Advances in Intelligent Systems and Computing 484

Neville A. Stanton
Steven Landry
Giuseppe Di Bucchianico
Andrea Vallicelli *Editors*

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

Springer

Advances in Intelligent Systems and Computing

Volume 484

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.

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

Neville A. Stanton · Steven Landry
Giuseppe Di Bucchianico
Andrea Vallicelli
Editors

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

Springer

Editors
Neville A. Stanton
Faculty of Engineering and the Environment
University of Southampton
Southampton
UK

Giuseppe Di Bucchianico
Dipartimento di Architettura
Università degli Studi "G.d'Annunzio" Chiet
Pescara
Italy

Steven Landry
Purdue University
West Lafayette, IN
USA

Andrea Vallicelli
Dipartimento di Architettura
Università degli Studi "G.d'Annunzio" Chiet
Pescara
Italy

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-41681-6 ISBN 978-3-319-41682-3 (eBook)
DOI 10.1007/978-3-319-41682-3

Library of Congress Control Number: 2016944169

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Advances in Human Factors and Ergonomics 2016

AHFE 2016 Series Editors

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

7th International Conference on Applied Human Factors and Ergonomics

Proceedings of the AHFE 2016 International Conference on Human Factors in Transportation, July 27–31, 2016, Walt Disney World®, Florida, USA

Advances in Cross-Cultural Decision Making	*Sae Schatz and Mark Hoffman*
Advances in Applied Digital Human Modeling and Simulation	*Vincent G. Duffy*
Advances in Human Factors and Ergonomics in Healthcare	*Vincent G. Duffy and Nancy Lightner*
Advances in Affective and Pleasurable Design	*WonJoon Chung and Cliff (Sungsoo) Shin*
Advances in Human Aspects of Transportation	*Neville A. Stanton, Steven Landry, Giuseppe Di Bucchianico and Andrea Vallicelli*
Advances in Ergonomics In Design	*Francisco Rebelo and Marcelo Soares*
Advances in Ergonomics Modeling, Usability & Special Populations	*Marcelo Soares, Christianne Falcão and Tareq Z. Ahram*
Advances in Social & Occupational Ergonomics	*Richard Goossens*
Advances in Neuroergonomics and Cognitive Engineering	*Kelly S. Hale and Kay M. Stanney*
Advances in Physical Ergonomics and Human Factors	*Ravindra Goonetilleke and Waldemar Karwowski*
Advances in The Ergonomics in Manufacturing: Managing the Enterprise of the Future	*Christopher Schlick and Stefan Trzcielinski*
Advances in Safety Management and Human Factors	*Pedro Arezes*
Advances in Human Factors, Software, and Systems Engineering	*Ben Amaba*
Advances in Human Factors and Sustainable Infrastructure	*Jerzy Charytonowicz*
Advances in The Human Side of Service Engineering	*Tareq Z. Ahram and Waldemar Karwowski*

(continued)

(continued)

Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries	*Sacit Cetiner, Paul Fechtelkotter and Michael Legatt*
Advances in Human Factors in Sports and Outdoor Recreation	*Paul Salmon and Anne-Claire Macquet*
Advances in Human Factors and System Interactions	*Isabel L. Nunes*
Advances in Human Factors, Business Management, Training and Education	*Jussi Kantola, Tibor Barath, Salman Nazir and Terence Andre*
Advances in Human Factors in Robots and Unmanned Systems	*Pamela Savage-Knepshield and Jessie Chen*
Advances in Design for Inclusion	*Giuseppe Di Bucchianico and Pete Kercher*
Advances in Human Factors in Cybersecurity	*Denise Nicholson, Janae Lockett-Reynolds and Katherine Muse*

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 factor 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 facing 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 the 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 facing human factors experts in rail and road, aviation, 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 15 sections over three volumes.

Section 1: Road and Rail—Ergonomic Analysis and Assistance
Section 2: Aviation—Human Factors in Aviation
Section 3: Road and Rail—Pedestrians and Intersections
Section 4: Road and Rail—Driver, Behavior, Distraction and Fatigue

Section 5: Maritime—Human Performance and Safety Assessment in the Maritime Domain
Section 6: Road and Rail—Vehicle Automation
Section 7: Road and Rail—Logistics and Passengers
Section 8: Road and Rail—Accidents and Pedestrian Modeling
Section 9: Road and Rail—Warning Systems/Public Transport
Section 10: Aviation—Human Factors in Aviation
Section 11: Road and Rail—Eco-Driving and Electric Vehicles
Section 12: Road and Rail—Education and Hazard Perception
Section 13: Road and Rail—Infrastructure
Section 14: Maritime—Users, Tasks and Tools in the Maritime Domain
Section 15: Road and Rail—Safety, Driver Psychophysiology and Eye Tracking

This book will be of interest and use to transportation professionals who work in the road and rail, aviation, 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 the book. These people are academic leaders in their respective fields, and their help is very much appreciated, especially as they gave their time freely to the project.

S. ur Rehman, Sweden
G. Walker, Scotland
K. Young, Australia

Aviation

M. Biella, Germany
C. Borst, The Netherlands
T. Edwards, USA
M. Feary, USA
A. Haslbeck, Germany
B. Hooey, USA
D. Kaber, USA
K. Latorella, USA
A. Majumdar, UK
L. Martin, USA
J. Mercer, USA
M. Mulder, The Netherlands
S. Verma, USA
K. Vu, USA

Maritime

A. Alkan, Turkey
D. Andrews, UK
D. Gray, USA
M. Grootjen, The Netherlands
T. Koester, Denmark
S. MacKinnon, Canada
M. Musio Sale, Italy
G. Praetorius, Sweden
A. Ratti, Italy

Southampton, UK — Neville A. Stanton
West Lafayette, USA — Steven Landry
Pescara, Italy — Giuseppe Di Bucchianico
Pescara, Italy — Andrea Vallicelli
July 2016

Contents

Part II Aviation—Human Factors in Aviation

Part III Road and Rail—Pedestrians and Intersections

Part X Aviation—Human Factors in Aviation

Part XI Road and Rail—Eco-driving and Electric Vehicles

Part I
Road and Rail—Ergonomic Analysis and Assistance

Modelling Human Factors for Advanced Driving Assistance System Design

Ata Khan

Abstract Although technological developments in experimental autonomous vehicles are impressive, industry experts are realizing that if automation in driving is to gain acceptance by drivers and to become a reality in real world driving environment, new generation driving assistance system (NDAS) shaped by human factors is necessary since it will provide transition to self-driving vehicles. The paper consists of five parts. In part one, the balance of demand pull vs. technology push is introduced and part two reports developments in driving assistance technologies. In the third part, transitions between human control and automation are described as high level "design" challenges. In part four, a Bayesian Artificial Intelligence (AI) model is presented that enables the NDAS to perform its functions. An example application based on driving simulator data from distracted driving study is presented to illustrate advanced driving assistance capabilities. Finally, in part five, conclusions are presented on how human factors-guided NDAS design is likely to enhance driver acceptance.

Keywords New generation driving assistance system (NDAS) · Advanced driving assistance system (ADAS) · Autonomous vehicle · Modelling · System design · Human factors

1 Introduction: Demand for Automation Versus Technology-Push

Experimental autonomous vehicles are undergoing tests in terms of "proof of technology". In parallel, initial public policy steps have been taken in a few jurisdictions to allow drivers to use their autonomous vehicles on public roads. So far, it appears that the proponents of these initiatives are technology developers and there is a lack of convincing evidence that there is a market demand for autonomous vehicles.

A. Khan (✉)
Carleton University, Ottawa, ON K1S5B6, Canada
e-mail: ata_khan@carleton.ca

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_1

In order for a balance to exist between demand for automation vs. technology-push, there should be a driver-perceived need for new technology and driver acceptance of well-designed systems.

Available published information on motorists' preferences for different levels of vehicle automation does not support high levels of automation. A recent study report published by the University of Michigan's Transportation Research Institute (UMTRI) presents results of a survey of 505 licensed drivers in the USA [1]. These show that: (1) No self-driving capability was the most frequent preference for vehicle automation. (2) Next was partially-self driving vehicle. (3) Completely self-driving vehicles was the least preferred choice. (4) Respondents' concern for riding self-driving vehicles was higher for completely self-driving vehicles as compared with the partially self-driving vehicles. (5) A strong preference was expressed to manually control completely self-driving vehicles, when desired. (6) Most respondents prefer that a notification should be given of the need to take control of a partially self-driving vehicle and the means for notification could be a combination of sound, vibration, and visual warnings.

The difficult part of autonomous technology vehicle assessment in terms of user acceptance is still to come. If automation in driving is to gain acceptance by drivers and to become a reality in real world traffic driving environment, new generation driving assistance system (NDAS) shaped by human factors should receive the necessary attention since it will provide transition to self-driving vehicles.

2 Levels of Technological Advances

Khan et al. [2] provided a projection of technological advances as a part of a paper on policy challenges of increasing automation in driving. On the technology development front, innovations have progressed as projected along the continuum between conventional fully human-driven vehicles and autonomous vehicles.

Technological developments continue to improve the driving assistance system [3]. The initial version of this system did not interact with the human driver. Next came the current advanced driver assistant (ADAS) noted as Level II technology in Fig. 1. Further research is expected to deliver the NDAS 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 believe that the next step is the development of 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 [4–6]. Table 1 presents the design features of the NDAS which will be a key component of the cognitive vehicle.

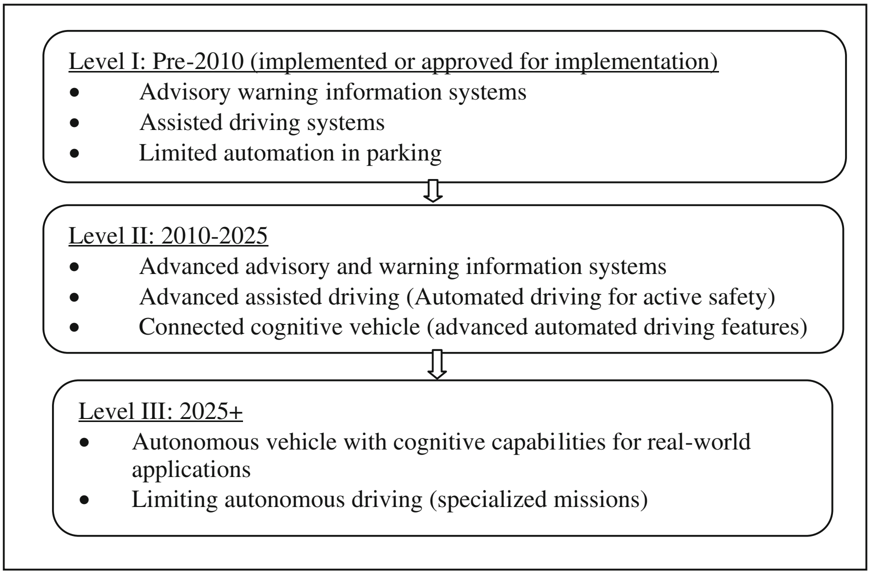


Fig. 1 Levels of technological advances and time frame [2]

Table 1 New generation driving assistance system of the cognitive vehicle

Design features	Driver assistance system capabilities
Extension of human driver capabilities	• Situational awareness (position, surroundings)
	• Ability to gather data and send out data
	• Ability to process data
	• Ability to cooperate/collaborate
	• Communication for active safety
	• Informing driver about situations (warning, advice)
	• Diagnostics capability
	• In case of crash, capability to send and receive information
	• Ability to provide non-distractive user interface for safe and efficient operation
	• Capability to perform user-requested infotainment tasks (not related to safety)

Source Adapted from [4]

A cognitive connected vehicle that will include the NDAS can function under human control and also in automated, highly automated, and fully autonomous mode as well [4]. The NDAS is expected to assist the human driver and if asked, will make decisions in driving. Among other component of the NDAS, the crash warning, active safety, adaptive cruise control, and lane-keeping systems are assigned safety, efficiency, and convenience roles.

3 Transitions Between Human Control and Automation

The NDAS should have cognitive features that mimic non-distracted and non-aggressive driving tasks. It is intended to assist the driver, and if necessary in dangerous conditions, it will have the capability to take corrective active safety action should the driver be incapacitated or highly distracted or if the driver selected the automation option. However, driving the cognitive vehicle does not take the driver out of the loop. The design attributes of NDAS should be influenced by human factors in driving. According to a recent news article, development of 'human-like' self-driving technologies is attracting investor capital [7].

Technological forecasts point in the direction of automation in driving in the long term and the availability of multi-functional high capability driving assistance in the short term. But further research is needed on how to integrate human and technology factors in order to make the human control and automation seamless and to overcome shared authority concerns in increasing automation in driving [8]. An attempt was made by the author of this paper to suggest the Bayesian artificial intelligence approach to the design of real-time transition from driving assistance to automation function [9].

4 Bayesian Artificial Intelligence

4.1 The Bayesian Approach

The Bayesian methodology enables system design and decision analysis when uncertain "states of nature" are encountered, but there are opportunities to refine knowledge of uncertain factors such as driving states, driver distraction and driver intention. According to Korb and Nicholson, artificial intelligent is the "intelligence developed by humans, implemented as an artefact" [10]. Specifically, the Bayesian artificial intelligence integrates two cognitive features of the NDAS. The first one is the descriptive artificial intelligence, which models a desired human action (e.g., non-distracted non-aggressive driving). The second is to model our rational view of what is "optimal" and apply it as a decision criterion [11, 12].

The Bayesian artificial intelligence approach is applied in three steps. 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.

Application of intelligent technology and human factors in the NDAS 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 ADAS models. The NDAS described in this paper has the capability to avoid rear as well as lateral crashes. The high level architecture of NDAS is shown in Fig. 2.

