the executive level for the implantation of new safety rules reflected senior managers' safety commitment:

> "I mean like I said these are the rules, and I know that [the C.E.O. is] involved on different committees like when there'd be a different procedure or safety rule that would come out, that it has to go to him to be approved right, it just can't be, one guy, I can't just ask a guy: O.K., can I do it this way now? It has to be, it has to go through the approval committee before we can change a process or procedure, so it goes right to the top…"

Last, participants described that safety focused memos and reports sent to all staff in the organization reflect senior managers' safety commitment. However, there was a substantial divide in participants' opinions of reports. A minority of participants stated that they read the reports and appreciated the information within, while a majority stated they did not read the information nor care for what was included. One participant who was positive about such reports stated that 59 safety tips were posted regularly, demonstrating managers' focus on safety: "they're always putting up safety tips of the week and stuff like that."

Authenticity. In responding to the question "what factors demonstrate authenticity of managers' safety commitment?", participant responses reflected three primary themes: (1) seeking to understand, (2) prioritizing safety before productivity, and 3) existence of a personal relationship.

Seeking to understand was described as a strong indication of authentic commitment to safety. One participant described that authenticity was a reflection of knowledge, particularly in asking questions when one is not knowledgeable, as well as taking the initiative to speak up when an unsafe act is witnessed:

> "a prime example right now is my assistant […] supervisor, […] she used to run her own, she used to have her own company, private company […] and when she got into [this …] industry she had no clue. But now she's picked my brain apart every day for the last five years […] she's picked other supervisors, other foremen, and past supervisor branches apart, she's a good supervisor. She's really authentic when it comes to safety and promotes it. If she sees something wrong, she's not afraid to give a friendly blast to the boys you know, stop right now, you know, see what you're doing wrong."

Specific to senior managers, one participant described the importance of safety being consistently prioritized over production:

> "…actually seeing them get in there and do that, our type of work, you would never see it but at the same time, if they asked us to do some work that they wouldn't do themselves, it's kind of hard to see anyway but it's easy for them to say one thing and mean another. It's like oh yah we don't want you to do that work, it's unsafe but really they're pushing the productivity envelope so much that it's like there's no other way for it to happen. If you want productivity, you're asking us to do this unsafely. So that's definitely upper management type thing, not our direct supervisor."

Last, participants described the importance of having a personal relationship in order to determine authenticity. One participant described that authenticity can best be determined by having a relationship with an individual and the experiences and discussions had with that person. In reflecting on his supervisor's authenticity, the participant stated "Well I know my supervisor fairly well, you can hear it in his voice. […] You can see there's concern on his face."

Another participant described a similar notion by referring to the importance of time and consistency in statements and actions in determining authenticity: "how do you learn to trust somebody? Just by, over time and they don't lie to you and they're straight up with you…."

In discussing authentic safety commitment of the organization's C.E.O., there were mixed responses. Some participants described an inability to determine authenticity of senior managers due to a lack of relationship, as described by one participant:

> "…you know once in a while he'll send a letter out that, you know, we're committed to safety and we want to do better and stuff like that, and you're thinking O.K., maybe he's serious, but, then you're thinking, the guy's just a picture you know, you never see him so you don't know if he's there for, you know he's there for the money but is he committed to this or he's just blowing wind here."

For other participants, although a relationship did not exist, inferences about senior managers' authenticity were made through indirect sources (newsletters, magazine articles, media interviews, information from others). One participant inferred authenticity from experiences observing the C.E.O. speak in the press and reading about the individual in the newspaper: "well he sounds like he's really, like I say he's genuine, he sounds like he's a real person. He doesn't talk or sound like an executive, he'd probably be the guy sitting next to you at lunch."

6 Discussion

6.1 Study Purpose

The purpose of this study was to better understand how employees form perceptions of senior managers' safety commitment. This includes the types of signals that inform perceptions of safety commitment and those that are most credible (i.e., discerned as authentic). The findings from this study provide a deeper understanding of employee perceptions as they relate to manager safety commitment and contribute to the development of management practices that promote positive safety outcomes.

6.2 Answering the Research Question

How do employees perceive senior manager safety commitment? In short, results suggest that senior manager safety commitment is interpreted through direct experiences (e.g., personal interactions with senior managers) and indirect experiences (e.g., making inferences about senior managers' decision-making through work experiences). Themes suggest that employees form perceptions from senior managers' engagement in safety leadership, consistency in safety leadership, allocation of resources to safety, and development and implementation of safety related policies and procedures.

Interpreting these results in context of signaling theory, senior managers who are personable and focus on safety at the individual level are more often perceived as authentically committed to safety. Signals that were consistent and exchanged in person were supported as the most likely to be interpreted as credible (versus through newsletter or e-mail correspondence). Furthermore, senior management messaging relayed in a manner that uses language suitable for all staff to understand and that focuses on the human element of work (versus business) emerged as factors that support employees' ability or willingness to receive the message as a demonstration of

safety commitment. Last, environment appears to impact success of signal interpretation whereby the existing climate (e.g., shared perceptions of supervisors safety commitment; shared perceptions of senior manager safety commitment) can impact the interpretation of senior manager signals. For example, when employees perceived a lack of safety commitment by senior managers, incoming e-mail messages promoting senior managers' commitment to safety were largely discarded.

6.3 Practical Implications

Most importantly, results support that senior manager efforts to build personal relationships with employees and demonstrate individualized consideration and value for employee safety most strongly signal senior managers' safety commitment. These findings are consistent with previous research (e.g., [10, 11]) and facilitate deeper understanding by considering the employee perspective. Thus, senior management should strive to ensure they maintain a physical presence within the organization in order to build relationships, particularly at the frontline level. Understandably, this initiative may be challenging for senior managers in large-scale organizations with thousands of employees. Where opportunity to engage in individual-level relationship building is limited, senior managers can drive this approach with other levels of management. Senior managers can also work to ensure existing policies and procedures supporting individual safety are appropriately maintained by other levels of management (e.g., stop work policy). Similarly, because production pressure was largely attributed to senior manager decision-making, senior managers need to take thorough steps to ensure other levels of management maintain a priority for safety over production and that plans and priorities are not imposing unreasonable expectations. This includes taking steps to assess and understand frontline employee perceptions to adequately determine if production pressure is perceived at this level and the nature of these perceptions.

Seeking to understand was also supported as a signal of senior managers' safety commitment and emphasizes the importance of this factor in previous findings [see 11]. Senior managers can work to ensure decision-making at a strategic level is supported by leanings from across the organization and that this leaning-based decision making is appropriately communicated to staff. Additionally, senior managers should drive this approach in lower level management in order to better promote perceptions of an organization that seeks to understand in order to make informed decisions.

The presence (versus absence) of resources at the frontline level was also supported to signal senior managers' safety commitment. As previously mentioned it is important that (1) senior managers support other levels of management to appropriately prioritize safety (including confirming that resources are well maintained and in safe operating condition) and (2) senior managers practice due diligence to ensure resources are adequate by assessing frontline employees perceptions in a manner that ensures anonymity. Using an anonymous assessment method provides a more comfortable reporting environment for staff and better ensures the results are an accurate reflection of the workers' reality.

Last, consistency between senior managers' statements, actions, and organizational outcomes was identified as a critical element in discerning credibility of senior

managers' commitment to safety. This finding lends to the support of the above-mentioned initiatives, particularly in the case of fostering these approaches in other levels of management to ensure outcomes perceived at the frontline level are aligned with senior managers' advertised vision.

6.4 Study Limitations and Conclusion

Limitations exist in the present study. First, for convenience of sampling, interviews were conducted at the worksite. It is possible employees engaged in impression management when describing management who are committed or not committed to safety [22]. Attempts to minimize participant impression management were made by explaining verbally and in writing the confidentiality of information disclosed in interview sessions. Second, because data collection occurred during one-day worksite visits, individual interviews were limited to a maximum of one hour. Interviews averaged around 46 min; thus, it is possible that time pressure impacted some participants ability to provide comprehensive answers for all interview questions.

This work provides a first step in understanding the signals of senior manager safety commitment, however, it is recommended that organizations use these findings as a reference and adopt investigative methods to fully understand the signals unique to their workplace. The recommended investigative methods include interviews and focus groups. When used appropriately, these methods can identify rich information, subsequently aiding in the identification of signals that support or detract from perceptions of safety commitment. Subsequently, this information can be used to develop initiatives tailored to the individual organization that support the processing of credible signals of safety commitment. Additionally, surveying senior managers' perceptions of safety commitment to identify alignment between manager and employee perceptions may be used to further enhance understanding and initiative development.

In conclusion, this study contributes to the organizational safety evidence-base in providing a greater understanding of the process of senior manager perception formation. As well, the study findings contribute to organizational practice through improved understanding of senior manager safety leadership and betterment of organizational initiatives to foster perceptions of senior manager safety commitment. These findings prompt additional research in this area to further understand credible signals of senior manager safety commitment and the applicability of signals across organizations of different size and type.

References

1. Association of Workers' Compensation Boards of Canada (AWCBC): Number of Accepted Time-Loss Injuries, by Industry and Jurisdiction, 2013–2015 (2015). http://awcbc.org/?page_id=14
2. INSAG: Tecdoc 1329 – Safety Culture in Nuclear Installations. (Safety Series No.75-INSAG-4). Vienna: International Atomic Energy Agency

3. Paté-Cornell, M.E.: Learning from the piper alpha accident: a postmortem analysis of technical and organizational factors. Risk Anal. **13**(2), 215–232 (1993)
4. United Stats of America, National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling: Deepwater: the gulf of oil disaster and the future of offshore drilling, report to the President (2011). http://www.gpo.gov/fdsys/pkg/GPOOILCOMMISSION/pdf/GPO-OILCOMMISSION.pdf
5. Christian, M.S., Bradley, J.C., Wallace, J.C., Burke, M.J.: Workplace safety: a meta-analysis of the roles of person and situation factors. J. Appl. Psychol. **94**(5), 1103–1127 (2009)
6. Clarke, S.: Safety leadership: A meta-analytic review of transformational and transactional leadership styles as antecedents of safety behaviours. J. Occup. Organ. Psychol. **86**(1), 22–49 (2013)
7. Zohar, D.: Safety climate: conceptualization, measurement, and improvement. In: Schneider, B., Barbera, K.M. (eds.) The Oxford Handbook of Organizational Climate and Culture, pp. 317–334. Oxford University Press, New York (2014)
8. Flin, R.: "Danger—men at work": management influence on safety. Hum. Factors Ergon. Manuf. Serv. Ind. **13**(4), 261–268 (2003)
9. Fruhen, L.S., Mearns, K.J., Flin, R.H., Kirwan, B.: From the surface to the underlying meaning-an analysis of senior managers' safety culture perceptions. Saf. Sci. **57**, 326–334 (2013)
10. Fruhen, L.S., Mearns, K.J., Flin, R., Kirwan, B.: Skills, knowledge and senior managers' demonstrations of safety commitment. Saf. Sci. **69**, 29–36 (2014)
11. Fruhen, L.S., Mearns, K.J., Flin, R., Kirwan, B.: Safety intelligence: an exploration of senior managers' characteristics. Appl. Ergon. **45**(4), 967–975 (2014)
12. Clarke, S.: Perceptions of organizational safety: implications for the development of safety culture. J. Organ. Behav. **20**(2), 185 (1999)
13. Zohar, D., Luria, G.: A multilevel model of safety climate: cross-level relationships between organization and group-level climates. J. Appl. Psychol. **90**(4), 616–628 (2005)
14. Zohar, D.: Thirty years of safety climate research: reflections and future directions. Accid. Anal. Prev. **42**(5), 1517–1522 (2010)
15. Crowe, E., Higgins, E.T.: Regulatory focus and strategic inclinations: promotion and prevention in decision-making. Organ. Behav. Hum. Decis. Process. **69**(2), 117–132 (1997)
16. Spence, M.: Signaling in retrospect and the informational structure of markets. Am. Econ. Rev. **92**, 434–459 (2002)
17. Spence, M.: Job market signaling. Quart. J. Econ. **87**, 355–374 (1973)
18. Connelly, B.L., Certo, S.T., Ireland, R.D., Reutzel, C.R.: Signaling theory: A review and assessment. J. Manag. **37**(1), 39–67 (2011)
19. Flanagan, J.C.: The critical incident technique. Psychol. Bull. **51**(4), 327–358 (1954)
20. Francis, J.J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M.P., Grimshaw, J.M.: What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol. Health **25**(10), 1229–1245 (2010)
21. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**(2), 77–101 (2006)
22. Leary, M.R., Kowalski, R.M.: Impression management: a literature review and two-component model. Psychol. Bull. **107**(1), 34–47 (1990)

Proposal of Ergonomic Intervention in Horizontal Traffic Signaling

Claudemilson dos Santos(✉) and Galdenoro Botura Jr.

FAAC-UNESP, Bauru, Brazil
claudemilson@fct.unesp.br, galdenoro@gmail.com

Abstract. The proposal aims to convert the two-dimensional drawing of a conventional traffic signal into a three-dimensional simulation called anamorphosis. Preliminary tests were conducted that revealed an indication that perception can be improved. For the tests, two groups of volunteers were used. The first one visualized the traffic signs in a traditional way, while the second visualized the same images, but with the anamorphic signaling, after this initial stage all the participants answered a questionnaire, where it was identified that those of the second group memorized the signaling better. The simulation of three-dimensionality could produce the effect of changing the instinctive behavior of the driver to reduce speed, or simply cause the driver to pay more attention to the front when encountering a 3D object in his way, even for a brief moment and instinctively.

Keywords: Cognitive ergonomics · Horizontal traffic markings · Anamorphosis · Ergonomic design

1 Introduction

1.1 Cognitive Ergonomics and Traffic Signs

According to Vidal and Carvalho[1] cognitive ergonomics focuses on adjusting human skills and limitations to operate machines, perform tasks and interact with the environment, but also observes the use of certain mental faculties, those that allow us to operate, to ratiocinate and to make decisions at work [1]. In the task of driving vehicles, mental effort involves the controls of the vehicle and the external environment. The attention of the driver is disputed between the position / trajectory of the vehicle itself and of others, obstacles, pedestrians, advertising messages, etc. Among so much information, this article highlights horizontal signage. The horizontal traffic signaling is the one that is used of the scroll track like support. It has a privileged position in the field of vision of the driver, since he does not need to look away from the lane for reading. For this reason, it has the function of guiding and regulating the flow of vehicles, issuing alert alerts, among others, as a complement to vertical signaling.

The CONTRAN establishes that the horizontal signaling is a subsystem of the road signaling composed of marks, symbols and legends, placed on the pavement of the

[1] The direct quotations contained in the text were translated into English by the author.

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_105

rolling track. Horizontal signaling is intended to provide information that allows road users to adopt appropriate behaviors in order to increase traffic safety and flow, order traffic flow, channel and guide road users. Horizontal signaling has the property of transmitting messages to drivers and pedestrians, enabling their perception and understanding, without diverting attention from the road bed. Due to its strong communication power, signaling must be recognized and understood by all users, regardless of their origin or frequency of use [2].

However, some studies indicate that horizontal signaling does not satisfactorily perform its function. In a study carried out by Cucci Neto [3] in the city of São Paulo, about the perception of the legend "DEVAGAR" by the drivers, showed that only 4% of the drivers interviewed remembered having seen it seconds after passing through it. This realization led to the elaboration of a series of questions about the reasons that lead to this low perception: would it be the excessive use of subtitles throughout the city? Are the drivers not very perceptive for this type of signaling? Could the use of better defined technical criteria improve the performance of this type of signaling? [3]

One of the possible reasons for the low index of low perception of the legend "DEVAGAR" is the phenomenon of selective vision. When an observer is preoccupied with some specific detail in his field of vision, the other details that do not pose immediate risk or that do not collaborate to establish meaning are despised; although its image is projected onto the retina, it is not recorded by the conscious mind, so it is not recorded by memory. The same does not occur with an obstacle in the lane, because the selective view issues an alert that dispenses a rational analysis of the "threat", provoking the instinctive reaction to reduce the speed. A frequent user of a section can walk it daily and even then, ignore the existence of this signal. But if a new and unknown obstacle is placed, interpreted by the mind as a three-dimensional form, attention will be awakened. This aspect is being used to redesign the legend, conveying the illusion that it is some kind of three-dimensional object. The development proposal is to test the perspective technique called "ANAMORPHOSIS" to improve the usability of the horizontal signaling system, improving perception and memory indices, by provoking the observer with a reaction that attracts attention instinctively.

In the work of driving a vehicle, decisions are also influenced by physical factors related to the vehicle and the environment. All human behavior originates in sensory stimuli. We react to light, to sound, to movement, to heat, finally, to every stimulus perceived by the senses; and with accumulated experience we construct meanings and our conceptions of space and form. Reality is therefore a mental construction processed by visual intelligence. Hoffman [4] argues that visual intelligence relates to other aspects of our intelligence, but it assumes a preponderant role in the process of mental construction.

Hoffman says that the visual intelligence takes up almost half of the cortex of your brain. Normally, it is closely linked to your emotional intelligence and your rational intelligence. It builds the elaborate visual realities in which you live, move, and interact. It transmits these constructions to their emotional and rational intelligences, which use them as raw materials of later constructions. The emotional world you inhabit is, like your visual world, the product of your own constructive genius. (...) If a percentage of cortex is a measure of something, then visual intelligence is a major facet of who we are as a species, and its understanding is a key to what we can become [4].

The construction of reality happens at all times, not just through visual information. The other senses are also recruited, such as the olfactory, sound and balance; to a greater or lesser degree. Reality is not a result of the sum of all variables, but moderated by an equation that involves emotions, memories, and other mental factors not yet fully understood. The manipulation of our visual constructions can be observed in the various media, to which all exposed. Hoffman [4] unravel this strategy as follows:

Consider (…) marketing and advertising, which daily manipulate our acquisitive habits with sophisticated images. Corporations spend millions annually on billboards, packaging, magazine ads, and television commercials. (…) If you intend to sell something, it is certainly essential to understand visual intelligence, for effective visual planning. (…) Perhaps the most surprising conclusion from the research on vision is this: vision is not merely a product of passive perception, it is a clever process of active construction [4].

If instead of selling something through visual manipulation, we are concerned with influencing good behavior, a good application would be traffic signaling, aiming, for example, at reducing speed. According to Thaler and Sustein (apud SCHWARTZMAN [5]), a push in the right direction can bring the desired benefit without the driver bothering to follow a rule. "In" Nudge "they advocate what they call" libertarian paternalism. "The basic idea is that human beings are anything but the rational and well-informed agents imagined by the classical economics textbooks." [5]. In practice, people often make the worst possible decisions about the future. They do so largely because of mental biases that are predictable and quantifiable. What Thaler and Sunstein suggest is that we take advantage of another feature of the real world: the inevitability of making certain choices, so as not to authorize people to "nudge" people in the right direction.

This theory applies to several fields. Self-control problems? In fact, the brain relies on two "systems" to react to the world, one automatic and one reflective. The latter is slow and inefficient. But he is also the only one capable of accounting and planning for the future. It is, for example, that leads us to want to be thin to prevent diseases. The problem is that when we eat, we get carried away by the automatic system. And it only for when the dish (or the pantry) will end [5].

Exploring the concept of Nudges applied to traffic, several alternatives emerge to shape certain behaviors, in this section, the proposal of the present study is focused on the horizontal signaling, exploring the instincts of the automatic system mentioned above. To do so, it is necessary to understand the concept of anamorphosis and its applicability in traffic.

1.2 The Anamorphosis - Definition and Constructive Principles

Anamorphosis is the name given to a type of conical perspective whose projection is made in a plane other than the vertical plane. The vertical plane is commonly used to draw exact perspectives, as it is a plane approximately equivalent to the plane of the retina, that is, they form similar images. However, it is possible to perform a projection in any plane without damaging the image projected on the retina, it is enough that the observer positions itself in the programmed point of view. Figure 1 shows the

difference between the exact perspective and the anamorphosis. The present gray object is projected in the frontal plane (perspective) or not geometric plane (anamorphosis) without the image of the retina changing.

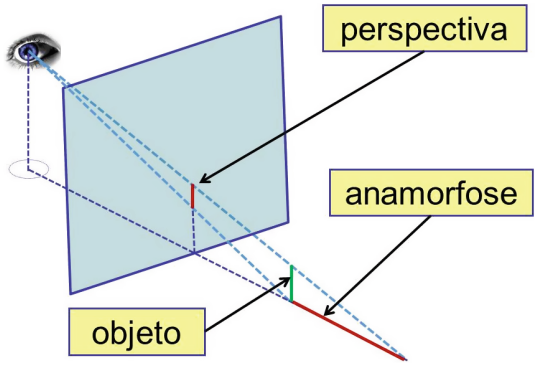


Fig. 1. Difference between the exact conventional perspective and anamorphosis.

If we position an observer standing on a horizontal plane where the signaling is to be inscribed, the inscription must undergo an anamorphic distortion so that it is visualized as a perspective object. However, what is observed in the conventional horizontal signaling, present in the traffic manuals, the distortion is only realized in the height dimension of the legend. Anamorphic distortion assumes that the height and width ratios of the inscription change to be viewed as a perspective.

Figure 2 shows how the anamorphic distortion of a rectangular object projected on the raceway would be.

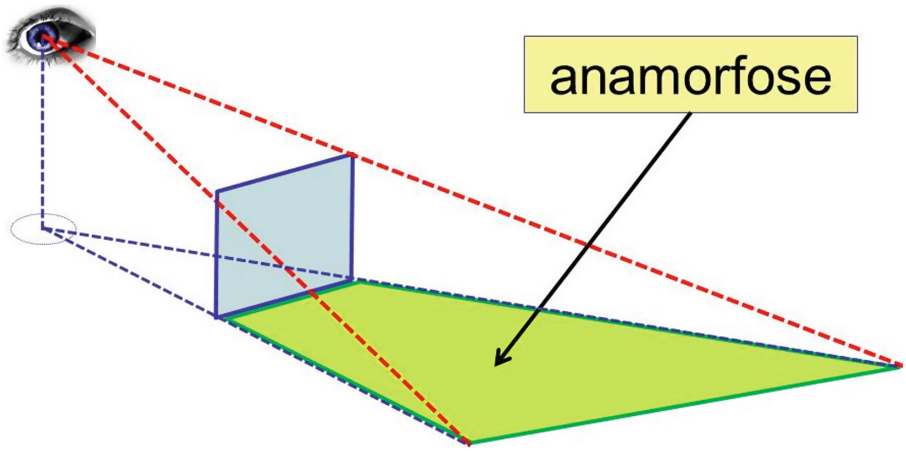


Fig. 2. Anamorphic distortion of a panel.

The inscriptions or drawings contained in a surrounding rectangle must undergo the proportional anamorphic transformation is not to be only visualized but understood by the drivers.

This technique has been known since the Renaissance and has been used by many different types of artists. Recently, advertising has also adopted this technique on panels located on soccer fields next to the goal, positioned to cause the three-dimensional illusion from the TV camera point of view [6]. The following figures (Figs. 3, 4, 5 and 6) demonstrate how the anamorphic method works compared to the traditional method applied to horizontal signaling.

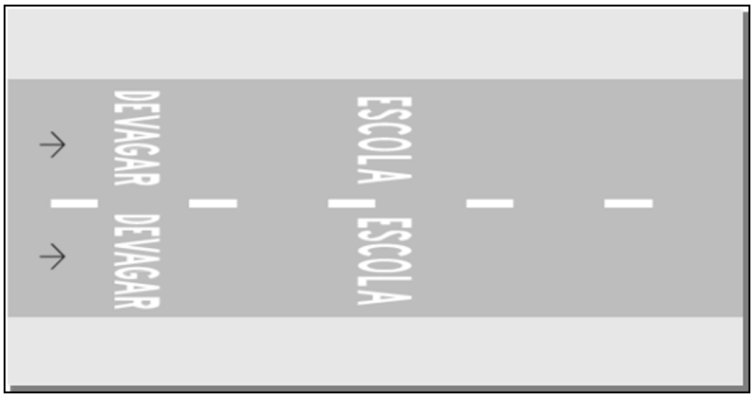


Fig. 3. Legend implementation by the traditional method

Fig. 4. Visualization of the traditional legend by the driver.

Fig. 5. Anamorphic projection of the legend DEVAGAR

In Fig. 3 you can observe the caption projected according to the current regulation. Note that the letters are slightly stretched. The extent to which this distortion occurs varies with the regulatory speed of the track. It does not consider other aspects of vision in perspective, for example, the reduction of measures of width of the image in proportion to the distance of the observer. Figure 4 represents the driver's view. Note that the perspective shapes the superscription, as if it were contained in a trapeze.

Figure 5 below shows how the anamorphic legend in the soil would be applied. The caption suffers deformation which corrects the driver's perspective by increasing the length of the caption at the top. The result from the point of view of the driver is a kind of three-dimensional illusion, in which the legend seems to be inserted in an imaginary vertical plane, as if it were floating in front of the driver. Figure 6 shows how the approximate view of this anamorphosis would be.