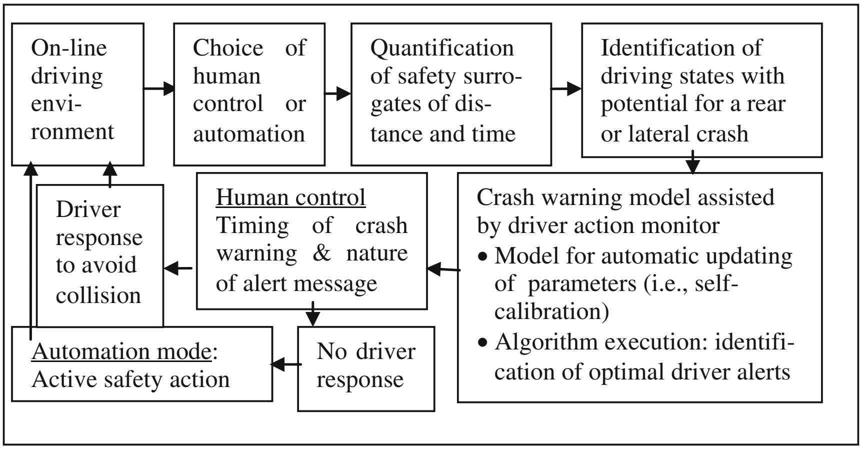


Fig. 2 High level architecture of driving assistance system's safety function

4.2 Modelling Human and Automation Interface

The seamless automated transition from human control to active safety can be modelled as shown in the architecture presented in Fig. 2. Figures 3 and 4 show the transition model algorithm that can operate the NDAS.

The operation of the NDAS is briefly described here. In the vehicle-following mode (i.e. travel in the longitudinal direction), the distance between the subject vehicle and the leading vehicle and speeds of these vehicles are monitored on a real time basis. If the longitudinal distance is less than or equal to 1.5 times the critical distance needed to avoid a collision in "non-ideal" driving condition, the algorithm

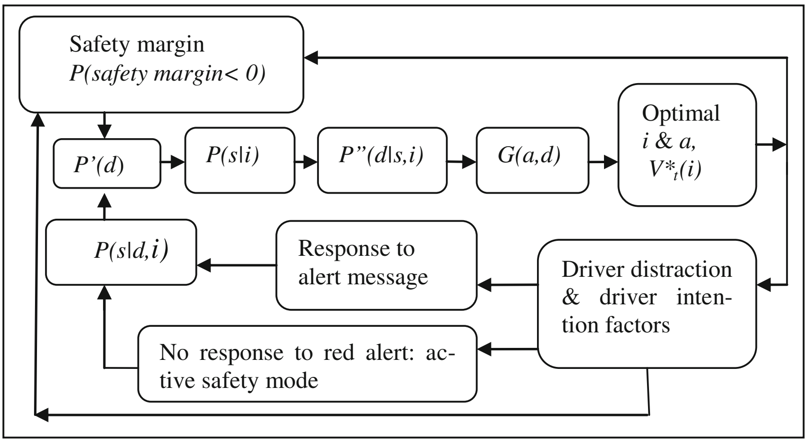


Fig. 3 Collision warning and active safety system

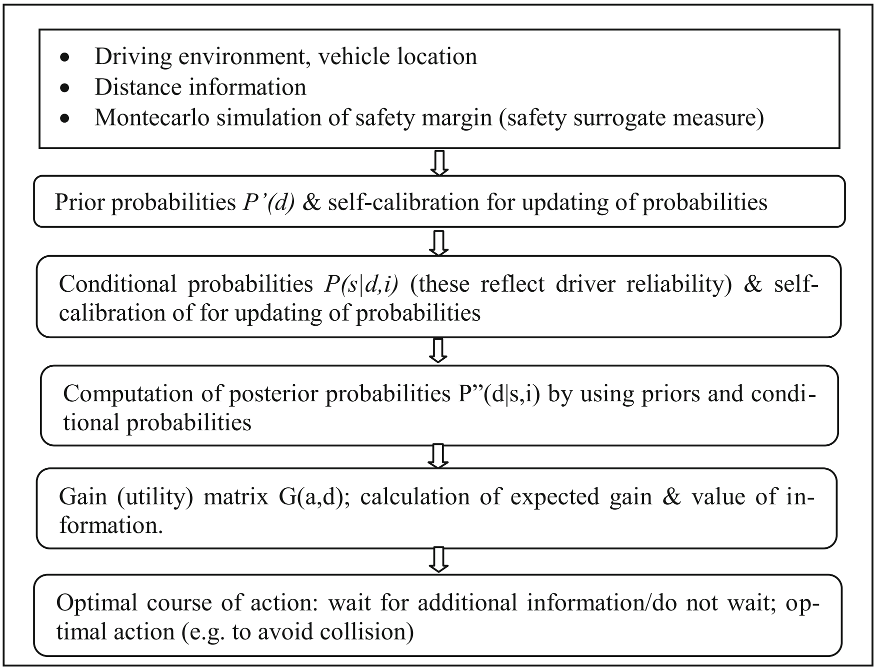


Fig. 4 Algorithm for new generation driver assistance system

is launched. Likewise, if the normal lateral distance is exceeded, the safety algorithm is launched. The components of the transition model are presented in Fig. 4 and described briefly in this paper. For details, please see Khan [11–13].

In the case of longitudinal direction rear crash, states of driving are defined by the distance between the following and the leading vehicles: d_c, $d_{1.25c}$, $d_{1.5c}$:

d_c (critical distance, if exceeded would lead to a rear crash—but can be avoided if the required action is taken by the driver),
$d_{1.25c}$ represents 1.25 times the critical distance, and
$d_{1.5c}$ is 1.5 times the critical distance.

The number of states as well as the 1.0, 1.25, 1.5, etc. multipliers can be changed by the designer.

Possible surrogate measures (i.e., readings) on the states of longitudinal driving condition are: s_0 (no new reading, if i_0 is selected), s_c (corresponds to d_c), $s_{1.25c}$ (corresponds to $d_{1.25c}$), and $s_{1.5c}$ (corresponds to $d_{1.5c}$). The corresponding multipliers for the lateral collisions are s_c (corresponds to d_c), $s_{1.5c}$ (corresponds to $d_{1.5c}$), and s_{2c} (corresponds to d_{2c}) [9, 11–13].

Driver information (alerts) are: i_0 (early applicable warning issued on the basis of initial information), i_w (the waiting mode so as to acquire and analyze additional safety surrogate information on the dynamics of vehicle-following or lane migration/lane change/merge and then issue the appropriate warning, if applicable) [9, 11–13].

A Montecarlo simulation module estimates safety margins within the driving environment. As driving progresses, the prior probabilities *P'(d)* and the conditional probabilities $P(s|d, i)$ are updated automatically. As noted in Fig. 4, these reflect driver reliability. The other probabilities, namely the marginal probability $P(s|i)$ and posterior probability $P''(d|s, i)$, are computed internally by the algorithm. Details of the model are reported by Khan in Refs. [9, 11–13].

The driving assistance can analyse a number of alternatives noted next: a_0 (no driver alert to be issued), a_a (amber alert—calls for higher than normal deceleration in the longitudinal direction or steering action to increase transverse separation distance), a_r (red alert—calls for serious emergency deceleration for avoiding a rear crash or steering action for preventing a lateral collision).

The crash warning-active safety component is intended to work with a driver action monitor. The system can be enabled to automatically update key driving parameters, namely the probabilities of safety surrogates (i.e., distances or times), as well as the probability of driver's awareness of distance and time. This adaptive self-calibration capability is reported by Khan [11–13].

For example, if a driver is distracted and the probability of crash is high, and to make matters worse, the driver does not show the intent to take corrective action, the driver monitoring part of the system will immediately modify the driver alertness parameter.

The transition model shown in Figs. 3 and 4 can identify optimal driver alerts in terms of the timing of rear or lateral crash warning (i.e., immediate or wait for an indication of driver intent) and the nature of alert message (e.g., no alert message, amber alert, red alert). If a highly distracted or a disabled driver does not respond, the system automatically initiates active safety action and informs the driver accordingly.

In order to apply the algorithm for identifying optimal *i&a* and the associated value of information $V^*_t(i)$, on a real time basis, the utility or gain $G(a, d)$ matrix has to be defined by the designer following consultations with policy experts. For details, please see Khan [9, 11–13].

4.3 Example Application to Distracted Driving

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 a shock wave 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.

The driving simulator provides driving trajectories of each driving mission on a split second basis. The time stamp provides a complete log of all driver actions in terms of speed, brake action, deceleration/acceleration, headway, lane keeping, etc.

For illustration of the application of NADS, a driving trajectory of a distracted driver was selected (Figs. 5, 6, 7, 8 and 9). At time stamp 84 s, the driver was travelling at 104.8 km/h and at that time driver did not perceive and react to the hazard within 1 to 1.25 s, which is common for young non-distracted drivers. Figure 5 shows that at time stamp 84 s, the driver was expected to perceive that the leading vehicle applied brakes suddenly and that quick action was necessary to avoid a collision. But the driver did not apply brakes until time stamp 88 s. This perception-reaction time of 4 s is very high as compared to about 1 s for an alert driver. Another indicator of driver distraction was that the lane migration (i.e. distance from the lane centre-line) was about 0.32 m out of 1.3 m (which is about 25 % of the one-half of the lane width).

At time stamp 88 s, brakes were applied on an emergency basis that delivered a deceleration rate of about 8.5 m/s^2. Figures 5, 6, 7, 8, 9 show the driving state descriptors during the braking action and then returning to a manageable driving state.

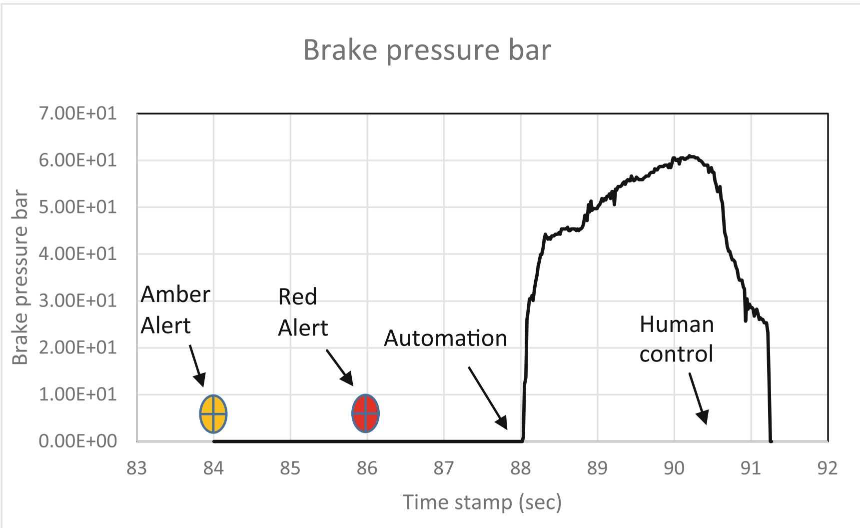


Fig. 5 Distracted driver warnings, transition to automation, and return of control to human driver

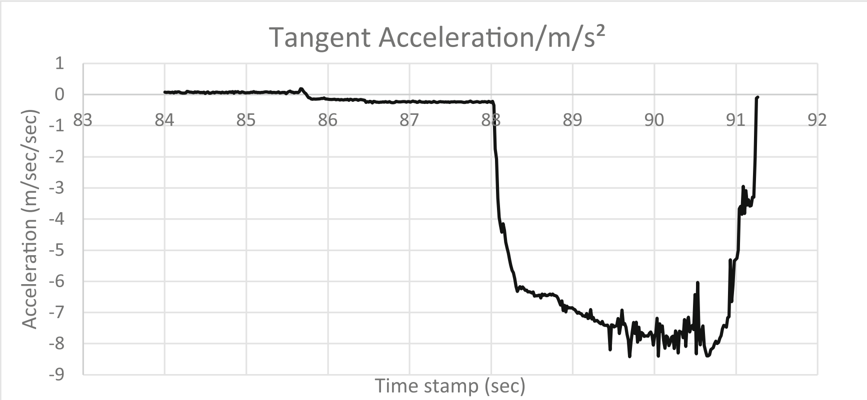


Fig. 6 High level of deceleration

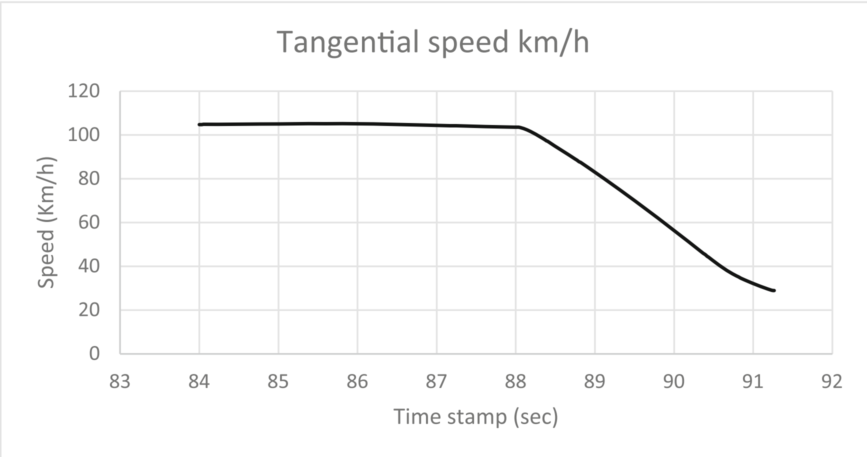


Fig. 7 Speed profile

Under non-ideal driving condition, the critical distance required for braking was 150.6 m vs. available distance of 56.3 m. The ideal driving condition can be characterized as very high tire-pavement coefficient of friction, new brakes, anti-lock brakes, and high ability of the driver to push the brake pedal. Calculations show that the conditional probability $P(d|s, i)$ was 0.37 at the critical distance level (i.e. at 1.0c), 0.51 at 1.025c, 0.92 at 1.25c and 1.0 at 1.5c. The profile of conditional probability presented in Fig. 10 shows that the driver was highly distracted.