The anamorphic signaling design methodology requires the data presented in Fig. 7. The dimension "x" refers to the distance of the observer to the anamorphic panel, it must be dimensioned according to the speed of the range and must increase in proportion to the speed. This data is arbitrary, however, it is recommended that the point of view (PV) be distant enough not to cause sudden deviation maneuver if the driver believes that he is facing a real three-dimensional object. The greater this distance, the three-dimensional illusion will be seen longer with the vehicle in motion.

The "y" dimension shall be determined according to the average height of the walking vehicles. The dimensions "a" and "b" and "c" configure the resulting trapezoid in which the signaling image must be fitted, deforming the original rectangle of dimensions "a"–"h". In a static configuration, simply follow the geometric process above to find the resulting trapezoid that will produce the three-dimensional effect on the desired signaling. The "x" dimension must be large enough for the speed effect to be minimized and the illusion seen longer.

Fig. 6. Visualization of anamorphic signaling

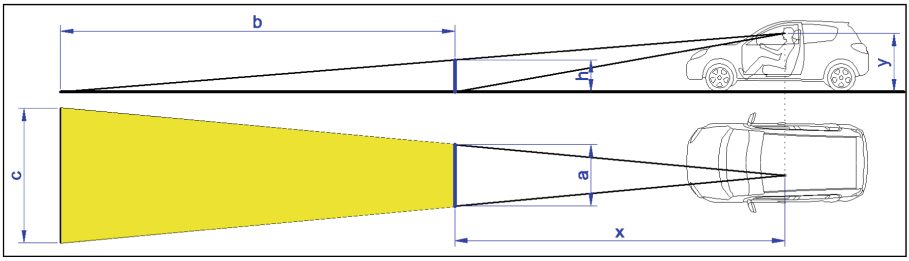


Fig. 7. Sizing of anamorphic signaling

2 Development and Results

Two groups of volunteers (group A and group B) were analyzed for confirmation of preliminary tests. For group A, photographs taken from the point of view of the driver in actual traffic situation were displayed images of the standard horizontal signaling. While group B was exposed to the same set of images, with the only difference that the legend DEVAGAR was inscribed with the anamorphic technique.

The caption analyzed was the inscription DEVAGAR.

Figure 8 shows the legend DEVAGAR, entered in the standard way. Figure 9, on the other hand, shows the anamorphic signaling of the legend DEVAGAR.

After viewing the images, both groups answered the same questionnaire. The questionnaire addressed which signs were present in the images, among them, the legend DEVAGAR. The initial results indicated that only 5% of group A remembered to have seen the traditional legend DEVAGAR. In group B, this index rose to 25% of the participants. The difference found between the readability of the legend slowly in the standard model and in the anamorphic model indicates that anamorphic signaling once implanted may increase traffic safety and decrease the number of accidents.

Fig. 8. Legend DEVAGAR in the pattern signaling.

Fig. 9. Anamorphic legend of the inscription DEVAGAR.

The analysis of the anamorphic signaling implantation project can also be used to test cognitive ergonomic design methods. It is an opportunity to test design methodology concepts, classified as behavioral design. Concepts like those of Thaler and Sustein [7] called "nudge", or in an approximate translation, "pushing in the right direction". Gibson's concept of affordances, Gestalt theory, and so many studies on visual perception and its influence on human behavior can be put to the test in such cognitive ergonomics projects. In Lockton [8] highlights the importance of involving designer and ergonomists in a project team, they claim that:

(…) there are varying degrees of success so that designers' intentions regarding the use of the product are realized. In designing products, designers are also designing user activities, which does not occur independently of the product. "Going further, he notes that" somehow, consumer behavior is shaped by products as far as the products are Shaped by consumer behavior [8].

(…), there is little general guidance available for design team briefing on how to influence user behavior, particularly in contexts where it would result in social benefit: this facet of the design process is not explicitly covered by the variety of tools available For designers of ergonomics (…) [8].

It has been argued … that every project is intended to influence user behavior in the sense that the artifacts that surround us contain socially constructed 'scripts' for users (…) - for example, Chair in a workstation, would be influencing the user to "follow the script" and sit down. However, systems designed intentionally to influence behavior other than what is usually associated with the situation, or in situations where a user does not Would have a strong idea of what to do (e.g. with an unfamiliar interface), represent a higher grade for the intentional designer [8].

Although anamorphic signaling is considered promising, the perfect illusion is only obtained in a single PV, but our visual system consists of binocular vision. It also requires that the PV be positioned at an exact point, and this must be static. In the case of the proposed signaling we have the PV in motion and in variable positions. Although these are unfavorable considerations, it is not yet known if such negative effects are enough to make the proposal unfeasible, so it is recommended that studies addressing PV in motion should be deepened.

3 Conclusion

Preliminary tests indicated that the perception of the legend DEVAGAR can be improved through the anamorphosis facilitating the perception, initial condition necessary to emit response. The simulation of tridimensionality could produce the effect of altering the driver's instinctive behavior by slowing down. Particularities such as slope of the track, its curves and geometry, makes it necessary that each deployment requires a case study, however, there are no theoretical difficulties to generate the anamorphosis, its geometric principles are demonstrated in this article and there are no factors that make it impossible to create an application That generates the distortion for particular situations, using the scheme of Fig. 7. The obstacles that involve the movement signaling as the variation of height, look and speed still need more precise tests, if after these tests it is verified that this Strategy does not create additional risks, it may be proposed to regulate it with the competent bodies.

The next step of the research is to create anamorphic signaling in a real traffic situation. In this way the driver's response to this new signaling model can be analyzed.

References

1. Vidal, M.C., Carvalho, P.V.R.: Ergonomia Cognitiva: raciocínio e decisão no trabalho. Virtual científica, Rio de Janeiro (2008)
2. Contran - Conselho Nacional De Trânsito, Manual Brasileiro de Sinalização de Trânsito: Vol. IV, Sinalização Horizontal, 1° edição. CONTRAN, Brasília (2007)

3. CUCCI Neto, J.: A research on the perception of the legend “DEVAGAR” by the drivers. Technical Notes SP09 / 92, NT 154/92 (1992)
4. Hoffman, D.D.: Inteligência Visual: como criamos, o que vemos. Campus, Rio de Janeiro (2000)
5. Schwartsman, H.: Empurrõezinhos. Coluna. Jornal Folha de São Paulo, 03 Feb 2011. Disponível em http://m.folha.uol.com.br/colunas/helioschwartsman/869841-empurroezinhos.html, último acesso em 30 Aug 2014
6. Sotto, P.O., Santos, C.: Anamorfose na comunicação visual. In: XXI CIC UNESP, 2009, São José do Rio Preto - SP. Congresso de Iniciação Científica da UNESP (2009)
7. Thaler, R.H., Sustein, C.R.: Nudge: o empurrão para a escolha certa. Elsevier, Rio de Janeiro (2009)
8. Lockton, D., Harrison, D., Stanton, N.A.: The design with intent method: a design tool for influencing user behaviour. Appl. Ergon. **41**(3), 382–392 (2010). ISSN 0003-6870, http://dx.doi.org/10.1016/j.apergo.2009.09.001

Road and Rail: Transportation Logistics

Collective Cars in Cities: Optimal Management in Real Time Through DES Approach

Jennie Lioris(✉)

ENPC-ParisTech, Paris, France
jennie.lioris@cermics.enpc.fr

Abstract. A decentralised management of an autonomous demand responsive transport mode is considered, where more than one passengers may be associated to a vehicle. No advanced seat reservation is required by clients appearing in the network in a continuous Poisson stream aiming at an immediate departure. A potential waiting time is associated with customers, in case no vehicle is available by their arrival, as well a tolerated trip detour threshold. Meanwhile, vehicles travel on the network transporting passengers while searching for prospective clients. Two different schemes, employing fully autonomous and manually driven cars are considered. A comparative study is presented, quantifying the performance of each structure. Moreover, the efficacy of each vehicle scheme is compared with a similar one employing a greater number of vehicles in service where each vehicle has an increased passenger capacity. Client-vehicle interaction forms an asynchronous, dynamical, spatio-temporal system the evolution of which is represented through an event-driven approach. Client metrics such as client initial waiting time, abandonment rate, trip detours etc. are measured as well vehicle indicators providing information on the vehicle productivity.

Keywords: Demand responsive transport · Travel demand management · Discrete-event simulation · Fully autonomous cars · Performance evaluation · Parameter optimisation

1 Introduction

Complex demand patterns are influencing city transport and are closely related to congestion. The distribution of the activities in urban areas (suburbs extent) explicitly influence demand. A scattered demand is associated when many activity centres exist while a concentrated type of demand is involved when a major activity centre is associated. However, changes in the spatial distribution of the activities cause changes to trip destinations. While urban transport aims at improving the accessibility of the demand generated within the urban environment, the impact of spatial development and journeys related to passengers and freights results to contradictory effects, where the number of the involved complex trips is increasing. Thus, *complex paths* are associated with downtown areas, occurring mostly in Europe, where shorter trips may also take place and for which slow transport modes can be used (cycling, walking etc.). On the other hand, the urban sprawl observed in the U.S. cities and European suburbs,

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_106

where longer trips are involved, increasing traffic and potentially deteriorating congestion. Moreover, *aggregating mobility* within limited areas in city centres, may also lead to heavy traffic. Most of the journeys can be classified as *daily*, *professional*, *personal*, *tourist*, *freight* etc. trips and hence they are *predictable* or *unpredictable* ones. Travel patterns are also related to the journey assignment (selection and availability of trip paths).

Although public transport is considered as essential element when managing urban traffic, in some cases it is related to a high level of *unattractiveness* turning commuters towards the use of private vehicles. *Poor performance* of public transport, complex service patterns, *lack of network coverage*, *personal convenience and security*, *reliability* are some of the principal elements encouraging the use of individual cars.

Since demand varies within time and space one would wonder how wise is to consider that transport services associated with fixed routes and timetables will efficaciously improve mobility while remaining long-term profitable.

Additionally, establishment of new public transport operations, may be associated with increased traffic along the routes served by the new transportation services.

Furthermore, parking constraints implicitly influence levels of traffic. The *cruising* time spent by cars searching for parking in large metropolitan areas can be quite important. There is a correlation between transport policy and transport planning. The way a transport scheme is planned depends upon the purposes of the policy.

Conventional taxis, providing services of similar quality as private cars, without parking limitations can successfully respond to *spatio-temporal* demand variations but at a cost which in general is felt as being quite expensive for a regular daily use.

A system of urban "collective cars" intelligently utilising vehicle capacity could provide comparable services to conventional taxis or individual vehicles at affordable rates. Indeed, it would offer a door-to-door service without the need of driving in a hostile environment, searching for available parking space, whereas additional advantages in terms of cost, environmental impacts and traffic conditions can be expected by simply raising the occupancy rate of cars. Vehicle itineraries will be optimised according to the current vehicle state, demand and present traffic conditions while detours will be controlled. Moreover, the idle vehicle management will be considered in order to keep car as busy as possible.

A decentralised management is envisaged, where each car will be optimised independently of the others. No prior reservations are required by clients who simple arrive at the network searching for a vehicle to bring them to their destination. Similarly, vehicles move within the network, transporting clients to their destinations while searching for potential new customers when the conditions allow for it. Different strategies can be constructed, based on different approaches for controlling the system and/or its dimensioning, the performance of each one will be evaluated within a discrete event decision tool. Various metrics qualifying and quantifying the service quality are calculated such as the client detours and initial waiting time, travel duration, vehicle activity etc. Hence, two transport schemes are going to be studied, both employing the same number of vehicles which are also of the same passenger capacity, and the same service duration. One structure utilises conventional manually driven cars while the other one employs self-driving vehicles moving on specific road lanes.

The current work is an extension of the comparative study presented in [1] where a similar transport system is considered, involving either manually or self-driving cars, and associated with different resource dimensioning. More precisely, in [1] 3,744 vehicles of five available passenger seats are employed serving a centripetal geometry of demand, generating approximatively 15,400 clients per hour, all over the network. The two structures are compared and the system employing fully autonomous cars presents higher performance, especially for the client measurements such as client waiting times and thus client queue lengths at nodes, abandonment rate, trip detours etc. Despite the fact that the fully autonomous vehicle scheme transports more customers (117,088 versus 112,883 for the manually-driven vehicle structure) the average vehicle occupancy is of 1.47 passengers for the fully autonomous vehicle scheme towards 2.15 for the conventional human-driving cars. On average, the self-driving cars transport only one passenger during the 60% of their service time. This paper examines the exerted influence of the system dimensioning, especially when reducing the number of vehicles in service as well their capacity, on the system performance for the same network and demand pattern.

Thus, the remainder of this paper is organised as follows. Section 2 briefly presents the system model, concerning client and vehicle behaviour and describes the utilised network information. Section 3 introduces the employed data such as demand pattern (geometry and intensity), resources (vehicles in service, related passenger capacity and associated service duration) as well other client characteristics (diversion threshold, max waiting time etc.). Section 4 analyses the system performance for each employed vehicle structure and quantifies the associated scheme efficacy. Finally, Sect. 5 accentuates the key points of the current study while it refers future goals of this work.

2 System Approach - Study Scope

Clients appearing everywhere at the network according to a Poisson stream and wishing an immediate departure, search for a vehicle able to bring them to their destination. If no such car is found by the time a client arrives at a node, potential customers decide whether to wait and for how long, according to their preferences. Customers who are not served by the end of their waiting period quit the system. A diversion threshold is associated with each client representing the tolerated detours regarding the corresponding direct trip, between the client origin and destination node. The employed demand (intensity level and geometry) follows the desired pattern to be studied, representing early hour movements, when the majority of commuters wishes to join the city centre.

Moreover, vehicles travel along the network, bringing passengers to their destinations while searching for potential new clients when the conditions allow it. Each vehicle is characterised by the passenger seat capacity and service duration. A car ceases its service if there is no passenger on board and consequently it has accomplished all its tasks. Idle vehicle management is considered where an appropriate station node has to be defined as well a related maximum parking duration. If no new client is found by the end of that station period the related decision will be revised and potentially a new parking node will be allocated to the vehicle.

The network model is represented as a directed graph the vertices of which correspond to the network nodes and the arcs represent the network links. Events occur at nodes while vehicles travel on links. Random travel times are associated with each network link, following probability distributions corresponding to the actual traffic state. These laws may change within time in order to represent traffic variations. However, for the current study, decisions are based on average values of these metrics while the realised vehicle movements follow the given laws.

One of the goals of this work, is to examine the impact of the resources when facing a given demand and consequently evaluating the resulting system effectiveness and related costs as a function of the demand level (offer curve). In fact, the performance of such an innovative transport mode will be influenced by the related demand and thus trying to predict its effectiveness is a quite complicated task. Moreover, since different transport modes exist, the demand which potentially may be captured by the collective cars structure, relies on the performance characteristics of the other transport alternatives (travel quality, associated costs). Consequently, an optimised system for any demand level is required, in order to establish the system efficacy.

3 Building Data

The "collective cars" transport mode is going to be applied at the Paris region. Due to the impossibility of publishing results utilising real and recent information, a set of fictitious but sufficiently realistic data is going to be employed. It is of worth to emphasise, that the presented evaluation methodology does not depend upon the involved numerical input values.

Figure 1 illustrates the considered network of Paris area (including suburbs), comprised of 288 nodes and 674 links. The employed centripetal demand geometry corresponds to trips occurring early in the morning when most of the involved trips are originated in the outskirts of the town and heading towards the centre of Paris.

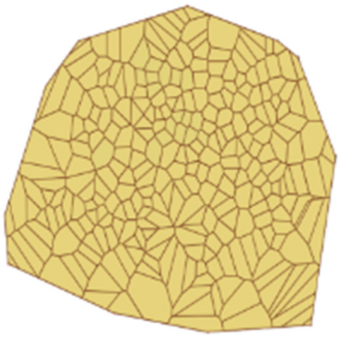


Fig. 1. Network: Île-de-France

The maximal waiting time associated with each potential client is of ten minutes. Each dialogue between a client and a car lasts 30 s while 10 s are allocated to each entering/exiting passenger. The number of vehicles in service is 3,168, each one is of three available passenger seats. The duration of each implementation is of eight hours. This duration is proved to be long enough to provide low variance results under the adopted stochastic model. The max detour threshold tolerated by clients is taken equal to 1.9.

4 Performance Measurement

Two structures of "collective cars" are going to be modelled. A comparative analysis of the effectiveness of each scheme is going to be discussed while also referring the performance of the system employing 3,744 autonomous vehicles of five passenger seats, developed in [1]. All experiments are utilising the same inputs concerning the network topology and demand pattern (geometry and intensity).

4.1 Network Statistics

For each network node, the total number of vehicle visits is measured. The upper plot of Fig. 2 (blue curve) corresponds to the structure employing 3,168 conventional human-driving cars of capacity three seats while the lower one (purple plot) illustrates the corresponding case when self-driving vehicles are considered. Obviously, autonomous cars join network nodes more frequently than conventional vehicles. Similarly, the number of times that vehicles went through each network link is computed.

The green plot in Fig. 3 represents the human-driving vehicle structure while the orange plot is related to the autonomous vehicles. As expected, self-driving cars cross fairly often the network arcs.

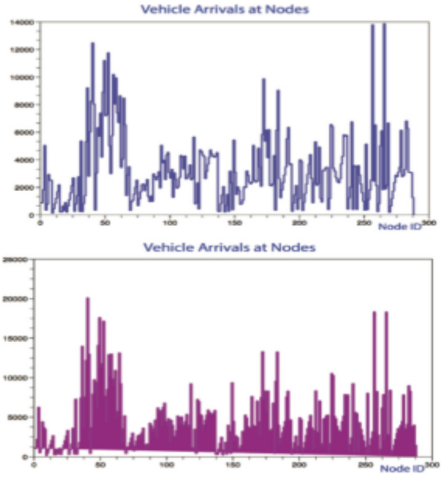


Fig. 2. Vehicle node visits

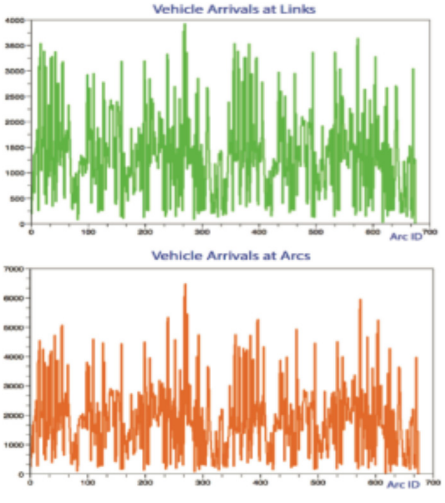


Fig. 3. Vehicle link visits

4.2 Client Waiting Time

Figure 4 represents the distribution of the client waiting time everywhere in the network. The lighter green histogram (upper plot) corresponds to the manually-driven cars while the darker one (lower plot) represents the autonomous vehicles. Again, smaller values are associated with the self-driving cars. Moreover, it would be of interest to examine the average client waiting time per intersection node. Figure 5 depicts the mean client waiting time per network node ± the standard deviation, for the manually driven cars (upper plot) and the autonomous vehicles (lower plot). Thus, on average 231 s of client waiting time are associated with the human-driving cars (red line) where only 104 s correspond to the self-driving ones (brown horizontal line).

Figure 6 presents the average client queue length at each node of the network. The upper plot corresponds to the scheme employing conventional cars and for which the average value of the client queue size for the entire network equals to 3.42 customers. The lower plot depicts the client queue length for the structure utilising self-driving cars. The average queue length all over the network is equal to 1.53 clients.

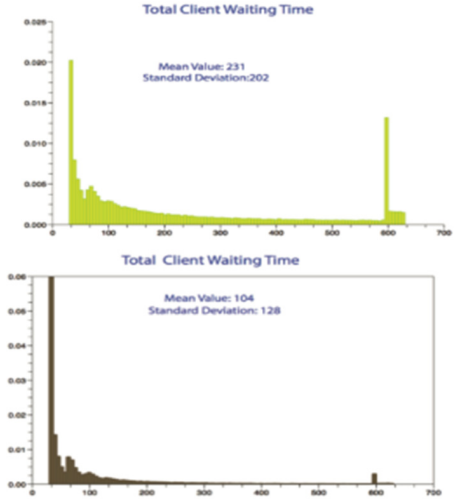


Fig. 4. Distribution of the client waiting time

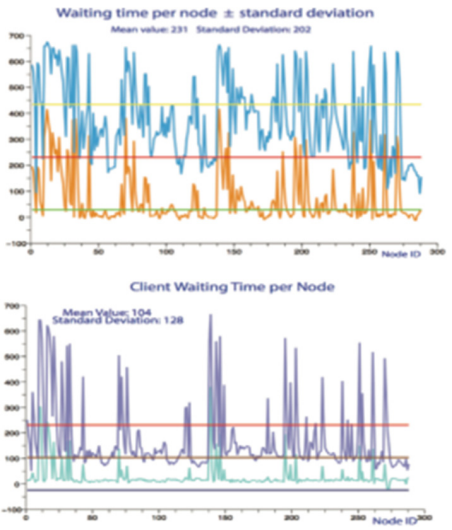


Fig. 5. Average client waiting time per node ± standard deviation

4.3 Client Abandonment Rate

Examination of the client abandonment rate will reveal *critical* nodes. As previously discussed, clients who are not immediately served by a vehicle, wait during 10 min and if until the end of this period no vehicle is found they abandon the system. The average client abandonment rate, for the whole network is 2.04% for the fully autonomous cars (lower plot in Fig. 7). The related value for the manually driven cars is 10.91% (upper plot in Fig. 7). This structure of conventional cars cannot accommodate the demand and consequently increased values of client abandonment are associated.

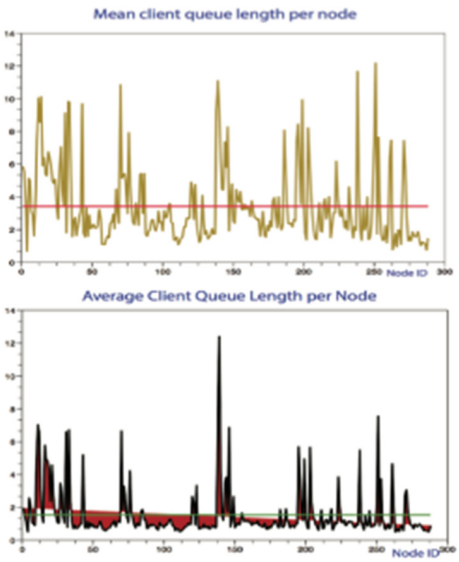


Fig. 6. Average client queue length per node

4.4 Quality of Service

An important indicator characterising the provided *quality of service* is the *total detour ratio*. It is defined as the ratio of the sum of the realised duration of the client trip plus the client initial waiting time (if any) over the duration of the direct client trip (the one the client would have made if he had taken a classical taxi or a private vehicle). Figure 8 illustrates the distribution of the total detour, for the two vehicle structures (red plot: manually driven cars, purple plot: self-driving vehicles). Obviously, the fully autonomous cars manifest a better performance (the average detour value is of 1.57) versus the human-driving cars (average value equal to 1.81), where the diversion thresholds for both structures is taken equal to 1.9.

4.5 Vehicle Occupancy

During the current experiment, the structure employing fully autonomous vehicles transported on total 113,851 passengers to their final destination while the scheme using manually driven cars delivered 100,551 clients. The average number of passengers per fully autonomous car is 1.82 where the mean value of the total number of transported passenger per fully autonomous car is 35.94. Concerning the human driving cars, the average vehicle occupancy equals to 2.47 (greater value regarding the autonomous cars) while the average value of the transported passengers to their destination per vehicle is equal to 31.74. These results are in accordance with the analysis qualifying the service quality. Higher (but still acceptable) diversions are associated with the structure employing conventional cars due to the increased number of vehicle passengers having different destinations.

Figure 9 presents the vehicle occupancy as a function of the service time. The max available passenger capacity of each car employed by both structures is three. Hence, on average each autonomous car (lower plot in Fig. 9), was transporting one passenger during the 38% of the service time, two passengers for the 36% of the service while it was making use of its full capacity during the 24% of the service period. As presented in [1], for the structure employing 3,744 fully autonomous car, each having a passenger capacity of five seats, on average each vehicle was carrying one passenger for the 60% of the service time, two passengers during the 24% of the service time, three passengers for the 8% of the service time, four passengers during the 3% of the service time and five passengers during the 1% of the service time. Each such vehicle was travelling at an empty state for the 5% of its service time. The related client abandonment rate for that structure was equal to 0.25%. As a result, the current structure makes better use of the vehicle capacity and keeps vehicles busier, while providing high quality os service (limited detours) but since it utilises less vehicles in service, the corresponding client abandonment rate is increased.