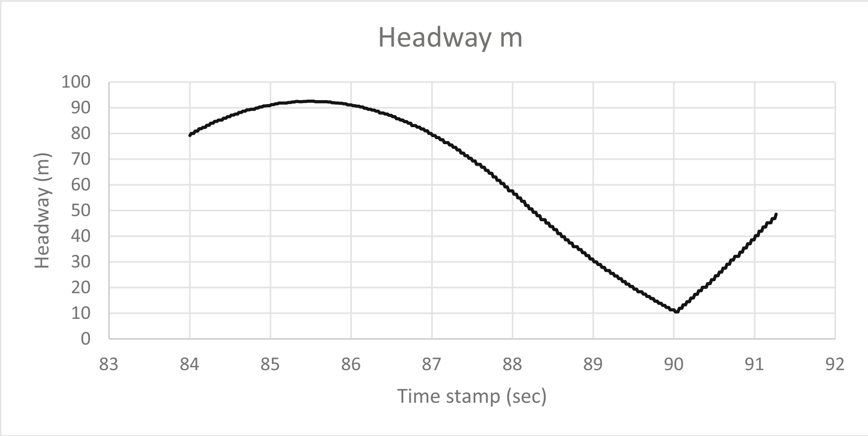


Fig. 8 Headway profile

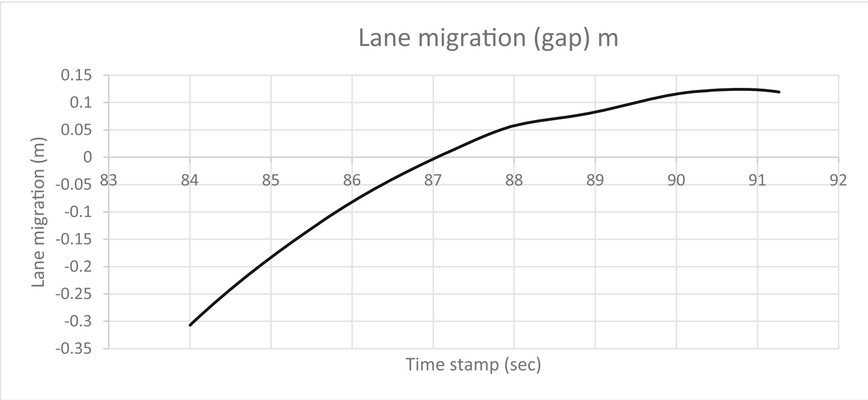


Fig. 9 Distracted driver's lane departure pattern

In the scenario that the driving assistance system was available and turned-on, the following could be the optimal decisions (Fig. 5):

- At time stamp 84 s, do not wait for additional information, give amber alert.
- At time stamp 85 s, keep amber alert, and wait for additional information.
- At time stamp 86 s, do not wait for additional information, and give red alert.
- At time stamp 87 s, wait for additional information before initiating active safety action.
- At time stamp 88 s, apply active safety.

An earlier time for automation could be considered at time stamp 87 s.

Upon driver's preference, control of the vehicle could be returned to human control at time stamp 90.5 s.

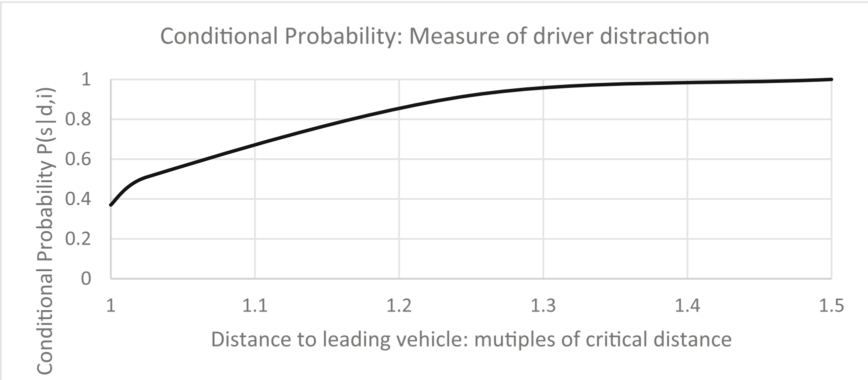


Fig. 10 Conditional probability (an indicator of driver reliability in assessing distance)

5 Conclusions

1. Technology-push does not appear to shape consumer demand for new in-vehicle technologies.
2. A new generation driving assistance system has an important role to play in the transition to automation in driving.
3. The design of the driving assistance, if guided by human factors, is likely to enhance driver acceptance. Consequently, safety benefits will be achieved.
4. The driving assistance example case illustrates the integration of intelligent technology, Bayesian artificial intelligence, and abstracted human factors.

Acknowledgments This paper is based on research that was 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. Schoettle, B., Sivak, M.: Motorists’ Preferences for Different Levels of Vehicle Automation. University of Michigan Transportation Research Institute, UMTRI-2015-22 (2015)
2. Khan, A.M., Bacchus, A., Erwin, S.: Policy challenges of automation in driving. IATSS Res. **35**, 79–89 (2012)
3. Hannes, E.: Paving the way to self-driving cars with advanced driver assistance systems. In: Worldwide Systems Marketing for Advanced Driver Assistance Systems (ADAS), Texas Instruments (2015)
4. 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)
5. Heide, A., Henning, K.: The “cognitive car”: a roadmap for research issues in the automotive sector. Annu. Rev. Control **30**, 197–203 (2006)

6. 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)
7. Traffic Technology Today: News: Autonomous Vehicle Software Developer Secures Funding for 'Human-Like' Driving, 5 Feb 2016
8. Cummings, M.L., Ryan, J.: Shared Authority Concerns in Automated Driving Applications. www.web.mit.edu/aeroastro/labs/halab/papers/cummingsryan_driverless2013_draft.pdf (2013)
9. 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, pp. 39–48. Springer, Berlin (2014)
10. Korb, K.B., Nicholson, A.E.: Bayesian Artificial Intelligence. Chapman & Hall/CRC, UK (2004)
11. Khan, A.M.: Cognitive connected vehicle information system design requirement for safety: role of Bayesian artificial intelligence. Syst. Cybern. Inf. **11**(2), 54–59 (2013)
12. 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)
13. Khan, A.M.: Design of adaptive longitudinal control for cognitive connected vehicle. In: Proceedings, ITS World Congress, Orlando, USA (2011)

Effects of Driver Characteristics and Driver State on Predicting Turning Maneuvers in Urban Areas: Is There a Need for Individualized Parametrization?

Matthias Graichen and Verena Nitsch

Abstract In future, advanced driver assistance systems (ADAS) may be able to adapt to the needs of the driver, thus reducing the risk of information overload in complex traffic situations. One way of achieving this may include the use of predictive algorithms that anticipate the driver's intention to perform a certain traffic maneuver based on vehicle data, such as acceleration and deceleration parameters. In order to explore whether the predictive quality of such algorithms may be mitigated by individual driver-specific parameters such as driver characteristics (i.e. emotional driving [ED] and uncritical self-awareness [US]) as well as driver state (specifically stress), an empirical test-track study was conducted with N = 40 participants. The results indicate that maximum longitudinal and lateral acceleration vary significantly depending on driver characteristics. Moreover, analyses of the collected data suggest that incorporating psychological aspects into driver models can promote new insights into driving behavior.

Keywords Maneuver prediction · Driver intention · Driver model · Driver characteristics · Driver state · Driver behavior · Driver assistance · Urban · Intersection

M. Graichen (✉) · V. Nitsch
Human Factors Institute, Universität der Bundeswehr München,
Werner-Heisenberg-Weg 39, 85579 Neubiberg, Germany
e-mail: Matthias.Graichen@unibw.de

V. Nitsch
e-mail: Verena.Nitsch@unibw.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_2

1 Introduction

Efforts in design and development of advanced driver assistance systems (ADAS) focus currently on challenges of urbanization, which is expected to increase significantly over the next few decades [1–3]. Inner-city traffic in metropolitan areas exhibit great complexity and diversity, not only with respect to the various road design features (e.g. architecture, type of traffic regulation, number of turning lanes, etc.) and other environmental aspects (visibility; e.g. [4]), but mainly in terms of number, types and dynamics of different traffic participants. Since the driver is necessarily confronted by a large amount of information in these complex traffic situations, it is of paramount importance to minimize the amount of information that is transmitted to the driver by the ADAS in order to ensure high levels of user acceptance and thereby also frequent voluntary usage of these assistive systems [5]. This can be achieved by reducing unnecessary warnings and making them adaptive to drivers' intentions and maneuver-planning [6]. This appears particularly necessary in highly complex traffic situations that include approaching and passing through intersections.

Urban intersections are a major accident hotspot: 25 % of accidents with personal injuries occur at intersections, 16 % while turning [7]. Here, the driver has to detect, identify and assess a large amount of visual stimuli correctly and within a brief time span [8], while scanning areas with decreasing distal proximity to the car at the beginning of the approaching phase and immediate proximity after reaching the intersection and performing the actual driving maneuver [9]. During this time, adaptive ADAS would only provide information about potential crossing cyclists or pedestrians if it was predicted with high certainty that the intention of the driver was to turn—but not in the case of drivers passing without turning, as no critical interaction with the other traffic participants is to be expected in this case. Numerous studies have demonstrated the potential of using vehicle data for predicting (turning) maneuvers (for a general overview see [10]).

As part of the on-going efforts to enhance the performance and adaptability of ADAS, (sensory) perception, evaluation and prediction of traffic situations and driving maneuvers play an important role. The present study aimed to explore the extent to which the predictive quality of vehicle data-based algorithms might be mitigated by specific driver characteristics that are assumed to influence the driving behavior during the approach to the intersection. For this purpose, a prediction algorithm for turning maneuvers based on Bayesian networks is used [11]. A model is stipulated based on two essential parameters: (1) speed models, e.g. acquired through (a) clustering of (desired) speed profiles of different drivers approaching intersections, or (b) obtained from path curvature, and (2) a maximum acceleration parameter (longitudinal acceleration after performing the turning maneuver for 1a, or lateral acceleration when during the turning process for 1b). Using the intelligent driver model (originating from traffic flow modeling), potential speed profiles for the maneuver types 'going straight', 'stopping', 'turning (right)' and 'turning (right), but stopping' can be modelled on the basis of any previously shown speed data of the driver (alternatively, speed profiles can also be modeled using modelled curvature data, without any driving data). When taking into account the last shown speed data point of the driver, the probability for any of these maneuvers can be computed.

The relationship of these parameters to psychological aspects lies within the initialization by the (individual) driver and corresponding driver operations (e.g. pedal activity and acceleration). These may be influenced by individual preferences and characteristics (as can be inferred from literature, e.g. [12, 13]). The following analyses investigate speed, acceleration and the underlying operational processes of the driver with respect to individual driver characteristics and driver state (in particular stress), which serve as input for the stated algorithm.

2 Method

2.1 Sampling and Participants

In order to measure and control for relevant driver characteristics, a pre-sampling process was applied. As a first step, participants possessing a valid driver's license could apply for the study by completing a validated online questionnaire about driving-related personality (VIP [14]), driving style (adapted from [15]), driving experience and demographic information. In a second step, potential participants out of 186 applications were selected based on extreme (high or low) scores on the VIP subscales emotional driving (ED) and uncritical self-awareness (US). These subscales were chosen, as their items concerned contents relevant to speed behavior and its sensitivity to driver state. In addition, only participants were included if they had indicated a certain amount of driving experience. Specifically, in reference to [16], participants were selected if they had acquired their driver's license before 2014 and had driven more than 30,000 km overall or more than 25,000 km in the last 12 months.

The resulting sample consisted of 40 participants (8 female), with a mean age of 36.78 years (Min = 20, Max = 61, SD = 13.14). They were assigned to one of four groups according to their scores on the ED and US subscales (N = 10 in each group). Participants had an average total driving experience of 354,625 km (SD = 287,244).

2.2 Research Design

The study was conducted as a 2 (approaching speed) * 3 (driving maneuver) * 4 (driver characteristic) repeated-measures mixed design with two consecutive runs (labelled as 'time') on a closed test track. Approaching speed was manipulated within-subjects with intersection approaching speeds of 50 or 70 km/h (labelled as ‘apprspeed’). Approaching intersections with 70 km/h was considered an upper limit in the context of driving in urban areas, but also appeared to alter the driver state by inducing stressful experiences when participants were reminded repeatedly and in short intervals to maintain that speed after each turning and when driving through curves. The within-condition of approaching speed was systematically varied, meaning that half of each group was instructed to approach the intersection with 50 km/h in the first two runs and 70 km/h in the second two runs, and vice versa. Each participant performed three types of driving maneuvers (left or right turn, or going straight). The sequence of maneuver types had to be held constant due to the structural requirements of the test track. Overall, there were twelve intersection maneuvers for each participant, with four runs for each direction and one repetition at each level of approaching speed. Finally, the factor driver characteristic was manipulated between-subjects, with participants being assigned to one of four groups based on their scores on the VIP subscales ED and US (labelled as ‘VIP’).

2.3 Measurements and Hypotheses

In order to check for efficacy of stress induction and accompanying changes in driver state through approaching speed requirements, the two scales ‘threat’ and ‘challenge’ from the PASA questionnaire [17] were used, which measure (cognitive) primary appraisal of stressful situations.

Since the present study aimed to explore the extent to which driver characteristics and driver state can predict turning maneuvers during the approach phase, dependent variables were selected that focused on longitudinal driving behavior as no lateral movements are to be expected to indicate turning maneuvers before entering the intersection. The dependent variables of interest in the present study are speed, speed profiles, and therewith-associated values for deceleration or acceleration, and—pertaining to these variables—pedal activity (e.g. releasing of acceleration pedal, pressing of braking pedal, or length of no-pedal-activity). The latter analysis can also serve as assessment of the general predictive potential for predicting turning maneuvers.

Table 1 lists a summary of the investigated hypotheses. The first hypothesis is exploratory, while hypotheses two to four relate to different phases of pedal activity (H2—releasing accelerator pedal, H3—no pedal-activity, and H4—pressing brake pedal).