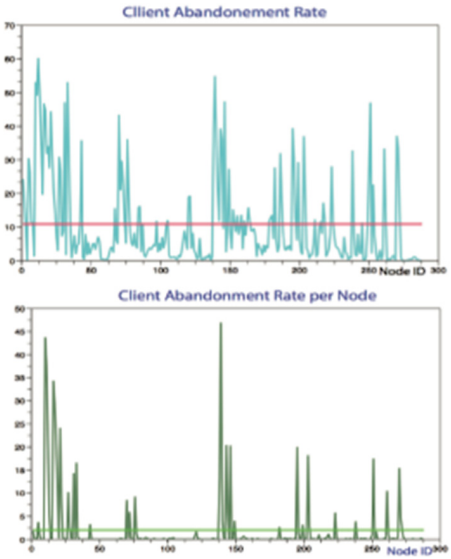


Fig. 7. Client abandonment rate per node

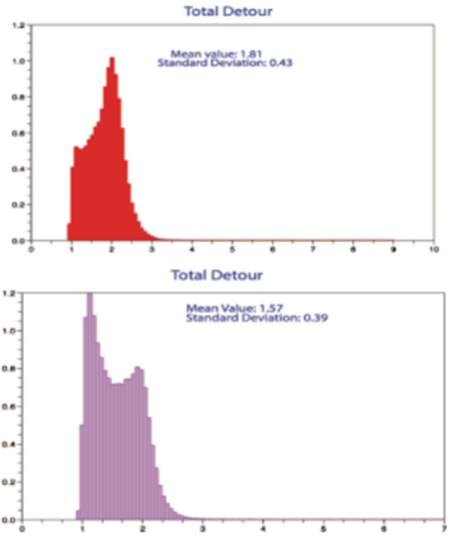


Fig. 8. Total detour

4.6 Vehicle Activity

Figure 10 presents the percentage of time spent by vehicles on travelling, doing various operations at nodes (dialogues entries, exits) and being idle. The upper plot of Fig. 10 corresponds to the scheme employing manually driven vehicles while the lower plot depicts the activity of the fully autonomous cars.

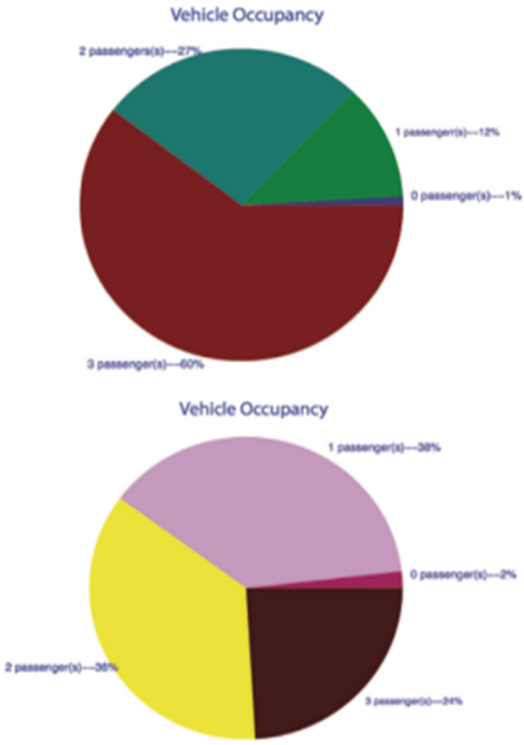


Fig. 9. Vehicle occupancy

The structure utilising self-driving cars is associated with a higher number of served clients and is related to lower values of client abandonment rate and trip detours regarding conventional vehicles. However, the corresponding vehicle idle duration is doubled for the autonomous cars in comparison with the equivalent structure using conventional cars. Despite that, the travel time and the period allocated to various operations is increased for the fully autonomous vehicles. The time spent by self-driving cars on travelling (88%) and being idle (4%) are improved, regarding the scheme employing 3,744 autonomous cars, each one of five available passenger seats, where the corresponding values were 87% and 8% respectively. However, vehicles spent less time on doing various operations at nodes, for the structure of the 3,744 autonomous cars. This is a reasonable result since the vehicle occupancy of the current structure (one employing 3,168 cars) is increased and consequently more entries and exits occur (operations consuming time).

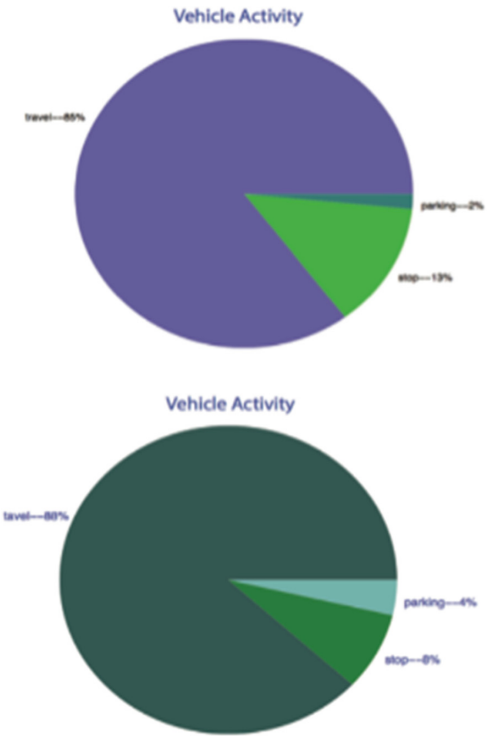


Fig. 10. Vehicle activity

5 Discussion-Future Work

Need to access numerous activities in more complex and sprawled urban environment implies an increase in demand for transport and potentially leading to congestion. Limited public transport infrastructure is not always utilised within an optimal way. Moreover, poor services, complex provided travel patterns (connections, decreased confort, frequent stops, delays etc.), reliability are some of the elements turning public transport unpopular and consequently automobile dependency increases. Furthermore, continuous spatio-temporal changes in demand patterns, which cannot be satisfied by the fixed lines and timetables of the public transport services often leads to heavy traffic.

Taking into consideration all these elements, a “collective cars” transport mode, utilising available vehicle capacity while associated with optimised itineraries, could form a successive alternative to private vehicles. Two schemes of “collective cars” are considered both employing the same number of vehicles, while all cars are of the same passenger capacity. However, one structure employs manually driven cars while the other is using fully autonomous ones which circulate on specifically dedicated lanes and thus respecting the current regulation for road safety. Among the purposes of this work is to examine the impact of the system dimensioning in the system effectiveness.

The performance of each structure is quantified and thus compared with each other. Hence, both vehicle structures (self-driving, manually driven cars) considered in this work, in comparison with the ones of a previous study utilising a greater number of vehicles in service while each car is of higher passenger capacity, involve slightly higher detour values but still acceptable ones by clients and increased values of abandonment rate. Despite that fact, the vehicle activity, concerning the travel and idle durations, is now improved. Additionally, at present the vehicle capacity is better utilised and consequently reduced costs may be associated for clients and cars. Once more, the fully autonomous vehicle scheme presents higher performance on almost all statistical metrics regarding the conventional car structure.

Future works aims at the study of optimal tuning of pertinent parameter values or reasonable trade offs, with the purpose to achieve an optimised system functioning according to desired system configurations, for the scheme of the self-driving cars. Hence, series of simulations need to be performed for different values of each parameter, the analysis of which will evaluate the system efficacy. Additionally, important information can be transmitted when the fully autonomous cars are also connected, regarding traffic state and potentially demand. This will positively impact the construction of the vehicle itinerary, the management of idle cars and potentially allow clients to define appropriately their waiting time at nodes.

Acknowledgments. The author is indebted to Professor Guy Cohen for expert advise, guidance, support and all interesting and mostly stimulating discussions concerning this work.

References

1. Lioris, J.: Performance appraisal of a collective cars system employing fully autonomous vehicles. J. Intell. Transp. Syst. (2017)
2. Baccelli, F., Cohen, G., Gaujal, B.: Recursive equations and basic properties of timed petri nets. J. Discrete Event Dyn. Syst. **2**, 415–439 (1992)
3. Sun, J., Huang, Z.: Autonomous dial-a-ride transit introductory overview. J. Softw. **4**(7), 766–776 (2009)
4. Ciari, F., Balmer, M., Axhausen, K.: Large scale use of collective taxis: a multi-agent approach. In: 12th International Conference on Travel Behaviour Research, Jaipur, India (2009)
5. Diana, M., Dessouky, M.: A new regret insertion heuristic for solving large-scale dial-a-ride problems with time windows. Transp. Res. Part B **38**(6), 539–557 (2004)
6. Dial, R.: Autonomous dial-a-ride transit introductory overview. Transp. Res. Part C **5**, 261–275 (1995)
7. Fu, L.: A simulation model for evaluating advanced dial-a-ride paratransit systems. Transp. Res. Part A **36**, 291–307 (2002)
8. Balemi, S., Kozak, P., Smedinga, R.: Discrete Event Systems: Modeling and Control. Birkhauser, Basel (1993)
9. Olsder, G.J.: On structural properties of min-max systems. In: 11th International Conference on Analysis and Optimization of Systems, pp. 237–246. Springer (1994)
10. Cieslak, R., Desclaux, C., Fawaz, A., Varaiya, P.: Supervisory control of discrete-event processes with partial observations. IEEE Trans. Autom. Control **33**(3), 249–260 (1988)

11. Baccelli, F., Cohen, G., Olsder, J., Quadrat, J.P.: Synchronization and Linearity – An Algebra for Discrete Event Systems. Wiley, New York (1992)
12. Ho, Y.C., Cao, X.: Perturbation analysis and optimization of queueing networks. J. Optim. Theory Appl. **19**, 559–582 (1991)
13. Wu, Y.: Agent Behaviour in Peer-to-Peer Shared Ride Systems. The University of Melbourne, Parkville (2007)
14. Mes, M., Der Heijden, M.V., Van Hartent, A.: Comparison of agent-based scheduling to look-ahead heuristics for real-time transportation problems. J. Eur. Oper. Res. **181**(1), 59–75 (2007)
15. Fountoulakis, M., Bekiaris-Liberis, N., Roncoli, C., Papamichail, I., Papageorgiou, M.: Highway traffic state estimation with mixed connected and conventional vehicles: microscopic simulation-based testing. Transp. Res. Part C (2017)
16. Roncoli, C., Bekiaris-Liberis, N., Papageorgiou, M.: Use of speed measurements for highway traffic state estimation: case studies on NGSIM data and highway A20, Netherlands. Transp. Res. Rec. **2559**, 90–100 (2016)
17. Bekiaris-Liberis, N., Roncoli, C., Papageorgiou, M.: High-way traffic state estimation with mixed connected and conventional vehicles. IEEE Trans. Intell. Transp. Syst. **17**, 3484–3497 (2016)

Human Factors in the Design of Automated Transport Logistics

Matthias Neubauer(✉) and Oliver Schauer

Logistikum – University of Applied Sciences Upper Austria,
Wehrgrabengasse 1-3, 4400 Steyr, Austria
{matthias.neubauer,oliver.schauer}@fh-steyr.at

Abstract. Initial ideas related to automated driving date back to the 1950's. Recent developments and manufacturer's forecasts promise the application of fully automated driving in the near future. What are expectations of logistics stakeholders and how will automated driving change transport logistics? This paper reports initial findings from probing a test region for automated and connected driving related to transport logistics. Thereby, application scenarios are detailed in a stakeholder-oriented manner. Furthermore, the empirical findings are examined in terms of technology acceptance factors observed.

Keywords: Automated transport systems · Technology acceptance · Human factors

1 Introduction

The advent of automated driving has gained significant importance within the logistics domain influencing economic-, environmental- and social sustainability aspects of transportation. Prominent logistic examples like the European truck platooning challenge or the automated Container Terminal Altenwerder in Hamburg, Germany, provide evidence for the applicability of automated driving. Thereby, truck platooning aims at improvements like increased road safety, reduced fuel consumption, increased driver comfort, or increased attractiveness of the truck driver job profile. Container terminal automation targets for example at improving the usage of stacking ground, optimizing the handling of cargo, or reducing labor costs. Today, such automation endeavors should not be regarded as isolated. It is necessary to consider them as connected and interoperable when designing sustainable supply chains across diverse stakeholders and new business models within the logistics domain. In this context, stakeholder involvement and participation is an important ingredient for the successful design and implementation of new solutions.

Aside rather technical discussions related to automated driving and its applications, automated driving has gained public attention and triggered important discussions related to acceptance, safety, legal issues, insurance issues, or even ethical issues. Also governments and (inter-)national funding agencies have adopted automated driving as a key enabler for designing the mobility of the future. This paper reports on findings observed within an Austrian project. The project aims at probing a test region for automated and connected driving in the central area of Upper Austria. The test region

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_107

shall build upon requirements derived from regional industry and from regional infrastructure operators. Within the probing phase, a stakeholder-oriented elicitation approach is taken to gain findings related to the applicability and acceptance of automated driving in Upper Austria. In the initial project phase, the following potential application cases have been identified: (1) Automation of in- and outbound logistics; (2) Automation of multimodal transport and handling of cargo; (3) Automatic maneuvering of special vehicles at company premises, ports, or airports; (4) Automation related to the last-mile and municipal scenarios. This paper presents empirical findings observed while detailing the given application cases in a stakeholder-oriented manner. Furthermore, the empirical results are examined in terms of technology acceptance factors related to the automated driving application scenarios.