Table 1 Summary of hypotheses

H1:	It is expected that drivers with high ED-scores intend generally to drive with higher speed and/or try to maintain speed longer and will decrease at lower time-to-intersection (TTI)/distance-to-intersection (DTI) compared to drivers with low ED-scores
H2:	Drivers with high ED-scores control actively their speed longer by means of pedal activity, and thereby will release the accelerator pedal at lower TTI/DTI (a), and will drive faster at this position (b)
H3:	Drivers with high ED-scores will avoid longer phases without pedal activity, thereby duration or driven distance without pedal activity will be shorter (a), and the loss of speed will be lower (b)
H4:	When pressing the brake pedal the position of TTI or DTI will be lower for drivers with high ED-scores (a), and the corresponding speed at this position will be higher (b). Speed values when entering the sequence of braking are expected to be higher (c). Consequently, speed loss is expected to be higher (d), as high-ED drivers need to decelerate more to reach a comfortable speed level for turning. Based on the general necessity for decreasing speed before turning maneuvers (contrary to the intention of drivers with high ED-score to maintain speed), maximum values of deceleration values (e) are expected to be higher before entering the intersection
H5:	Due to the intention of maintaining higher speeds, drivers with high ED-scores will experience greater lateral acceleration while turning
H6:	Due to the intention of maintaining higher speeds, it is expected that drivers with high ED-scores will show greater maximum acceleration values after performing the turning maneuver to reach faster preferred levels of higher speed

2.4 Driving Scenario

The study was conducted on a closed-off test track (see Figs. 1, 2 and 3). The route was designed with the same four rounds: approaching the intersection from a western starting point (1), right turn maneuver (2), following the circuit and re-approaching the intersection from east (3), left turn maneuver (4), following the circuit and re-approaching the intersection from east (5), and finally going straight over the intersection (6) back to the starting point.

Fig. 1 View from north-west (right turn maneuver before scenery-vehicle)

Fig. 2 View from east (left turn maneuver behind scenery-vehicle)

Fig. 3 Satellite view on test track (*Origin* Google Earth). Free area in the *middle* and *curvy roads* on the left side were used for familiarization

2.5 Procedure

Upon arrival, participants completed questionnaires pertaining to demographics, personality (BFI-10 [18]), sensation seeking (BSSS [19]) and traffic-related locus of control (T-LOC [20]). A cover story ("Adjusting of vehicle chassis for urban intersections") was used in order to avoid undue influence of participants' expectations. Participants familiarized themselves with the vehicle and test track circuit by performing a series of basic driving maneuvers (cautious, as well as strong acceleration and deceleration; driving through a constricted alley, slalom course, and curves; left and right turning maneuvers). The experimenter was seated on the right rear seat of the vehicle.

Each run began at the same starting point, at which participants were instructed regarding the approaching speed requirements. After the first of each approaching speed condition, participants completed the PASA questionnaire in order to evaluate the experience of situationally-induced stress. At the end of the session, participants were informed about the true purpose of the study and received monetary compensation for participation. Overall, the experiment took each participant about 45 min.

2.6 Data Processing

All data from CAN-interface GPS-data, and questionnaires were stored in a PostgreSQL database [21]. To access the data from the database for further processing, a comprehensive framework was programmed in R. Approximate "arrival measures" with respect to DTI and TTI were computed in three steps: (1) Detecting the GPS-coordinates for (artificial) positions-of-interest as reference points (here: latest meaningful position before entering the intersection, likely to stopping line), (2) computing the linear (spherical) distance between these coordinates and the actual GPS-position of the driver at each timestamp (using R package 'geosphere'), and (3) computing the arrival measure by subtracting the driven distance/time from the driven distance/time when reaching the minimum GPS-distance from step 2.

2.7 Manipulation Check

In order to ascertain whether changes in approaching speed requirements affected stress levels, PASA reliability and scores were examined for each speed condition. After excluding item two of the PASA questionnaire ("This situation is important to me") due to response ambiguity, Cronbach's alpha indicated high reliability (0.72) in runs with approaching speed of 70 km/h, and moderate reliability in the 50 km/h condition (0.6) (for interpretation of Cronbach's alpha according to common practice see e.g. [22]). Means for scales 'threat' and 'challenge', and scores for primary appraisal were significantly higher in conditions with high approaching speed to the intersection (see Table 2). As there was no sequential effect of driving with 50 or 70 km/h at first, the results indicate that stress induction was successful.

Table 2 Scale characteristics for primary appraisal of PASA

	M_{50kmh} (*SD*)	M_{70kmh} (*SD*)	$t_{df\,=\,39}$	*p*	*d*
Threat	4.8 (1.65)	6.03 (3.0)	−2.37	.011	−0.37
Challenge[a]	6.1 (3.22)	7.5 (3.1)	−1.93	.03	−0.31
Primary Appraisal[a]	10.9 (4.09)	13.53 (5.06)	−2.61	.006	−0.41

[a]Without item two

3 Results

Present analyses focus on the relationship between VIP and vehicle data, while analyses of the questionnaire for driving style, BFI-10, BSSQ or T-LOC are not reported at this point. In the following, participants with high or low scores in VIP subscales will be labeled according to their initials, e.g. high ED or low US. For the investigation of approaching behavior with respect to group assignments, objective data and subjective data were merged. All variables have been tested for univariate outliers using z-scores according to the thresholds proposed by [23]. To test for statistical significance of dependent variables (dv), linear mixed models (see Eq. 1) were performed using R-packages afex, lsmeans and multcomp. Tests were computed with an alpha of 5 %.

$$dv \sim VIP * apprspeed * time + Error(id/(apprspeed * time)). \qquad (1)$$

All tests were performed using both TTI and DTI. Data from maneuvers of left and right turning were separated (no going straight maneuvers were analyzed at this point). For analyses and preceding visual exploration of key parameters of pedal activity, sequential data were computed and sorted (using R packages ‘TraMineR’, and ‘TraMineRextras’) by minimum position of braking activity, see Fig. 4. All

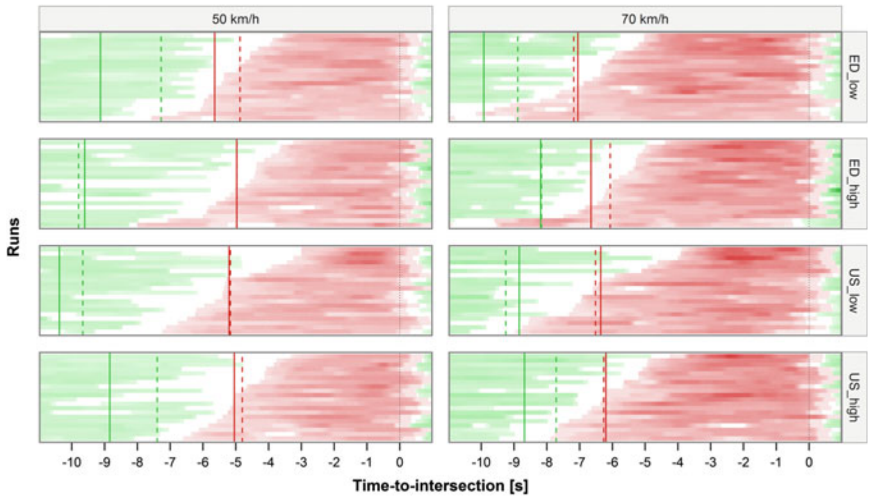


Fig. 4 Pedal sequences for right turning. *Rows* represent runs. *Colours* represent type of pedal activity (*green*—acceleration, *red*—braking, *white*—no pedal activity). Gradient represents intensity of pedal activity (high gradient–high intensity). *Coloured, vertical lines* represent first (*dashed*) or second (*solid*) run in blocks of approaching speed

Table 3 Model statistics for VIP from linear mixed models

	DTI				TTI			
	Left		Right		Left		Right	
	F	*p*	*F*	*p*	*F*	*p*	*F)*	*p*
accprel_ pos	2.08	.12	2.08	.12	3.00	.05**	0.57	.64
accprel_speed	2.62	.07**	3.88	.02*	2.37	.09**	3.55	.03*
wopedal_length	1.47	.25	1.29	.3	1.37	.27	1.53	.23
wopedal_speedloss	0.24	.87	0.72	.55	0.55	.65	1.29	.3
brake_pos	1.19	.33	0.65	.59	1.54	.22	0.92	.44
brake_length	0.29	.83	0.61	.61	0.4	.76	0.78	.51
brake_speed	1.91	.15	1.24	.31	1.91	.15	1.23	.32
brake_speedloss	0.98	.42	0.34	.8	0.97	.42	0.57	.64
max_decel	1.88	.15	0.62	.61	1.88	.15	0.62	.61
max_latacc	12.10	<.001*	5.80	.003*	12.13	<.001*	6.72	.001*
max_lonacc	1.84	.16	2.33	.09**	1.77	.17	2.43	.09**

*$p < 0.05$
**$p < 0.1$

visualizations were compiled using R-package ‘ggplot2’. A summary of statistical key values for the model component ‘VIP’ can be seen in Table 3. Due to outliers, within-group degrees of freedom vary between 25 and 35.

3.1 Speed Profiles (H1)

As this hypothesis was intended for exploration, no differentiation between consecutive runs of the same participants and no additional adjustments for t-testing (e.g. repeated measures) between VIP groups were performed. There were mostly no deviations from normal distribution (tested using Shapiro-Wilk test) throughout the trend of speed data against DTI, but in speed data against TTI. As no significant differences between speed profiles of groups with low or high scores in US could be found, no further analyses were performed here in the context of exploration.

Participants with high scores in ED drove slightly faster, particularly when approaching the intersection followed by a left turning maneuver. A significance level of 5 % was surpassed at DTIs between −10.1 and −3.3 m before reaching the (artificial) stopping line. Corresponding values for TTI can be seen in Fig. 5.

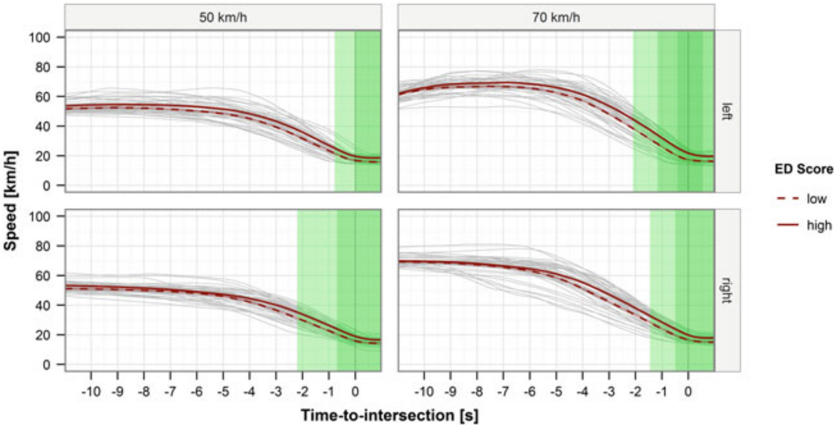


Fig. 5 Speed profiles and areas of significance ($\sim$ *green*, saturation $\sim$ p < [0.1; 0.05; 0.01])

3.2 Pedal Activity (H2–H4)

The results of the tests of hypotheses H2 to H4 are summarized in the following:

H2:	(a) No significant differences in positions of releasing acceleration pedal (labelled as 'accprel_pod') could be found in terms of DTI in left and right turning maneuvers. In terms of TTI, values almost reach significance in left turning maneuvers
	(b) Speed values (labelles as 'accprel_speed') between groups were significantly different when approaching intersection with 70 km/h and turning right (see right column in Fig. 5), slightly significant in cases of left turning
H3:	(a) In phases without pedal activity (between releasing accelerator pedal and pressing brake pedal), there were no significant differences in length of phases (labelled as 'wopedal_length') measured as time (using TTI) and driven distance (using DTI)
	(b) There were also no significant differences in speed loss (labelled as 'wopedal_speedloss')
H4:	(a) There were no significant differences between minimum positions of pressing brake pedal (labelled as 'brake_pos')
	(b) There were no significant differences in length of braking activity (labelled as 'brake_length')
	(c) There were no significant differences in speed values (labelled as 'brake_speed') at minimum position of pressing brake pedal
	(d) There were no significant differences in speed loss (labelled as 'brake_speedloss')
	(e) There were no significant differences in maximum deceleration values

3.3 Driving Through Intersection and Re-Accelerating (H5–H6)

In the following, the results of the analyses pertaining to hypotheses H5 and H6 are summarized:

H5:	In all conditions (left and right turning) related to both measures DTI and TTI (see Fig. 6) values of maximum lateral acceleration are significantly different between groups
H6:	No significant differences in maximum longitudinal acceleration could be found in left turn maneuvers. In right turn maneuvers (see Fig. 7) values almost reach significance, in relation to measures of both DTI and TTI

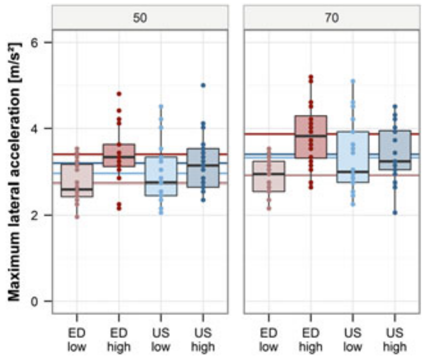


Fig. 6 Maximum lateral acceleration while turning left (means—*coloured lines* in background)

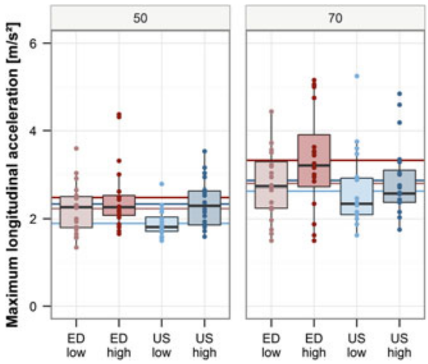


Fig. 7 Maximum longitudinal acceleration after turning right (means—*coloured lines* in background)

4 Discussion

The present study focused on the effect of driver characteristics and driver state on different parameters when approaching intersections and their potential of mitigating prediction algorithms. Most hypotheses regarding driver operations underlying to speed behavior were not confirmed by the empirical data. Only values for speed when releasing the accelerator pedal, values for maximum lateral acceleration while turning ($p < 0.001$) and values for maximum longitudinal acceleration after turning ($p < 0.1$) show significant differences. The last two variables represent essential parameters to create speed models for the used prediction model of [11]. Therefore, it can be stated that conventional parameters used in prediction algorithms for turning maneuvers can be significantly differentiated between groups of drivers based on driver characteristics as indicated by the VIP.

Regarding the initial research question ("Is there a need for individual parametrization of prediction algorithms?"), the present statistical results do not indicate that the consideration of detailed individual differentiations is necessary, but relating to the essential parameters of [11], there were significant differences in maximum longitudinal and lateral acceleration. Thus, in future work, it is planned to extend the model (Bayesian network) of [11] by additional nodes (e.g. VIP), as is suggested in Fig. 8. The integration of knowledge about individual characteristics is also supported by the grouping results of exploratory cluster analysis of speed profiles (Figs. 9 and 10), which shows good differentiation between ED. Regarding driver state more elaboration is necessary.