The paper is structured as follows. Section 2 presents related work in the field of technology acceptance models and learnings from human factors engineering when automating transportation means. Section 3 describes the elicitation approach regarding the transportation logistics scenarios and the related findings. Section 4 synthesizes the findings of Sects. 2 and 3. The remainder of the paper draws conclusions and discusses future work.

2 Related Work

2.1 Technology Acceptance Models

Research on technology acceptance aims to develop and apply instruments to assess and predict the acceptance of existing and upcoming technologies. A prominent model, commonly referred to, is the Technology Acceptance Model (TAM) originally proposed by [1]. TAM argues that the 'Attitude Toward Using' a technology depends on the variables (i) 'Perceived Usefulness' and (ii) 'Perceived Ease of Use'. Furthermore, TAM states that the 'Behavioral Intention to Use' a system is influenced by both, the 'Attitude Toward Using' and the 'Perceived Usefulness'. Finally, TAM reasons that the 'Behavioral Intention to Use' a system influences the 'Actual System Use' (cf. Fig. 1).

Based on the original TAM, [2] proposed TAM2 - an extension of the TAM. TAM2 additionally includes constructs related to social influence (i.e. subjective norm, voluntariness, and image) and cognitive instrumental processes (i.e. job relevance, output quality, results demonstrability) [2]. The relationships among these constructs are depicted in Fig. 1.

A further TAM extension has been proposed by [3] in their 'Unified Theory of Acceptance and Use of Technology' (UTAUT). UTAUT builds upon TAM/TAM2 and synthesizes eight individual user acceptance models. UTAUT argues that four direct determinants influence the 'Behavioural Intention' and the 'Usage Behavior', i.e. 'Performance Expectancy', 'Effort Expectancy', 'Social Influence' and 'Facilitating Conditions'. The relationships among them are depicted in Fig. 2 as well as moderators (gender, age, voluntariness, and experience) influencing these relationships.

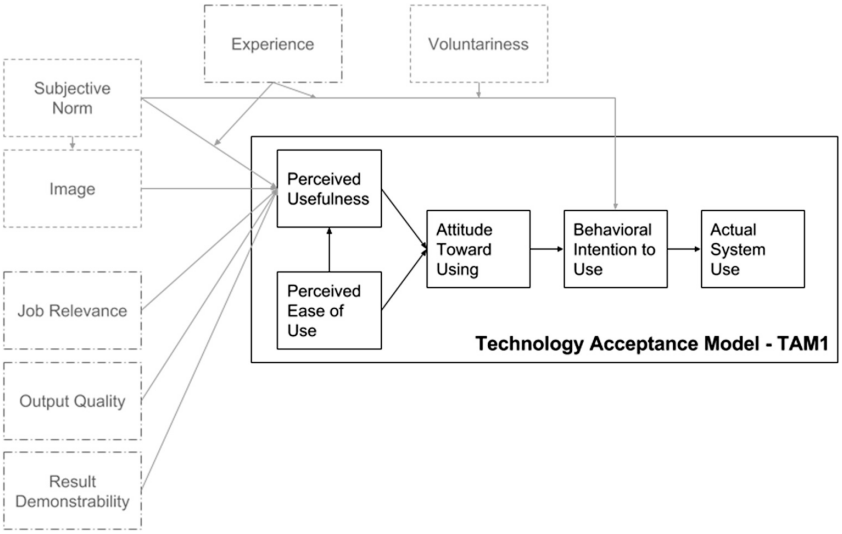


Fig. 1. TAM 2, adapted from [2]

In 2008, [4] present TAM3. TAM3 additionally defines influencing factors on 'Perceived Ease of Use'. This extended model argues that individuals anchor their perception to different general technology-related beliefs, which are adjusted while using a system and gaining hands-on experience. As anchors the model defines the determinants 'Computer Self-Efficacy', Perceptions of External Control', 'Computer Anxiety' and 'Computer Playfulness'. Determinants that are adjusted while using a system represent 'Perceived Enjoyment' and 'Objective Usability' (cf. [4]).

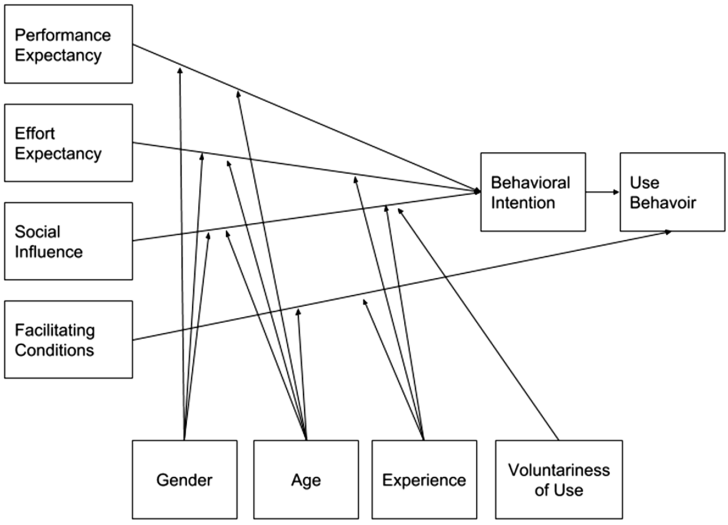


Fig. 2. UTAUT, adapted from [3]

The presented technology acceptance models originate from the field of information systems science. Thus, the models are widely adopted related to information systems evaluation. In the automotive context, [5] propose a theoretical car technology acceptance model (CTAM), which builds upon the UTAUT model and extends the model with the determinants 'Anxiety', 'Perceived Safety', 'Self Efficacy', and 'Attitude Towards Using Technology'.

[5] define the 'Attitude Towards Using Technology' as '*an individual's overall affective reaction upon using a system.*' As such, this determinant reflects '*the beliefs of the user regarding system usage and its effects. It addresses attributes of in-car information systems that lie beyond pure functionality and tries to determine the attitude towards using the system through affective items like fun and likeability*' [5].

[5] consider the determinant self-efficacy as '*a person's belief in his/her ability and competence to use a technology to accomplish a particular task. It can play a major role in how one approaches a goal or a task as it is developed from external experiences and self-perception. In the car context, people with high self-efficacy - those who believe they can perform well - are more likely to view a difficult task as something to be mastered rather that something to be avoided.*'

Anxiety is defined by [5] in the car context as *'the degree to which a person responds to a situation with apprehension, uneasiness or feelings of arousal. Anxiety was derived from a computer anxiety construct and addresses the fear to use a computer because of a general technology inexperience'.*

Finally, [5] define perceived safety as *'the degree to which an individual believes that using a system will affect his or her well-being'* (Fig. 3).

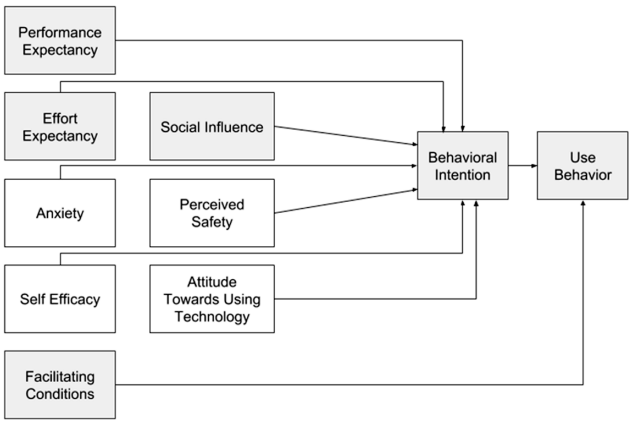


Fig. 3. Car Technology Acceptance Research Model (the grey determinants are inherited from the UTAUT), adapted from [5]

More recent work related to technology acceptance in the context of automated transport systems is presented by [6] and [7]. [6] explore how user acceptance variables might influence the use of a public automated road transport systems (ARTS). The results of [6] indicate that '*all three UTAUT constructs impact on intention to use*

ARTS. Performance expectancy is the strongest predictor, suggesting that the most important factor that people will consider in deciding whether or not to use an ARTS is how well they believe it will perform in comparison to other public transport systems. Social Influence and Effort Expectancy also had an impact on behavioral intentions, indicating that the influence of other people, and perceptions of how difficult the system is to use will also both influence the decision to use an ARTS.' Furthermore, they argue that the UTAUT framework may be applied to increase the understanding of user's behavioral intentions regarding automated vehicles.

In addition, [7] investigates the acceptance of self-driving cars and discusses idealized versus realistic portrays of self-driving cars and their impact. Thereby, [7] introduces the Self-driving Car Acceptance Scale (SCAS) featuring 24 statements that aim to assess individuals' acceptance of self-driving cars. The SCAS instrument is inspired by existing technology acceptance models and theories related to acceptance of automation. SCAS comprises the following dimensions:

- perceived reliability of automation/trust;
- cost of automation;
- appropriateness of automation/compatibility;
- enjoyment of to-be-automated task;
- perceived usefulness of automation;
- perceived ease of use of automation;
- experience with automation; and
- intention to use automation.

The results of [7] indicate that '*people may be more accepting of self-driving cars under idealized rather than realistic scenarios during the initial deployment of the technology*'. Furthermore, the results reveal that future work regarding yet-to-be determined factors is required, which accounts the variance in people's a priori acceptance of self-driving cars. For example, the data suggested '*a tendency for less acceptance among older, more experienced drivers. Familiarity with self-driving car technology and exposure to articles about self-driving cars were both positively related to acceptance*' [7]. Finally, [7] stresses results from human factors research, which show that '*unrealistic expectations about automation performance indeed increases initial acceptance, but trust and acceptance may be irreparably harmed after interactions with imperfect automation reveal its previously undisclosed shortcomings to users.*'

2.2 Learnings from Human Factors Engineering

[8] discuss the evolving role of automation in transportation and present human factors lessons learned from different modes of transportation – aviation, road, rail – when automating functions of transportation means. In their discussion, [8] identify the following (non-exhaustive) cross cutting themes to be addressed when automating transportation means (cited from [8]):

- ***System safety, security, alerting and readiness***
  There is a need to address safety and security concerns related to automated systems. Pilots and controllers (cross-modally, operators) must mitigate risk on a regular and ongoing basis. Automation (and technology in general) cannot be implemented before its fairly mature, in terms of functional safety and reasonable cost-benefit ratios. With cooperative adaptive cruise control, auto-braking and an auditory alarm are necessary to reduce crashes.
- ***Operator-system interaction and transfer of control***
  Vulnerabilities exist in operators' interactions with automated systems. There is a need to mitigate failures in human-machine interaction, especially with operator understanding and expectations of the system. The distinction between system and operator responsibility may not be clear, especially with partially-automated vehicles. In particular, it is unclear what role the driver should play when an automated system fails. We must consider operator-vehicle interface design issues, especially with L2 and L3 vehicles. The most effective hand-off driving strategies incorporate nonvisual components.
- ***Operator monitoring, engagement and workload***
  Issues still exist related to driver status monitoring and driver engagement. Drivers are more likely to engage in secondary tasks when automation affords them the opportunity. Drivers operating vehicles with lane-keeping automation were found to be less engaged, or even completely disengaged, from the driving task. We still need to better understand driver operation of L1 vehicles, since these systems are now more widely available in the light vehicle market. Automated systems can reduce workload during normal operations, but may add complexity and workload during demanding situations. A systems-theoretic computational model can represent operator/team behaviors and workload in the presence of emerging and changing technology.
- ***Operator education and training***
  Lack of practice can result in degradation of basic knowledge and skills. Education is a necessary component of automated systems, even for systems that only assist a driver. There is a need to address issues related to unintended use and consequences. While collision warning and autonomous emergency braking systems are not designed to replace any aspect of driving, overreliance and misinterpretation may still be a concern. Licensing and training requirements must be developed.

3 Eliciting Relevant Automated Transport Logistics Scenarios

The previous section considers related work in the field of technology acceptance. The related work informs the design and evaluation of automated systems. This section presents findings from empirical work conducted within an Austrian project aiming at probing a test region for automated and connected driving in the central area of Upper Austria. This project primarily investigates potential needs and benefits of automated and connected driving within the logistics domain. The empirical results comprise

findings gained within a world café setting and semi-structured interviews. Subsequently, the method applications and the corresponding results are summarized.