Furthermore, the present article elaborates sophisticated methods of analyzing and visually exploring driving behavior in complex driving situations using colored sequence tiles and gradient, which allows for differentiated insights into data and could be transferred to the analyses of other critical situations in urban areas. Additionally, driving behavior was dissected in different phases on the basis of elementary driver operations such as pedal activity. For example, visualizations

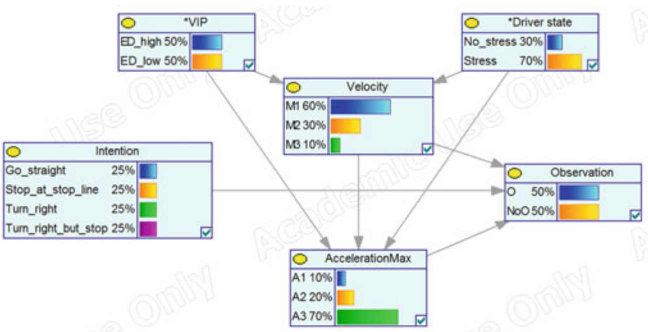


Fig. 8 Extended prediction model (built with GeNIe by BayesFusion). New nodes (*) for 'VIP' (as measure for driver behavior) and 'Driver state' added

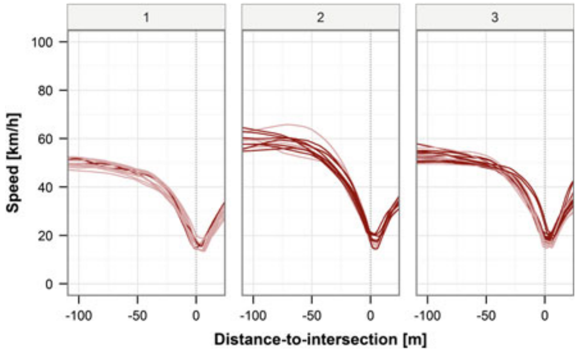


Fig. 9 Clustered groups (using k-means) for left turning and approaching speed of 50 kmh. Most speed profiles of drivers with low ED-scores were assigned to group 1

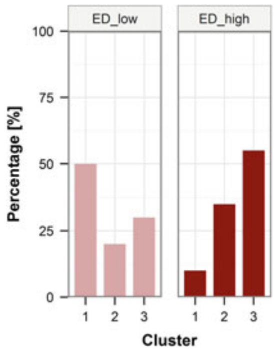


Fig. 10 Percentage of speed profiles in each cluster group for each ED score

from present data show (see Fig. 4), that approaching intersections begin with cautious release of the accelerator pedal (0), up to the complete release of the accelerator pedal (1). This is followed by a phase without pedal activity (2), and pressing the braking pedal (3) until reaching the intersection. This basic (and plausible) sequence might hold true in situations without preceding vehicles or with distant preceding vehicles as well as high familiarity with the route and thereby habituation to comfortable speed according to individual preferences. Otherwise, the number of changes in pedal activity are expected to be higher and the length of phases without pedal activity to be shorter. Also, lack of familiarity with the route might lead to early releasing of the brake pedal to allow for coasting at lower than usual speeds.

In conclusion, statistical analyses only partially show significant differences in underlying operations for "designing" driving behavior, but machine learning techniques can take complete profiles into account and even integrate into algorithms, thereby enhancing the prediction model with psychological aspects and thus promote new insights.

Acknowledgments Research was conducted in the context of the project "Recognition of driver intention—prediction of driver behavior" as part of the German research initiative UR:BAN (http://urban-online.org). UR:BAN is supported by the German Federal Ministry of Economics and Technology on the basis of a decision by the German Bundestag.

References

1. Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C., Winner, H.: Three decades of driver assistance systems: review and future perspectives. IEEE Intell. Transport. Syst. Mag. **6** (4), 6–22 (2014)
2. LSE Cities, LSE Cities Report: July 2012–September 2014. London: LSE Cities, London School of Economics and Political Science (2014)
3. Manstetten, D., Bengler, K., Busch, F., Färber, B., Lehsing, C., Neukum, A., Petermann-Stock, I., Schendzielorz, T.: "UR:BAN MV human factors in traffic"—a German research project to increase safety in urban traffic. In: 20th World Congress on Intelligent Transport Systems (ITS World Congress), pp. 596–605, Tokyo, Japan, 14–18 Oct 2013. Curran, Red Hook, NY (2014)
4. Charlton, S.G.: Restricting intersection visibility to reduce approach speeds. Accid. Anal. Prev. **35**(5), 817–823 (2003)
5. 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 Traffic Psychol. Behav. **18**, 47–57 (2013)
6. Diederichs, F., Pöhler, G.: Driving maneuver prediction based on driver behavior observation. In: 2014 AHFE International, 5th International Conference on Applied Human Factors and Ergonomics, 19–23 July 2014, Kraków, Poland (2014)
7. Statistisches Bundesamt (Destatis): Verkehrsunfälle - Fachserie 8 Reihe 7 – 2014. Wiesbaden (2015)
8. Werneke, J., Vollrath, M.: What does the driver look at? The influence of intersection characteristics on attention allocation and driving behavior. Accid. Anal. Prev. **45**, 610–619 (2012)
9. Graichen, M., Färber, B.: Approaching urban intersections: visibility, glance behavior and potential analysis for predicting turning maneuvers. In: IEEE Intelligent Vehicles Symposium (IV): Gothenburg, Sweden, June 19–22 2016. IEEE (2016)
10. Liebner, M., Klanner, F.: Driver intent inference and risk assessment. 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. Springer International Publishing, Cham (2016)
11. Liebner, M., Klanner, F., Baumann, M., Ruhhammer, C., Stiller, C.: Velocity-based driver intent inference at urban intersections in the presence of preceding vehicles. IEEE Intell. Transp. Syst. Mag. **5**(2), 10–21 (2013)
12. OECD/ECMT, Young Drivers: The Road to Safety. OECD Publishing and European Conference of Ministers of Transport, Paris (2006)
13. Ahie, L.M., Charlton, S.G., Starkey, N.J.: The role of preference in speed choice. Transp. Res. Part F Traffic Psychol. Behav. **30**, 66–73 (2015)
14. Schmidt, L., Piringer, A.: Wiener Testsystem: Verkehrsspezifischer Itempool (Kurzbezeichnung VIP). SCHUHFRIED GmbH, Mödling (2012)
15. Arndt, S.: Evaluierung der Akzeptanz von Fahrerassistenzsystemen: Modell zum Kaufverhalten von Endkunden. VS Verlag für Sozialwissenschaften, Wiesbaden (2011)
16. Gstalter, H., Fastenmeier, W.: Reliability of drivers in urban intersections. Accid. Anal. Prev. **42**(1), 225–234 (2010)

17. Gaab, J.: PASA—primary appraisal secondary appraisal. Verhaltenstherapie **19**(2), 114–115 (2009)
18. Rammstedt, B., John, O.P.: Measuring personality in one minute or less: a 10-item short version of the big five inventory in English and German. J. Res. Pers. **41**(1), 203–212 (2007)
19. Hoyle, R.H., Stephenson, M.T., Palmgreen, P., Lorch, E.P., Donohew, R.: Reliability and validity of a brief measure of sensation seeking. Pers. Individ. Differ. **32**(3), 401–414 (2002)
20. Özkan, T., Lajunen, T.: Multidimensional traffic locus of control scale (T-LOC): factor structure and relationship to risky driving. Pers. Individ. Differ. **38**(3), 533–545 (2005)
21. Pereira, M., Lietz, H., Beggiato, M.: Development of a database for storage and analysis of behavioural data. In: Stevens, A., Krems, J., Brusque, C. (eds.) Driver Adaptation to Information and Assistance Systems, pp. 301–317. The Institution of Engineering and Technology (IET), London (2014)
22. Westland, J.C.: Structural Equation Models. Stud. Syst. Decis. Control **22**(5), 152 (2015)
23. Field, A.: Discovering statistics using IBM SPSS statistics: and sex and drugs and rock ‘n’ roll, 4th edn. Sage, Los Angeles (2013)

Using Adaptive Interfaces to Encourage Smart Driving and Their Effect on Driver Workload

Stewart Birrell, Mark Young, Neville Stanton and Paul Jennings

Abstract In-vehicle information systems (IVIS) aimed at supporting green driving have increased in both number and complexity over the past decade. However, this added information available to the driver raises significant ergonomic concerns for mental workload, distraction and ultimately driving task performance. Adaptive interfaces offer a potential solution to this problem. The Smart driving system evaluated in this study (which provided in-vehicle, real-time feedback to the driver on both green driving and safety related parameters via a Smartphone application) offers a comparatively simple workload algorithm, while offering complexity in its levels of adaptively on the display, with the theoretical aim to limit driver visual interaction and workload with the system during complex driving environments. Experimental results presented in this paper have shown that using the Smart driving system modulates workload towards manageable levels, by allowing an increase in driver workload when under low task demands (motorway and inter-urban driving) but not increasing workload when it is already at moderate levels (urban driving). Thus suggesting that any increase in workload can be integrated within the driving task using the spare attentional resource the driver has available.

Keywords Green driving · In-vehicle information systems (IVIS) · Mental workload · Adaptive interfaces

---

S. Birrell (✉) · P. Jennings
WMG, University of Warwick, Coventry, UK
e-mail: S.Birrell@Warwick.ac.uk

P. Jennings
e-mail: Paul.Jennings@Warwick.ac.uk

M. Young
Human-Centred Design Institute, School of Engineering and Design,
Brunel University, Uxbridge, UK
e-mail: M.Young@brunel.ac.uk

N. Stanton
Engineering and the Environment, University of Southampton, Southampton, UK
e-mail: N.Stanton@soton.ac.uk

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_3

1 Introduction

In-vehicle information systems (IVIS) have increased in number and complexity with the advancement of enhanced infotainment features, continuation of brand identity via reconfigurable displays and the connected car. Over the past decade in particular, such motivations have become more and more focused on the environmental and economic costs of road transport [1]. One way in which these costs of driving can be reduced is by adopting an 'eco-driving' style, with many manufacturers now offering in-vehicle information displays to provide feedback on such behaviors. Examples are the Ford Fusion Hybrid SmartGauge with EcoGuide, Honda Insight Eco Assist and the Nissan Leaf Eco Indicator (Fig. 1).

However, this added information available to the driver raises significant ergonomic concerns for mental workload, distraction and ultimately driving task performance. Meanwhile, road safety remains a high priority alongside these other concerns [2]. In an effort to better understand the potential impact that in-vehicle systems could have in improving safety and economy, a UK project called 'Foot-LITE' undertaken. The project aimed at developing a system to encourage 'Smart'—that is safe and environmentally friendly—driving behaviors. Foot-LITE differs from existing products offered by vehicle manufacturers in three important ways. First, it runs on a nomadic device (i.e., a Smartphone platform) rather than being original fit, collecting data wirelessly from the vehicle diagnostic systems and an adapted lane departure warning camera. Secondly, the system provides feedback not just on eco-driving principles, but also on safe driving maneuvers, and it attempts to balance its advice between the two, which may or may not be in conflict (see [1] for a discussion). Finally and of interest to this paper, Foot-LITE was designed to be an adaptive interface which varied the presentation of visual and audio information in accordance with (basic) workload measures. In addition the interface was developed and tested according to ergonomic design principles, and rigorously evaluated in simulator and on-road driving trials to evaluate if the desired behavior changes were achieved while avoiding negative consequences of workload or distraction. The next section briefly reviews the design development and the evidence for these claims, before going on to discuss how Foot-LITE might, in future, be more dynamically responsive to driver mental workload.

Fig. 1 Selection of vehicle manufacturers' in-vehicle Eco driving aids (*left to right* Ford Fusion, Honda Insight and Nissan Leaf)

2 Design Development

As stated above the Foot-LITE in-vehicle interface was developed using ergonomic principles, which as outlined in Fig. 2 originated with the completion of a cognitive work analysis (CWA) and state-of-the-art review, from this six design concepts were established and three were evaluated in a rapid prototyping study (RPS). Results from the RPS suggested that two designs be took forward to simulator testing, a more convention dashboard inspired display and a conceptual interfaced based on Ecological Interface Design (EID) principles. Both these designs were evaluated in a driving simulator before the EID interface was selected (based on subjective workload and increased peripheral detection response rate) and iterated based on participant feedback, until the final EID design was released for on-road trials (Fig. 3).

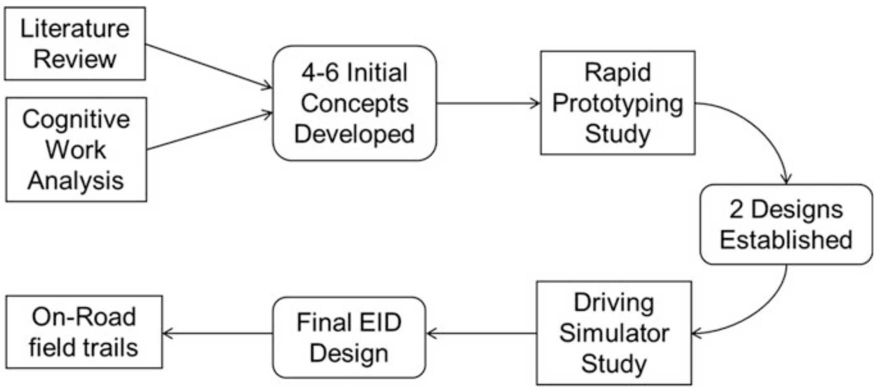


Fig. 2 Ergonomic development of the Foot-LITE interface

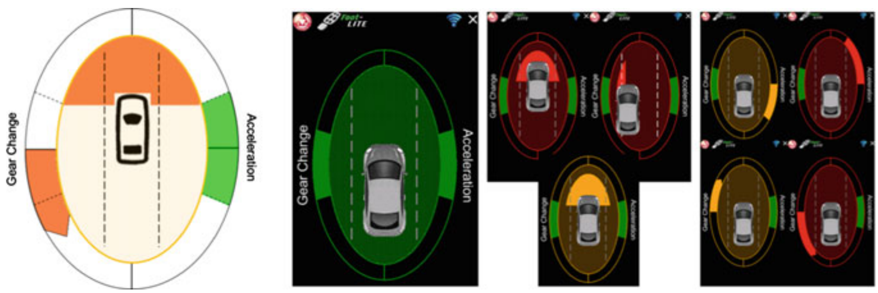


Fig. 3 Theoretical depiction of the EID interface (*left*) and version used for on-road trials (*right*)

2.1 Ecological Interface Design (EID)

In order to facilitate the design process, a Cognitive Work Analysis (CWA) [3] was previously conducted for the Foot-LITE project [4]. Based on the output of this CWA, a concept human-machine interface (HMI) was generated for the present study drawing on principles of Ecological Interface Design (EID) [5]. EID is an approach to interface design that was introduced specifically for complex socio-technical, real-time, and dynamic systems. It has been applied successfully within a number of work environments, including process control, nuclear, petrochemical, military and aviation domains [5].