3.1 Exploring the Field in a World Café Setting

In an initial step, potential stakeholders of a test region for automated and connected driving were invited to a workshop held in December 2016. Overall, 47 representatives participated at the workshop. The representatives were carefully selected in order to ensure a broad spectrum of different stakeholders. The participants comprised decision makers from technology providers (hard- and software), automotive industry, road and motorway operators, automotive suppliers, logistic providers, and public servants responsible for regional road/city infrastructures.

The workshop design applied the world café method [9]. The set-up comprised four tables, each exclusively dedicated to one of the following topics: (i) Automation of in- and outbound logistics; (ii) Automation of multimodal transport and handling of cargo; (iii) Automatic maneuvering of special vehicles at company premises, ports, or airports; (iv) Automation related to the last-mile and municipal scenarios.

These four cases origin from the research project proposal. The proposal is based on related work targeting the automation of cargo handling and transport, as well as on coarse grain insights gained in an initial stakeholder dialog during the project proposal phase. Overall, the world café took 2 h. The first round took 1 h, the remaining 3 rounds each 20 min. The following guiding questions guided the table discussions:

1. What are potential application cases or scenarios relevant for my organization?
2. What are concrete challenges and needs related to factors such as technology, infrastructure, and general conditions regarding economic, social and legal issues?
3. Which benefit(s) do I expect for my organization? What is important and what is necessary for my organization?

The results gained at each table were documented on flip charts, and table hosts summarized these results after each table round. Finally, the table hosts presented the overall results in a plenary session. Subsequently, the results are summarized.

Table #1 – Automation of In- and Outbound Logistics. Two types of relevant application scenarios have been identified at this table – (i) short range in-bound logistics scenarios, and (ii) long distance freight transport between distribution centers. Type 1 refers to transport routes of around 1 km. These routes typically comprise operating premises and public roads. The transport of goods needs to be carried out between (1) distribution centers of Logistics service providers and adjacent industrial enterprises, or (2) between nearby operating sites of a manufacturer. In the discussion, general requirements were identified such as: 24 h/7 days applicable; Low noise emission; Automated refueling; Automated loading and unloading; Ensuring public acceptance in the region; Ensuring load safety and failure safety; Or certification & testing of automated solutions.

Type 2 concerns application examples for truck platooning. For logistics service providers and vehicle manufacturers, the testing of freight transport, for example,

between distribution centers, which takes into account different road types and the transition between these (motorway → federal road → company premises), is of special interest. As potential benefits particularly the management of the driver's shortage due to the reduction of the necessary driver personnel, the reduction of personnel costs, the increase in transport capacities or the reduction of fuel costs were stated.

Table #2 – Automation of Multimodal Transport and Handling of Cargo. At this table, participants discussed the automation of multimodal transport and transshipment in the context of port areas (ship-rail-truck) and cities (rail-bus-truck-passenger car aircraft). The following potential applications were identified: (i) switched shuttle services, e.g. between distributors and port areas, (ii) autonomous loading and unloading between different modes of transport, and (iii) autonomous distribution networks. In this context, a variety of challenges was discussed, ranging from legal aspects, technical interoperability (V2V, V2I), development approaches for testing autonomous vehicles, end-to-end supply chain considerations, and safety aspects. The identified opportunities and benefits of automation include increasing security and efficiency, developing new business models and reducing costs.

Table #3 – Automatic Maneuvering of Special Purpose Vehicles at Company Premises, Ports, or Airports. Two types of relevant applications have been identified at this table – (i) Autonomous in-house transports on the site (swap body, trailer, etc.), (ii) Delivery scenarios considering different 'Automated Guided Vehicle (AGV) maturity'.

Type 1 relates, for example, to internal transports such as the movement of swap bodies with special trucks or the autonomous rigging of trailers at a site. Discussions included 'support processes' such as the automated refueling/charging process (fuel cell, battery replacement by robots) or the handling of technical failures and order peaks by means of 'parallel infrastructure' of conventional vehicles and teams.

Type 2 deals with the topic of the takeover of trailers or entire semi-trailers at interfaces of road sections/operating areas with different 'maturity' of the 'AGV-Readiness'. Furthermore, participants raised important topics such as the certification or the approval of AGV maturity of road sections. In addition, the digital infrastructure (geographic information system data) and vehicle-related data (vehicle geometry, steering behavior) were discussed. For example, a pilot or the truck driver could take over an autonomous vehicle at the factory gate (analogous to shipping) or a special vehicle takes over the trailer (keyword 'trailer park').

Table #4 – Automation Related to the Last-Mile and Municipal Scenarios. In the discussion on this topic, two key fields of application crystallized as promising – (i) Versatile, easy-to-use light commercial vehicles for the support of municipal cleaning and maintenance activities, (ii) special purpose automated vehicles for the delivery in urban areas.

Type 1 includes the use of automated special vehicles for road or sidewalk cleaning, as well as snow clearing or mowing. Vehicles used to carry out this work have (i) a slow basic speed as well as relatively easy to plan routes, (ii) Core working time, (iv) operate at times with little traffic (nighttime), which increases the feasibility in the

urban environment. Urban operators can act more flexibly by increasing capacities at short notice (e.g. use for snow clearing), taking into account occupational safety (e.g. in night operation). Furthermore, the intelligent strategies for route planning and fleet management should be included.

Type 2 refers to delivery in cities. Small and agile driverless vehicles could move more easily through pedestrian zones. In this case, the flexible applicability was emphasized, which enables the usage outside typical working hours.

3.2 Detailing Scenarios in Semi-Structured Interviews

In addition to the world café, semi-structured interviews were conducted in February 2017. These interviews were intended to detail application scenarios and the interest of organizations related to a test region for automated and connected driving. At the time of writing, overall 26 interviews were conducted and documented. The interviewees are high-level representatives from 15 technology providers for automated solutions, from 10 scenario providers and from 1 insurance company.

The interview guide related to the semi-structured interviews comprises, amongst others, questions related to (i) relevant application scenarios, (ii) the strategic role of automated driving for the organization, (iii) the expected usefulness, and (iv) expected technological requirements towards automated driving. The interview results confirm the relevance of the initial world café scenarios and detail procedures and current technologies in use. Furthermore, they provide deeper insights into the role of automated driving for the different stakeholder groups.

Overall, expectations of scenario providers regarding enabling technologies for automated driving seem to be very optimistic. 5 scenario providers explicitly stated that they consider the technology as mature, e.g. for truck platooning. However, they also remarked that the success of automated driving applications will depend on resolving legal aspects, building public awareness related to certain applications, and to clearly describe the potential as well as functionality of automated vehicles within the design of efficient logistics procedures. Actually, all scenario providers explicitly stressed the importance of the return on investment as well as the efficiency and sustainability improvement.

Compared to the scenario providers, the statements of the technology providers draw a rather realistic or even cautious picture regarding today's abilities of automated driving. Overall, the 15 technology providers appreciate a test region, since it allows them to implement and improve their technologies within relevant logistics applications. Furthermore, a test region supports them by adequately evaluating their solutions with respect to technology acceptance criteria such as safety, usefulness or performance expectancy. In addition, they consider the knowledge exchange between different technology providers as vital. They also expect a test region to be beneficial as communication platform between diverse, interdisciplinary stakeholders for drawing realistic expectations towards automated driving applications.

4 Examining the Role of Technology Acceptance Within the Gained Transportation Logistics Scenarios

In this section, the gained empirical results are examined in terms of technology acceptance aspects. For this reason, the results were clustered according to the technology acceptance dimensions of the related work described in Sect. 2. Following the main findings will be described.

4.1 Technology Acceptance Aspects Discussed Within the World Café

Perceived Usefulness. Related to the usefulness of automated vehicle applications, application scenario providers considered solutions as useful that (i) enable 24/7 operation, (ii) ensure low noise and CO2 emissions, (iii) support automated refueling, (iv) support automated loading and unloading, (v) reduce (personnel) costs, (vi) increase transport capacities, (vii) reduce fuel costs, (viii) increase security and efficiency, or (ix) support to flexibly increase capacities at short notice.

Job Relevance. The participants agreed that automated driving impacts job profiles and needs to be tackled adequately. Thereby, the definition of new job profiles, increasing qualification and job enrichment were stated as important aspects.

Social Dimension. Participants stated that ensuring the public acceptance when testing and implementing automated vehicle solutions is vital. They proposed intense communication with stakeholders when testing and implementing certain solutions. Intense communication additionally was considered essential to reduce fears and increase awareness. Thereby, also aspects such as changes in job profiles and impact on unemployment rates (fear of losing one's job) were stated.

Perceived Safety. Ensuring safety is a major requirement when automating transport operations. Automated solutions must not harm people and need to ensure e.g. load safety or failure safety. Related to safety, the application scenario providers considered the reliability of solutions important. In this context, also ensuring security against hacker attacks was identified an important aspect. Furthermore, adequate human-machine-interaction and dedicated trainings were considered relevant in order to ensure safety.

4.2 Technology Acceptance Aspects Observed in the Interviews

In addition to the technology acceptance aspects observed within the world café results, the semi-structured interviews especially revealed different expectations of scenario and technology providers towards automated driving. This finding links to Sect. 2, in which idealized versus realistic portrays of self-driving cars were identified as relevant in the context of technology acceptance. Unrealistic expectations may harm technology acceptance on the long run. Therefore, the authors consider clear and intense

communication towards potential stakeholders as beneficial when designing automated transport logistics solutions.

Two technology providers explicitly addressed the role of human computer interaction and rigor technology evaluation. Furthermore, another one stressed the appropriateness of automation multiple times in the interview. In addition to the importance of safety, one technology provider explicitly stated the importance of ethics as determining factor for technology design.

5 Conclusion

Automated driving has gained significant importance within the logistics domain. The design of automated driving applications needs to consider technology acceptance aspects to ensure successful implementations and widespread adoption of such solutions. This paper presented initial findings gained while probing a test region for automated and connected driving in Upper Austria. These findings were linked to related work in the field of technology acceptance. Thereby, stakeholders identified in a world café setting aspects related to perceived usefulness, perceived safety, social dimension, and job relevance as relevant within logistics application scenarios. In addition, the importance of managing different expectations between diverse stakeholders, i.e. realistic un-realistic expectations towards automated transport logistics applications, was observed as relevant in the design of automated transport logistics solutions.

Future work may consider the implementation, rigor evaluation and improvement of automated transport logistics scenarios. Findings gained within these efforts should be documented in terms of design guidelines and success stories. In this way, the communication of results related to automated transport logistics could be disseminated among relevant stakeholders.

Acknowledgments. This research was partially supported by the research field 'Smart Hyperconnected Logistics Systems' funded by the State of Upper Austrian as part of the research program 'FTI Struktur Land Oberösterreich'.

References

1. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. **13**(3), 319–340 (1989)
2. Venkatesh, V., Davis, F.D.: A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag. Sci. **46**(2), 186–204 (2000)
3. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view. MIS Q. **27**(3), 425–478 (2003)
4. Venkatesh, V., Bala, H.: Technology acceptance model 3 and a research agenda on interventions. Decis. Sci. **39**(2), 273–315 (2008)

5. Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., Tscheligi, M.: Predicting information technology usage in the car: towards a car technology acceptance model. In: Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2012, pp. 51–58. ACM, New York, NY, USA (2012). doi:10.1145/2390256.2390264
6. Madigan, R., Louw, T., Dziennus, M., Graindorge, T., Ortega, E., Graindorge, M., Merat, N.: Acceptance of automated road transport systems (ARTS): an adaptation of the 5UTAUT6 model. Transp. Res. Proc. **14**, 2217–2226 (2016)
7. Michael A.N.: Acceptance of self-driving cars. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting 60, vol. 1, pp. 1449–1453 (2016)
8. Allahyar, M., Becic, E., Chappell, S., Fisher, D., Lohrenz, M., Monk, C., Philips, B.: The evolving role of automation in transportation. Human factors lessons learned from the different modes. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting 60, vol. 1, pp. 1971–1975 (2016)
9. Brown, J., Isaacs, D.: The world café: living knowledge through conversations that matter. Syst. Think. **12**, 5 (2001)

Erratum to: Advances in Human Aspects of Transportation

Neville A. Stanton(✉)

Human Factors Engineering, Transportation Research Group,
Boldrewood Innovation Campus, University of Southampton, Southampton, UK
n.stanton@soton.ac.uk

Erratum to:
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 597,
DOI 10.1007/978-3-319-60441-1

The original version of the book was inadvertently published without the following corrections:

In Chapter 1, Page 2, the position of Fig. 1 has to be at the end of the subchapter after the sentence "Combining these different human error definitions, an Integrated Threat Model was developed for this study to structure possible errors during pilot's work on flight decks:"

In Chapter 80, the incorrect author name "Gabrhe" has to be corrected to read as "Gabrhel".