The EID approach is heavily based on two concepts—firstly, the Abstraction Hierarchy (AH) of the CWA [6], and secondly, the Skill-Rule-Knowledge (SRK) taxonomy [7]. The AH decomposes the work domain into five levels (functional purpose, abstract function, generalized function, physical function, physical form) to establish what type of information should be displayed, as well as where, when and how it should be presented, and finally how to integrate pieces of information which need to be associated. Vicente and Rasmussen [8] suggest that an 'EID interface should not contribute to the difficulty of the task, and at the same time, it should support the entire range of activities that operators will be faced with.' (p. 589). Thus it has been argued that interfaces designed following the EID framework will reduce mental workload when dealing with unfamiliar or unanticipated events [9].

The AH completed previously in the project suggested several aspects of safe and eco-driving that should be represented on the display, such as headway, lane deviation and cornering speed for safety, complemented by engine speeds and acceleration forces for eco-driving [4]. Figure 3 shows both the theoretical depiction of the EID interface developed at Brunel University, and also the final version developed with project partners and released for on-road trials. The principal aspects of the interface are the vehicle inside the oval, which represents mainly safety parameters, and the dynamic scales in the outer oval, which largely reflect eco-driving properties. The oval concept was based on Gibson and Crooks [10] notion of the 'field of safe travel', which was noted as '… a spatial field but it is not fixed in physical space. The car is moving and the field moves with the car through space.' (p. 456). The oval shape grew out of an initial concept of a 'virtual glass of water', drawing on the apocryphal driving test in which the candidate must not spill the water placed on the dashboard. The edge of the glass, in this case, represents the boundaries of safe travel in Gibson's terminology; instead of a circle, this has been distorted into an oval to allow for the greater longitudinal movement of the car. On the EID display, the vehicle is dynamic inside the oval and illustrates the movements of the car within its safety envelope of headway and lane position in the real-world. Meanwhile, the eco-driving factors are presented in gauges on the outer ring. In both safety and eco-driving cases, the driver's goal is to maintain the car within a 'Green Zone' of performance (in the middle of the display), to optimize each set of parameters. In addition, there is the facility for 'pop-up' messages to be delivered to the driver for more specific feedback about positive or negative aspects of their driving.

It is by no coincidence that the elements of the display pertain to low-level, operational elements of vehicle control as opposed to tactical or strategic tasks in models of vehicle control [11]. Such operational tasks correspond to skill-based levels of processing in Rasmussen's [7] taxonomy, and as such impose little in terms of mental workload. Information about rule-based, tactical levels of driving is provided in the pop-up messages, while knowledge-based, strategic feedback is reserved for the off-line system in order to minimize potential workload and distraction a priori.

Nevertheless, valid concerns remain about the use of in-vehicle information systems and their potential to distract the driver [12, 13], and it is acknowledged that the range of real-world scenarios that drivers may encounter could present workload issues that would not revealed in simulator trials alone. In order for the Foot-LITE EID to anticipate and cope with this range of events, the possibility of a mental workload manager for the in-vehicle system was explored – in other words, an adaptive interface.

3 Mental Workload and In-Vehicle Information Systems (IVIS)

As mentioned earlier, mental workload can be a particular issue with IVIS systems. Driver overload with an additional task or interface in the vehicle can adversely affect performance [12, 14], particularly if workload is already high or if the driver has a lower capacity to respond. Studies have shown that while conducting a difficult cognitive task (such as math addition), drivers spend less time looking at areas in the peripheries (such as mirrors and instruments) and instead focus on looking centrally ahead [15]. Even though time looking outside of the vehicle remained unchanged, these results suggested a change in drivers' allocation of attention.

There is a view in the literature that drivers may have up to 50 % spare attentional capacity during normal driving [16], meaning that under ordinary circumstances there may be enough spare capacity to interact with IVIS displays. Green and Shah [17] suggest that the goal of the distraction mitigation system should be to keep the level of attention allocated to the driving task above the attentional requirements demanded by the current driving environment, and that during 'routine' driving approximately 40 % of attention could be allocated to non-driving tasks. The key issue is that mental workload is constantly varying during a drive, and it is crucial to maintain sufficient spare capacity to deal with unexpected or emergency scenarios. Within that constraint, it may be possible to provide additional driving-related information in real-time without detriment, as has been seen with previous Foot-LITE EID studies [18].

3.1 Adaptive Interfaces

The concept of adaptive interfaces has been around for some time [19, 20]. These theories were adapted for the GIDS project (Generic Intelligent Driver Support system) [21], and more recently been applied in the driving context [12, 22]. Typically, the adaptive interface will use sensors to detect some parameter of the task context, and will infer the user's mental workload based upon this information. The sensors may be monitoring the user's physiological state [19] or other overt behavioral indices [12], or alternatively they may be measuring dynamic characteristics of the task in question. The interface itself then adapts by providing more or less information depending upon the outcome of the workload calculation.

The current Foot-LITE system has no capacity for monitoring the driver directly, but it does include a number of sensors related to the driving task, including GPS data, a forward-looking camera with object recognition, and numerous vehicle parameters from the on-board diagnostics. Thus it would be quite plausible to build a task-based driver workload model upon which to base the adaptive functions of the interface.

The literature on driver mental workload offers several task-related indicators of workload. Factors of the environment, such as traffic and road situation, as well as different elements of the driving task (e.g. vehicle control and guidance, navigation) can influence mental workload. For instance, steering appears to be a significant source of workload in vehicle control [23], while tuning a car radio or using a navigation system are amongst the most demanding of the conventional in-car tasks [24]. In terms of driving maneuvers, it is known that workload increases during a turn [25], particularly when emerging from a junction when the driver has to cross a lane of traffic. Mental workload also increases in towns and cities when compared to motorway or rural driving, due to the unpredictable nature of the former [26]. These high workload situations are also associated with accident involvement. This has led to the idea of constructing mental load maps of towns, in order to predict accident rates and so design appropriate interventions [27].

Indeed, this is the kind of approach taken by [22] with their adaptive interface. Situational factors were detected by an on-board geographical database, and a computational workload estimator bases its decisions on the assessment of those situations. Such situations included road type (urban, rural etc.), curvature, slope, junctions, and directions. If workload was deemed to have exceeded a set threshold, then incoming telephone calls are routed directly to voicemail without informing the driver. This adaptive interface showed promising results in terms of managing driver mental workload. Indeed, a similar system has recently been marketed by Volvo cars [13].

Whilst the underlying algorithms of the workload estimator are quite complex in the Piechulla et al. [22] demonstrator, the implementation of the adaptive interface is rather more straightforward. In contrast, the proposed Foot-LITE adaptive interface is somewhat more rudimentary in terms of its workload algorithms, but offers more complexity in its levels of adaptivity on the display.

The literature reviewed above was drawn upon in combination with pragmatic considerations about what parameters could be measured and other factors such as driving standards, to propose a set of rules for the adaptive interface. In our conception, driver mental workload has three levels:

- High—mental workload is deemed high when driving on urban roads with a high density of junctions (i.e., probably in a city or town center). Speed limits of these roads may be between 20 and 40 mph, but actual speeds will probably be around 0–25 mph. The drive is characterized by numerous stop/starts, frequent turns, or highly inconsistent speed profiles.
- Medium—medium mental workload situations may still be in an urban or inter-urban setting but in the absence of many junctions, and with fewer stops and turns. Speed limits are likely to be 30 or 40 mph, with probable speed ranges of 20–40 mph. The drive is characterized by lower mean driving speeds but more consistency in speed profiles.
- Low—on roads with speed limits of 50 mph or over, with relatively consistent speeds of approximately 45 mph and over, low junction density and low numbers of turns. This category also incorporates any extra-urban road of national speed limit, motorways and dual carriageways.

In addition to this categorization, there is provision for mental workload to 'step up' a level if any of the following factors are present: high degrees of speed inconsistency, increased number or magnitude of turns, or high densities of junctions.

With the rules for mental workload levels derived, the next step is to determine how these affect the adaptive nature of the display. Recall that the EID HMI has several components—the inner oval for safety related information (headway, lane departures), the outer oval for eco-driving feedback (acceleration, gear changing), and the pop-up messages for specific, event-related feedback. Furthermore, within the oval, the status of the eco-driving and safety parameters can be either green, amber or red. Each of these elements can be independently enabled or disabled on the display, providing various combinations of levels of information available to the driver.

There was wide acceptance within the Foot-LITE project that, in the event of any conflicts in advice from safety or eco-driving perspectives, the safety-related information should always take precedence. With that in mind, it was determined that the feedback provided at each level of workload should be as follows:

- High—only 'red' safety and eco warnings to be given; audio, amber warnings and pop-up feedback are disabled
- Medium—all safety warnings active (red and amber), only red eco-driving feedback is given, audio is active, pop-ups are presented if deemed safe to do so following an additional set of pop-up implementation rules (essentially if the driver is stopped or at a steady speed and not engaging in a maneuver)
- Low—all information active

Table 1 Summary of rules and algorithms for Foot-LITE adaptive interface

		MWL		
		High	Medium	Low
Road	Speed limit	20, 30 or 40	30 or 40	50+
Geometry	Probable speed	0–25	20–40	45+
	Actual speed	<40	<40	>40
	Road type	Urban	Urban/inter	Inter/extra
	Junctions	Yes	No	No
	Speed variability	High	Low	Low
Advice	Safety	Red	All	All
Given	Eco	Red	Red	All
	Audio	No	Yes	Yes
	Pop-ups	No	If safe	Yes

It is worth noting that in the high and medium workload categories, the only safety-relevant information which can be presented pertains to headway, since the lane departure warning camera is not active at speeds below 40 mph for separate technical reasons. The criteria for, and adaptive conditions of, the interface are summarized in Table 1.

4 Evaluation of Driver Workload Through the Design Process

From conception to evaluation a rigorous ergonomic procedure for the design of the in-vehicle Foot-LITE HMI was adopted, Fig. 2 outlines this process. The following section will discuss subjective workload findings (as assessed by the NASA-Task Load Index, or TLX [28] as a result of using the Foot-LITE Smart driving EID HMI in both simulated and real world driving. Statistical testing was conducted using SPSS 21.0 for Windows and significance determined using the Friedman and Wilcoxon Signed Rank tests.

4.1 Simulator Study

The first dynamic evaluation of the Foot-LITE adaptive interface was conducted using the Brunel University driving simulator, with results published regarding the selection between the conventional dashboard and EID interface in a previously [29]. Included in this paper is a summary of the methodology and previously unreported data on TLX ratings for the EID adaptive interface.

Table 2 Mean and sub-scale TLX ratings for a Control condition and when using the Foot-LITE EID HMI, in both Urban and Inter-Urban simulated driving

Mean	Urban		Inter-urban	
	Control	Foot-LITE	Control	Foot-LITE
Mental	60.2	57.8	47.4	47.7
Physical	33.0	31.6	23.0	25.4
Temporal	48.2	47.8	28.2	35.4
Performance	54.6	48.6	49.4	44.8
Effort	58.2	55.6	45.0	42.3
Frustration	46.0	43.6	34.8	33.3
TLX Rating	50.0	47.5	38.0	38.2

Twenty-five participants drove two different, 5-min, simulated scenarios. The first was an 'Urban' drive though a cityscape with numerous traffic light controlled intersections and other road users to interact with; the second a dual carriageway route with little traffic, termed 'Inter-Urban'. Two conditions were adopted, one a control (no feedback) the other with smart driving feedback being offered to the diver in real-time, in the vehicle via the Foot-LITE EID adaptive interface. The TLX was recorded after each of the randomly assigned condition and scenario combination. Results presented in Table 2 showed no significant ($p > 0.05$) difference between mean TLX ratings when using the Foot-LITE Smart driving advisor compared to the control when driving in either scenario. Interestingly, although not significant, a strong trend ($p < 0.1$) was observed for 'Own Performance' to be rated more preferably (i.e. lower TLX rating for a better self-rating of driving performance) in both urban and inter-urban driving when using the Smart driving advisor. This may be as a result of the positive feedback which was given to users via the default green display on the Foot-LITE HMI; either 'rewarding' them for corrective action taken to sub-optimal driving behaviors, or the positive reinforcement for already good performance.

Conclusions from the simulator study suggest that presenting the driver with real-time information on driving performance did not increase workload. It is not possible to suggest if this is directly a result of the adaptive nature of the interface, i.e. not increasing workload above 'manageable' levels, or confounding factors such as simulated driving. However this was a positive outcome for the interface, and suggests that system could be considered for further investigation on-road.

4.2 On-Road Field Trials

As part of an EU funded project called TeleFOT on-road field trials were conducted using the Foot-LITE system. Again results relating to driving performance [18] and glance behavior [30] have been published previously; this paper focuses on the previously unreported TLX ratings collected.

Forty drivers drove an instrumented vehicle, equipped with data loggers and the Foot-LITE system, around a mixed driving route being 40.1 miles (or 64.5 km) in length and taking approximately 1 h and 15 min to complete. The scenario encompassed three clearly defined sections of roadway, each taking a similar length of time to complete and only including one type of road category—Motorway, Urban and Inter-Urban. Two counter-balanced conditions were adopted, one as a control with no Smart driving feedback offered, the other where the Foot-LITE system was active. In order to obtain the TLX ratings for each individual section of roadway three predefined points were selected where participants were instructed to pull over and complete the questionnaire. These points were immediately after the motorway, urban and inter-urban sections allowing for only that specific category of roadway to be considered in the TLX ratings. This avoids common issues associated with questionnaire completion following a lengthy task, i.e. respondents express opinions on a generic combination of the route, or more typically what participants experienced in the preceding 5 min. This provides an accurate and reliable measure for workload ratings obtained in this study.