The erratum book has been updated with the changes.

The updated online version for these chapters can be found at:
http://dx.doi.org/10.1007/978-3-319-60441-1_1
http://dx.doi.org/10.1007/978-3-319-60441-1_80

© Springer International Publishing AG 2018
N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1_108

Author Index

A
Abdella, Galal M., 807
Abdur-Rouf, Khadija B., 807
Abenoza, Roberto, 1036
Ahmad, Bashar I., 423
Aksan, Nazan, 242
Alford, Chris, 325, 760
Alfredson, Jens, 79, 117
Al-Ghafli, Abdulla, 251
Al-Khalifa, Khalifa N., 807
Alkim, Tom, 842
Allison, Craig K., 906
Almeda, Josiel Catherine V., 1025
Altendorf, Eugen, 631, 672
Anwer, Izza, 444
Aoyama, Hisae, 70
Ariansyah, Dedy, 307
Auflick, Jack L., 927
Aust, Mikael Ljung, 373
Ayala, Helon Vicente Hultmann, 117

B
Bailey, Randall E., 15
Baltzer, Marcel C.A., 490
Bang, Magnus, 79
Barbieri, Fábio Augusto, 289
Barnes, Jo, 1097
Baxendale, Anthony, 538
Bazilinskyy, Pavlo, 457
Beaumont, Charles, 457
Befelein, Dennis, 525
Beggiato, Matthias, 1072
Bella, Francesco, 855
Bellem, Hanna, 551, 708
Bengler, Klaus. J., 618
Bengler, Klaus, 595, 684, 708, 730
Berloco, Nicola, 571
Bernal, Fernando, 264
Bernard, Tiziano, 38
Berthelon, Catherine, 878
Blomstrand, Niklas, 79
Bordegoni, Monica, 307
Botura Jr., Galdenoro, 1121
Bowers, Kate C., 1109
Boy, Guy A., 38
Bradley, Mike, 749
Brázdil, Milan, 275
Broström, Robert, 373
Brown, James W.H., 906
Brünken, Roland, 282
Burnett, Gary, 583

C
Cabrall, Christopher D.D., 695
Caleb-Solly, Praminda, 325
Canpolat, Yigiterkut, 631, 672
Carizio, Bethânya Graick, 289
Carmo, Maria Beatriz, 199
Cefkin, Melissa, 650
Chai, Chen, 963
Charchalakis, Periklis, 503
Cheng, Jianxin, 484
Chung, Younshik, 884
Clarkson, John, 749
Clarkson, P. John, 395
Coelho, Leandro S., 117
Colón, Enid, 867
Colonna, Pasquale, 571
Colucci, Benjamín, 867
Colwill, Ian, 503
Conaghan, Madeleine, 503
Conceição, Victor, 199
Costa, Nicole A., 157, 213
Crawford, Elise, 825
Cui, Wenjie, 720
Czekóová, Kristína, 275

© Springer International Publishing AG 2018

N.A. Stanton (ed.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 597, DOI 10.1007/978-3-319-60441-1

D
Dahlman, Joakim, 199
Dawson, Jeffrey, 242
de Ângelo, Juliana Cristina, 289
de Jonge, Reinier, 457
de Winter, Joost C.F., 618
de Winter, Joost, 457, 643
Dell, Geoff, 825
Di Nicolantonio, Massimo, 170
Diana, Marco, 1036
Diels, Cyriel, 415, 538, 742
Ding, Li, 895
Dixon, Kate, 1097
Domingo, Majella Cayley J., 1025
Dong, Hua, 985
dos Santos, Claudemilson, 1121

E
Eckerberg, Maria, 79
Eckstein, Lutz, 606
Elton, Edward, 503
Escobar, Claudia, 467
Etherington, Timothy J., 15

F
Fang, Weining, 128
Farid, Mehdi, 684
Fay, Daniel, 191
Feldhütter, Anna, 730
Fenoglio, Franco, 145
Ferraris, Simona, 145
Ferreira, Ana, 199
Fleming, Mark, 1109
Flemisch, Frank, 490, 631, 672
Fornells, Alba, 234
Friedrich, Max, 96
Fu, Shan, 58
Fuest, Tanja, 708
Fujisawa, Shoichiro, 319, 817
Furuta, Kazuo, 70

G
Gabrhel, Vít, 300, 834
Galy, Edith, 878
Ganotis, Ioannis, 618
Gao, Yidan, 963
García, Ricardo, 867
Gertman, David I., 662
Ghadban, Nazha R., 807
Giacaglia, Marcelo Eduardo, 1010
Gilkey, David, 227
Glohr, Tobias, 337
Godsill, Simon J., 423
Gold, Christian, 551
Gotardi, Gisele Chiozi, 289
Grant-Muller, Susan, 444
Guo, Beiyuan, 128

H
Halbey, Julian, 939
Hamouda, Abdel Magid, 807
Happee, Riender, 643, 695
Hasselberg, Andreas, 96
Herzberger, Nicolas D., 798
Hilz, Jana, 282
Hornberger, Paul, 618
Hu, Linna, 985
Huang, Siyuan, 307

I
Ikeda, Norihiro, 319, 817
Inagaki, Tomoyuki, 817
Inoue, Satoru, 70
Intini, Paolo, 571
Ito, Shin-ichi, 319, 817
Izumi, Toshiyuki, 998

J
Jarosch, Oliver, 551
Jie, Yuwen, 58
Jin, Xiaoping, 511
John, Richard S., 27
Josten, Johanna, 606

K
Kaber, David, 775
Kanno, Taro, 70
Kennedy, Kellie D., 15
Khan, Ata, 362, 560
Kift, Ryan L., 825
Kim, Juyoung, 884
Klamer, Linda, 79
Kluge, Annette, 3
Koskinen, Hanna, 1059
Kramer, Lynda J., 15
Krems, Josef, 1072
Kurečková, Veronika, 275, 300, 834

L
Lagatta, Jessica, 182
Lagerstrom, Elise, 227
Langdon, Patrick M., 423
Langdon, Patrick, 395, 749
Large, David R., 583
Lassen, Christian, 490
Lau, Mei Ying, 775
Ledin, Johanna, 79

Lee, Shawn, 467
Li, Haoyang, 511
Li, Tengye, 484
Li, Yan, 895
Lindvall, Sofia, 373
Lioris, Jennie, 1133
Liu, Aie, 895
Liu, Haoran, 662
Liu, Long, 985
Löber, Cathrine, 684
López, Daniel, 490
Lu, Yanyu, 58
Lu, Zhenji, 643
Lundh, Monica, 157, 213
Luo, Hong, 407
Luo, Ling, 511
Lützhöft, Margareta, 213

M
Ma, Ke, 349
MacKinnon, Scott N., 157, 213
Maeda, Setsuo, 474
Maier, Thomas, 337
Majumdar, Arnab, 48
Malivern, Melcior, 264
Malone, Sarah, 282
Man, Yemao, 213
Mansfield, Neil, 474
Marberger, Claus, 595
Mariani, Eliete, 1010
Marini, Robert, 242
Martens, Marieke H., 842
Martinussen, Laila M., 787
Matthias, Rebecca, 583
Michel, Britta, 383
Mielenz, Holger, 595
Miglianico, Denis, 1046
Miura, Toshiro, 998
Miyamoto, Takashi, 319
Møller, Mette, 787
Morgan, Phillip L., 325, 760
Morris, Andrew, 583
Mouchel, Mathieu, 1046
Mouzakitis, Alex, 395, 749
Moyart, Luc, 1046
Musso, Giorgio, 145
Muthumani, Arun, 583

N
Nagahama, Hideaki, 817
Naujoks, Frederik, 525, 551, 595
Neubauer, Matthias, 1145
Neukum, Alexandra, 525
Neves, Hali, 227
Nguyen, Kenneth D., 27

O
Ochieng, Washington, 48
Ogino, Hiroshi, 817
Okada, Tomohiro, 319

P
Papenfuß, Anne, 96
Parera, Núria, 234, 264
Parkhurst, Graham, 325, 760
Parnell, Katie J., 906
Paschoarelli, Luis Carlos, 289
Payre, William, 415
Pei, Yeqing, 511
Petermeijer, Sebastiaan M., 618
Petrovych, Veronika, 695
Philipsen, Ralf, 606, 939, 951
Pipe, Anthony, 325
Pipe, Tony, 760
Pirra, Miriam, 1036
Polastri, Paula Fávaro, 289
Politis, Ioannis, 749
Prato, Carlo G., 787
Pray, Ira, 662
Pritchard, James, 973
Psyllou, Elena, 48

Q
Qiu, Hanzhao, 128

R
Radlmayr, Jonas, 551, 595, 684
Rämä, Pirkko, 1059
Ran, Linghua, 407, 720
Ranieri, Vittorio, 571
Řezáč, Pavel, 275, 300
Rezac, Pavel, 834
Ricci, Raffaella, 145
Rizzo, Matthew, 242
Roa, Julio, 139
Roberts, Aaron P.J., 191
Rodrigues, Sérgio Tosi, 289
Rohde, Richard W., 108
Rosecrance, John, 227
Ruff-Stahl, Hans-Joachim, 87
Ruiz, Bryan, 867
Ruiz, Johnathan, 867
Ruscio, Daniele, 307, 916

S
Sahnoon, Iyad, 251
Salatino, Adriana, 145

Salazar, Miguel, 275
Sammonds, George, 474
Sanchez, Jacob, 264
Saringan, Alyannah Bianca L., 1025
Sato, Katsuya, 319, 817
Schauer, Oliver, 1145
Schmidt, Teresa, 606, 939, 951
Schreck, Constanze, 631
Schütz, Raphael, 672
Schwalm, Maximilian, 798
Segler, Christoph, 730
Shaw, Daniel, 275
Shaw, Susan, 467
Shawky, Mohamed, 251
Shergold, Ian, 325
Shi, Jintian, 349
Siebericchs, Sebastian, 3
Sikora, Ivan, 87
Silvestri, Manuel, 855
Simoes, Anabela, 1082
Skrypchuk, Lee, 395, 423, 749
Sodnik, Jaka, 435
Song, Tai-Jin, 884
Song, Zhenghe, 511
Sorokin, Lenja, 708
Spielman, Zachary, 662
Špiláková, Beáta, 275
Springer, Sabine, 1072
Stanton, Neville A., 191, 906
Stecher, Michael, 383
Stephane, Lucas, 38
Stipidis, Elias, 503
Stojmenova, Kristina, 435
Sueda, Osamu, 319
Susilo, Yusak, 1036

T
Takahashi, Kazuya, 817
Takeuchi, Kiyohito, 817
Tarander, Jasmine, 79
Tatsuno, Junya, 474
Thompson, Simon, 742
Tippin, Jon, 242
Toft, Yvonne, 825
Tovey, Michael, 538, 1036
Trabasso, Luís G., 79, 117
Traiteur, Justin, 662
Turetta, Felipe M.S., 117

V
Valdés, Didier, 867
Vallicelli, Andrea, 182
van der Geest, Xander, 457
van der Kroft, Koen, 457
Vaquer, Oriol, 234
Vinkhuyzen, Erik, 650
Vogel, Daniel, 87
Voinescu, Alexandra, 325
Voß, Gudrun M.I., 798

W
Wandtner, Bernhard, 595
Wang, Bo, 128
Wang, Li, 128
Wang, Xuesong, 963
Wasser, Joscha, 538
Waterson, Patrick, 1097
Weinbeer, Veronika, 684
Weßel, Gina, 631, 672
Wiedemann, Katharina, 525
Willemsen, Dehlia M.C., 842
Williams, Craig, 325, 760
Wilschut, Ellen S., 842
Witzlack, Claudia, 1072
Wold, Scott, 662
Wong, Yiik Diew, 963
Woodcock, Andree, 1036
Wysmuller, Steven, 662

X
Xiong, Tao, 484

Y
Yan, Yunhao, 407, 720
Yang, Yimu, 985
Yi, Jun, 307
Yoshida, Haruka, 70
Yu, Mengli, 407

Z
Zafarana, Antonio, 145
Zámečník, Petr, 275, 300, 834
Zhang, Bo, 842
Zhang, Jingyi, 650
Zhang, Xin, 407, 720
Zheng, Bowen, 511
Zheng, Yiyuan, 58
Zhou, Ronggang, 720
Ziefle, Martina, 606, 939, 951
Zimmermann, Andreas, 337, 383