Results presented in Fig. 4 show that when using the Foot-LITE system there was a significant increase in subjective workload over the control condition when driving in the motorway ($Z = -3.69$, $p < 0.01$) and inter-urban ($Z = -3.58$, $p < 0.01$) driving scenarios; although no difference was observed with urban driving ($p > 0.05$). When considering the control condition only (as a baseline for normal driving), the urban scenario was rated significantly higher workload compared to both motorway ($Z = -4.24$, $p < 0.001$) and inter-urban driving ($Z = -4.75$, $p < 0.001$). These TLX results could be considered suitable for use by

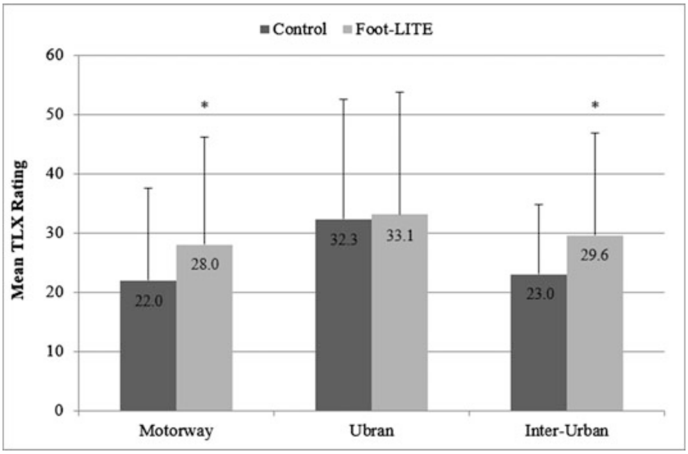


Fig. 4 Mean TLX rating for all participants across each road category, error bars represent standard deviation. *Asterisks* (*) indicates significant ($p < 0.01$) difference to the control

future research as a reference for real-world driving workload, as each road category was clearly defined and TLX recorded immediately following the completion of each category. This ensured accuracy and reliability of data.

Summarizing results from the on-road field operational trials shows that TLX rated workload increased significantly when using the Smart driving system during motorway and inter-urban driving, but not with urban driving. Additional research also showed that using the Foot-LITE system resulted in significant positive changes in driver behaviors, namely an increase in mean headway (distance to vehicle in front) and a 4.1 % increase in fuel efficiency [18]. As highlighted in Table 1, the adaptive nature of the Foot-LITE interface resulted in the driver being presented with a limited set of feedback when driving in the urban environment (or at speeds below 25 mph). The theoretical aim of which was to limit driver visual interaction and mental workload with the system during complex driving environments. This aim could be considered to be achieved with the Foot-LITE adaptive interface as there was no increase in subjective mental workload (as assessed using TLX) with urban driving.

The concept that drivers may have up to 50 % 'spare' attentional capacity to allocate to non-driving tasks during normal or routine driving [16, 17] was key to the design of the Foot-LITE smart driving interface, with it being designed to allow more information to be presented to the driver at times of low workload (motorway and inter-urban driving) but limiting this during high workload driving (urban). Also observed was a modulation towards manageable workloads with the Smart driving system, as there was no significant difference in workload observed when driving in the motorway (28.0), inter-urban (29.6) and urban (33.1) scenarios (Fig. 4). What the adaptive nature of the Foot-LITE EID HMI was hoping to achieve was that this increase would be easily integrated within the driving task using the spare attentional resource the driver has available–results from the on-road trials support this notion.

5 Conclusions and Future Directions

The proposed adaptive interface for the Foot-LITE system presented here has been based on first principles regarding driver mental workload, taking into account engineering constraints of the technology. Whilst resource limitations in the project have precluded any further testing or development of the adaptive nature of the interface thus far, the rudimentary algorithms do have intuitive appeal and are consistent with existing knowledge (and, indeed, other applications–e.g. [22]) of driver mental workload. Moreover, the prioritization rules for the display are ultimately founded on the skill-rule-knowledge elements of driving derived from the CWA, and the adaptive interface merely extrapolates these in a dynamic fashion. Thus, by limiting feedback during high workload only to skill-based, safety-critical tasks, the intention is to optimize both the beneficial effects on driving performance as well as mitigating any consequences of distraction.

Several questions remain unanswered about adaptive interfaces, both here and in general, such as how to manage the feedback for the driver so that they know which level of the system is operating—otherwise it could lead to uncertainty and unpredictability. Similarly, how to handle transitions between levels, as well as transient peaks or troughs in workload, needs to be answered. If, for instance, a driver is operating in a 'low' workload context, then the information from the display may be occupying spare capacity that could be used to deal with unexpected or emergency events.

However, experimental results presented in this paper have shown that using the Foot-LITE system during driving modulates workload towards manageable levels, by allowing an increase in driver workload when under low task demands (motorway and inter-urban driving) but not increasing workload when it is already at moderate levels (urban driving). This could be considered a success for the adaptive nature of the mental workload model utilized by Foot-LITE, by limiting increases in workload when it was deemed unacceptably high. This study again highlights the potential for adaptive interfaces in eliciting positive changes to driving behavior in real-world driving environments, whist mitigating the negative impacts of driver distraction and increases in driver workload.

Acknowledgments The Smart driving system used in this study was developed as part of the Foot-LITE project which was sponsored by the Technology Strategy Board, Engineering and Physical Sciences Research Council and the UK Department for Transport. The TeleFOT project, for which these trials were conducted as part of by MIRA Ltd. in the UK, is an EU sponsored project under the Seventh Framework Programme. Additional support was offered by Innovate UK through the WMG centre High Value Manufacturing Catapult. Specific thanks go to Mark Fowkes and Nadia Shehata who assisted greatly with the study.

References

1. Young, M.S., Birrell, S.A., Stanton, N.A.: Safe driving in a green world: a review of driver performance benchmarks and technologies to support 'smart' driving. Appl. Ergon. **42**, 533–539 (2011)
2. PACTS: Beyond 2010—A Holistic Approach to Road Safety in Great Britain. Parliamentary Advisory Council for Transport Safety, London (2007)
3. Vicente, K.: Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. Lawrence Erlbaum Associates, Mahwah (1999)
4. Birrell, S., Young, M., Stanton, N., Jenkins, D.: Cognitive Work Analysis for safe and efficient driving. Theor. Iss. Ergon. Sci. **13**, 430–449 (2012)
5. Burns, C., Hajdukiewicz, J.: Ecological Interface Design. CRC Press, Boca Raton (2004)
6. Rasmussen, J.: The role of hierarchical knowledge representation in decision making and system management. IEEE Trans. Syst. Man Cybern. **15**, 234–243 (1985)
7. Rasmussen, J.: Skills, rules, knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Trans. Syst. Man Cybern. **13**, 257–266 (1983)
8. Vicente, K., Rasmussen, J.: Ecological interface design: theoretical foundations. IEEE Trans. Syst. Man Cybern. **22**, 589–606 (1992)

9. Vicente, K.: Ecological interface design: supporting operator adaptation, continuous learning, distributed, collaborative work. In: Proceedings of the Human Centered Processes Conference, pp. 93–97 (1999)
10. Gibson, J.J., Crooks, L.E.: A theoretical field-analysis of automobile driving. Am. J. Psychol. **51**, 453–471 (1938)
11. Ranney, T.A.: Models of driving behavior: a review of their evolution. Accid. Anal. Prev. **26**, 733–750 (1994)
12. Donmez, B., Boyle, L.N., Lee, J.D.: Safety implications of providing real-time feedback to distracted drivers. Accid. Anal. Prev. **39**, 581–590 (2007)
13. Broström, R., Engström, J., Agnvall, A., Markkula, G.: Towards the next generation intelligent driver information system (IDIS): the Volvo car interaction manager concept. In: 13th ITS World Congress, London, 8–12 Oct 2006 (2006)
14. Horberry, T., Anderson, J., Regan, M., Triggs, T., Brown, J.: Driver distraction: the effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. Accid. Anal. Prev. **38**, 185–191 (2006)
15. Harbluk, J., Noy, Y., Trbovich, P., Eizenman, M.: An on-road assessment of cognitive distraction: impacts on drivers' visual behaviour and braking performance. Accid. Anal. Prev. **39**, 372–379 (2007)
16. Hughes, P.K., Cole, B.L.: What attracts attention when driving? Ergonomics **29**, 377–391 (1986)
17. Green, P., Shah, R.: Task time and glance measures of the use of telematics: a tabular summary of the literature. In: Safety Vehicles Using Adaptive Interface Technology (SAVE-IT) Project, Literature Review, Task 6, December 2004. The University of Michigan Transportation Research Institute (2004)
18. Birrell, S.A., Fowkes, M., Jennings, P.: The effect of using an in-vehicle smart driving aid on real-world driver performance. IEEE Trans. Intell. Transp. Syst. **15**, 1801–1810 (2014)
19. Byrne, E.A., Parasuraman, R.: Psychophysiology and adaptive automation. Biol. Psychol. **42**, 249–268 (1996)
20. Hancock, P.A., Verwey, W.B.: Fatigue, workload and adaptive driver systems. Accid. Anal. Prev. **29**, 495–506 (1997)
21. Michon, J.A. (ed.): Generic Intelligent Driver Support: A Comprehensive Report on GIDS, p 252. Taylor & Francis, Bristol (1993)
22. Piechulla, W., Mayser, C., Gehrke, H., König, W.: Reducing drivers' mental workload by means of an adaptive man-machine interface. Transp. Res. Part F **6**, 233–248 (2003)
23. Young, M.S., Stanton, N.A.: Taking the load off: investigations of how adaptive cruise control affects mental workload. Ergonomics **47**, 1014–1035 (2004)
24. Dingus, T.A., Antin, J.F., Hulse, M.C., Wierwille, W.W.: Attentional demand of an automobile moving-map navigation system. Transp. Res. Part A **23**, 301–315 (1989)
25. Hancock, P.A., Wulf, G., Thom, D., Fassnacht, P.: Driver workload during differing driving manoeuvres. Accid. Anal. Prev. **22**, 281–290 (1990)
26. Harms, L.: Variation in drivers' cognitive load. Effects of driving through village areas and rural junctions. Ergonomics **34**, 151–160 (1991)
27. Wildervanck, C., Mulder, G., Michon, J.A.: Mapping mental load in car driving. Ergonomics **21**, 225–229 (1978)
28. Hart, S., Staveland, L.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P., Meshkati, N. (eds.) Human Metal Workload, pp. 139–183. Elsevier, Amsterdam (1988)
29. Birrell, S.A., Young, M.S.: The impact of smart driving aids on driving performance and driver distraction. Transp. Res. Part F **14**, 484–493 (2011)
30. Birrell, S.A., Fowkes, M.: Glance behaviours when using an in-vehicle smart driving aid: a real-world, on-road driving study. Transp. Res. Part F **22**, 113–125 (2014)

Is It Me or Is It You? Assessing the Influence of Individual and Organizational Factors on Safety Performance in the North American Railway Industry

Dylan Smibert and Mark Fleming

Abstract The recent Lac-Mégantic disaster highlighted the safety-critical nature of the railway industry. The public inquiry identified both individual and organizational failures as causal factors for the incident; yet, very few studies examine both constructs simultaneously. The authors examined the impact of safety climate and individual differences on self-reported safety performance and safety records. The study examined 306 railway workers employed by a North American Class I Railway (M = 5.8 years of service). Personality traits accounted for significant incremental variance over safety climate in self-report measures of safety compliance (ΔR^2 = 0.18), safety participation (ΔR^2 = 0.13), safety knowledge (ΔR^2 = 0.24), and safety motivation (ΔR^2 = 0.15). Based on the findings, employers may want to consider personality factors when selecting training, or identifying interventions for employees within safety critical occupations. Limitations include a relatively small industry specific sample; therefore, the results may not generalize to all transportation employees.

Keywords Railway transportation · Safety climate · Personality · Safety performance · Workplace accidents and injuries · Safety and selection

1 Introduction

The catastrophic Lac-Mégantic railway accident in July 2013 prompted the public, Government, and industry to re-evaluate the perceived safety of railway transportation. An inquiry report by the Transportation Safety Board (TSB) of Canada

D. Smibert (✉) · M. Fleming
Saint Mary's University, Halifax, Canada
e-mail: Dylan.Smibert@smu.ca

M. Fleming
e-mail: Mark.Fleming@smu.ca

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_4

identified 18 distinct causal factors contributing to the accident [1]. Although not directly categorized in the report, each of the 18- factors can be sorted into organizational, individual, and mechanical antecedents. For example, failing to secure the proper number of handbrakes can have a combination of *organizational*, *individual*, and *mechanical* antecedents. That is, the organizational culture may value efficiency, which would promote a pressure to engage fewer manual handbrakes, leading to less lag between employee shifts. The individual factor of risk perception [2, 3] suggests that employees calculate the risk of rule non-compliance differently, thus, could invoke a different behavior in the number of handbrakes applied. Finally, a mechanical factor may be that there is variation in the mechanism of the handbrake, some brakes may take 30 spins to engage while some may take 50 spins to engage and there is no functional indicator other than a pull-test (assuming the test is properly implemented). From a safety system's perspective, it is important to understand the influence of all factors and antecedents when trying to understand safety behavior [4, 5].

Over the last half-century, safety professionals have made a considerable effort towards adopting a systems approach to organizational safety [5, 6]. For the most part, academics and practitioners agree that there are often multiple factors involved in safety incidents, and there is little to no utility in placing fault solely on an individual [6]. Therefore, the goal of this research is not to support the antiquated concept of accident-proneness and risky employees [7, 8]. However, avoiding research on individual differences and their relation to safety has hindered, rather than helped, the progress of a systems approach to safety. An over-emphasis on organizational and engineering factors has created a blind-spot to a number of initiatives and interventions with an individual difference focus that may be helpful for employees in safety-critical occupations. To the best of the authors' knowledge, this will be the first study to examine the relative influence of both organizational factors (i.e., safety climate perceptions) and individual differences (i.e., personality traits) on safety performance.

1.1 Safety Climate

Safety climate refers to an employee's perception of their organization's commitment towards safety [9]. Many transportation companies use safety climate surveys to the evaluate perceived supervisor commitment to safety and to identify opportunities for improvement. Consistently, past research has supported the link between safety climate and safety performance [4, 5, 10, 11]. A meta-analysis on safety climate and organizational safety [11] found a significant relationship between safety performance indicators of compliance (ρ = 0.43) participation (ρ = 0.50), and accidents and injuries (ρ = 0.22) with safety climate. In our study, the relationship between safety climate and safety outcomes are not being contested; however, we aim to gain a better understanding of the strength of this relationship in the presence of other related factors, like personality traits.

1.2 Personality Traits

Personality traits are relatively stable over time, influence our thoughts and behaviors, and can be shaped, to a certain extent, by our environment [12]. Popularized by recent research, the “Big-Five” personality traits are one of the most commonly used models [13]. The model suggests that everyone varies in the intensity of each of the five traits of personality (Conscientiousness, Extroversion, Emotional Stability, Agreeableness, and Openness) [14, 15]. For example, Joey, may be very extroverted, conscientious, and emotionally stable, but not very open to new experiences or agreeable. Based on normative data for men and women in the general population, an individual will fall on a spectrum between high and low for each of the five traits [14].

Personality traits have been explored in relation to safety outcomes for decades but has recently fallen out of academic favor in light of climate and culture research. [11] A meta-analysis by Clarke and Robertson [16] summarized the relationship between occupational injuries and the Big-Five personality traits. Both Conscientiousness ($\rho = -0.27$) and Emotional Stability ($\rho = -0.26$) were significant predictors of organizational injuries and vehicular accidents. Extroversion has also been reported as a significant predictor of vehicular accidents ($\rho = 0.24$), but not organizational injuries. The next section describes and connects each personality trait to safety performance as identified by past research.

Conscientiousness. Someone with high-Conscientiousness is dutiful, driven by a strong will, competent, and methodological [14]. Individuals with a low level of Conscientiousness are often unprepared, impulsive, and careless in their work behaviors [14, 15]. Past research has found a relationship between individuals with low Conscientiousness and increased injuries [16–18]. Additionally, Glendon and colleagues [18] suggest that participation in safety-related behaviors may be predicted by conscientiousness.

Emotional Stability. Low Emotional Stability, often referred to as Neuroticism, describes someone that is sad, embarrassed easily, angry, fearful, and emotionally unstable [14]. Low levels of Emotional Stability can lead to a preoccupation with personal anxieties, increasing the likelihood of on-job distraction resulting in higher injury rates [19]. This personality trait has a relationship with both organizational and vehicular accidents ($\rho = 0.26$) [16]. Additionally, negative affectivity (the tendency to experience negative emotions) has also displayed a relationship to increased occupational injuries [20].

Extroversion. Someone with high-Extroversion is full of energy, optimistic, highly social, and seeks thrills and excitement [14]. Extroverts can have lower levels of vigilance; therefore, they will be less engaged in identifying hazards [21]. Sensation and thrill seeking have also been related to increased accident rates [18, 22].

Agreeableness. An individual with high-Agreeableness is often cooperative, pleasant, tolerant of others, and helpful [14]. Low-Agreeableness (aggressive, stubborn, dominant, and cynical) has been shown to be correlated with injury involvement [16, 23]. In contrast, individuals who are high in Agreeableness will likely be more cooperative when teamwork is required to work safely. Further, a certain level of cynicism and stubbornness may actually prove to reduce accidents and injuries in relation to refusing unsafe work and questioning the status quo. Further research is required in the area of Agreeableness and safety performance.

Openness to Experience. Someone high in Openness is imaginative, curious, broad-minded, and proficient problem solvers. Openness to experience is one of the least explored traits in relation to safety outcomes [16], but it has been found to be an important in job performance and training proficiency ($\rho = 0.25$) [18, 24].

1.3 Safety Performance

Burke et al. [25, p. 432] defined safety performance as "actions or behaviors that individuals exhibit in almost all jobs to promote the health and safety of workers, clients, the public, and the environment." By conceptualizing safety performance as behaviors rather than just low base rate outcome variables (accident and injury rates), researchers better identify psychological factors associated with safe behavior [5]. In this study, we measure safety knowledge, motivation, participation, and compliance.

This research supports a systems approach to safety, whereby individual differences in safety performance are a part of the larger system that includes job features and organizational features [4, 6, 26].

2 Method

2.1 Participants

The sample consisted of 306 railway workers employed by a North American Class I Railway. The employees ($M_{age} = 36.3$ $SD = 10.3$) were mostly male (97 %), Caucasian (83 %), and educated (51 % high school, 42 % university level). The sample consisted of qualified conductors ($N = 132$), engineers ($N = 24$), and conductors in training (N = 150).

2.2 Design

The employees were recruited to participate in a study on personality and workplace performance through an email invitation sent to their organizational email address. A total of 4400 employees received an email in July 2014 and had two weeks to complete the survey for a chance to win $500 in gift-card prizes. The participants were asked to complete 25-min study that consisted of a personality assessment in addition to a number of safety related questionnaires (see measures).

2.3 Measures

Safety Motivation Questionnaire. This questionnaire asks employees to assess their Safety Climate (α = 0.95), Safety Motivation (α = 0.79), Safety Knowledge (α = 0.66), Safety Compliance (α = 0.77), and Safety Participation (α = 0.77) on a 5-point scale. Each scale consists of 3-items. Safety Climate refers to the employee's perceptions of the manager's commitment to safety. Safety Participation refers to the participant's involvement in organizational safety programs and Safety Compliance captures an employee's adherence to following rules and procedures. Safety Knowledge measures an employee's knowledge of safe work practices and Safety Motivation refers to an employee's commitment towards safety. [27, 28]

International Personality Item Pool (IPIP). The IPIP-NEO-PI-R personality inventory was developed, validated and made available to the public domain. There are 10-items for each of the Big-Five personality traits for a total of 50-times. Sample items for Emotional Stability (α = 0.86) include "I rarely get irritated" and "I panic easily"; Extroversion (α = 0.86) includes items like "I am the life of the party", and "I don't talk a lot"; Openness (α = 0.82) includes the items "I have a vivid imagination", and "I do not like art"; Conscientiousness (α = 0.81) includes items like "I am always prepared" and "I do just enough work to get by"; and Agreeableness (α = 0.77) includes items like "I accept people as they are" and "I make people feel at ease". Each of the IPIP personality traits is highly correlates with the NEO-PI-R traits (ρ = 0.88–0.92) [13].

3 Results

A hierarchical multiple regression analysis was used to assess the incremental validity of personality traits over safety climate for safety participation, compliance, motivation, and knowledge. Evaluation of assumptions were assessed, and there were no significant concerns of univariate or multivariate outliers, linearity, normality, or homoscedasticity. Demographic characteristics of age and service were non-significant variables the regression analysis. The means and standard deviations of the personality traits and safety performance indicators are outlined in Table 1.

Table 1 Descriptive statistics

Instruments	*N*	*M*	*SD*
Safety climate	295	3.66	1.18
Agreeableness	306	3.99	0.47
Conscientiousness	306	4.11	0.48
Extroversion	306	3.37	0.56
Emotional stability	306	4.05	0.52
Openness	306	3.55	0.52
Safety compliance	295	4.46	0.51
Safety participation	295	4.02	0.69
Safety motivation	295	4.64	0.43
Safety knowledge	295	4.39	0.48

3.1 Hierarchical Multiple Regression

See Table 2.

Safety Motivation. In order to examine whether personality traits account for additional unique variance beyond Safety Climate, Safety Climate was placed in step one and the personality traits was placed into step two of the hierarchical regression. The Big-Five personality traits, $R^2 = 0.22$, $F(6, 288) = 14.89$, $p < 0.001$, accounted for significant incremental variance ($\Delta R^2 = 0.15$) in Safety Motivation scores above Safety Climate $R^2 = 0.09$, $F(1, 293) = 28.93$, $p < 0.001$, $\beta = 0.17$, $t(293) = 2.94$, $p < 0.001$, $S^2_{ri} = 0.02$. All personality traits were significant predictors of Safety Motivation except for Openness. Agreeableness, $\beta = 0.19$, $t(288) = 2.91$, $p < 0.01$, $S^2_{ri} = 0.02$, Conscientiousness, $\beta = 0.21$, $t(288) = 3.24$, $p < 0.001$, $S^2_{ri} = 0.03$, Extroversion, $\beta = 0.21$, $t(288) = 3.23$, $p < 0.001$, $S^2_{ri} = 0.03$, and Emotional Stability, $\beta = 0.18$, $t(288) = 2.59$, $p = 0.01$, $S^2_{ri} = 0.02$.

Table 2 Hierarchical multiple regression with standardized betas, squared semi-partial correlations, and R^2 change for safety motivation and safety knowledge

Safety predictors	Safety motivation			Safety knowledge		
	ΔR^2	β	r_{si}	ΔR^2	β	r_{si}
Step 1						
Safety climate	**0.09**[a]	**0.17**[b]	**0.02**	**0.05**[a]	0.06	0.00
Step 2						
Personality	**0.15**[a]			**0.24**[a]		
Agreeableness		**0.19**[b]	**0.02**		**0.21**[a]	**0.03**
Conscientiousness		**0.21**[a]	**0.03**		**0.27**[a]	**0.05**
Extroversion		**0.15**[a]	**0.03**		**0.26**[a]	**0.04**
Emotional stability		**0.18**[b]	**0.02**		0.13	–
Openness		0.11	–		0.08	–
Total R^2	**0.22**[a]			**0.28**[a]		

Note Significant ($p < 0.05$) beta weights are in bold. [a]$p < 0.001$, [b]$p < 0.01$, [c]$p < 0.05$

Safety Knowledge. The Big-Five personality traits, $R^2 = 0.28$, $F(6, 288) = 19.60$, $p < 0.001$, accounted for significant incremental variance ($\Delta R^2 = 0.24$) in Safety Knowledge scores above Safety Climate $R^2 = 0.05$, $F(1, 293) = 17.10$, $p < 0.001$, $\beta = 0.06$, $t(293) = 1.09$, $p > 0.05$, $S^2_{ri} = 0.00$, which was not a significant predictor in the presence of the personality traits. All personality traits were significant predictors except for Emotional Stability and Openness. Agreeableness, $\beta = 0.21$, $t(288) = 3.42$, $p < 0.001$, $S^2_{ri} = 0.03$, Conscientiousness, $\beta = 0.27$, $t(288) = 4.36$, $p < 0.001$, $S^2_{ri} = 0.05$, and Extroversion, $\beta = 0.26$, $t(288) = 4.18$, $p < 0.001$, $S^2_{ri} = 0.04$.

Safety Participation. The Big-Five personality traits, $R^2 = 0.25$, $F(6, 288) = 17.47$, $p < 0.001$, accounted for significant incremental variance ($\Delta R^2 = 0.13$) in Safety Participation scores above Safety Climate $R^2 = 0.14$, $F(1, 293) = 48.38$, $p < 0.001$, $\beta = 0.27$, $t(293) = 4.84$, $p < 0.001$, $S^2_{ri} = 0.06$. All personality traits were significant predictors of Safety Participation except for Openness. Agreeableness, $\beta = 0.13$, $t(288) = 2.12$, $p < 0.05$, $S^2_{ri} = 0.01$, Conscientiousness, $\beta = 0.20$, $t(288) = 3.21$, $p < 0.001$, $S^2_{ri} = 0.03$, Extroversion, $\beta = 0.27$, $t(288) = 4.33$, $p < 0.001$, $S^2_{ri} = 0.05$, and Emotional Stability, $\beta = 0.16$, $t(288) = 2.42$, $p = 0.02$, $S^2_{ri} = 0.01$ (Table 3).

Safety Compliance. The Big-Five personality traits, $R^2 = 0.31$, $F(6, 288) = 23.20$, $p < 0.001$, accounted for significant incremental variance ($\Delta R^2 = 0.18$) in Safety Compliance scores above Safety Climate $R^2 = 0.14$, $F(1, 293) = 48.49$, $p < 0.001$, $\beta = 0.21$, $t(293) = 4.06$, $p < 0.001$, $S^2_{ri} = 0.04$. All personality traits were significant predictors of Safety Compliance except for Emotional Stability and Openness. Agreeableness, $\beta = 0.25$, $t(288) = 4.10$, $p < 0.001$, $S^2_{ri} = 0.04$,

Table 3 Hierarchical multiple regression with standardized betas, squared semi-partial correlations and R^2 change for safety participation and safety compliance

Safety predictors	Safety participation			Safety compliance		
	ΔR^2	β	r_{si}	ΔR^2	β	r_{si}
Step 1						
Safety climate	**0.14**[a]	**0.27**[a]	**0.06**	**0.14**[a]	**0.21**[a]	**0.04**
Step 2						
Personality	**0.13**[a]			**0.18**[a]		
Agreeableness		**0.13**[c]	**0.01**		**0.25**[a]	**0.04**
Conscientiousness		**0.20**[a]	**0.03**		**0.24**[a]	**0.04**
Extroversion		**0.27**[a]	**0.05**		**0.22**[a]	**0.03**
Emotional stability		**0.16**[b]	**0.01**		**0.14**[c]	**0.01**
Openness		−0.02	–		0.001	–
Total R^2	**0.25**[a]			**0.31**[a]		

Note Significant ($p < 0.05$) beta weights are in bold. [a]$p < 0.001$, [b]$p < 0.01$, [c]$p < 0.05$

Conscientiousness, $\beta = 0.24$, $t(288) = 4.02$, $p < 0.001$, $S^2_{ri} = 0.04$, Extroversion, $\beta = 0.22$, $t(288) = 3.70$, $p < 0.001$, $S^2_{ri} = 0.03$, and Emotional Stability, $\beta = 0.14$, $t(288) = 2.19$, $p = 0.03$, $S^2_{ri} = 0.01$.

4 Discussion

This study has examined the relative influence of safety climate perceptions and individual differences on safety performance. In short, we were able to assess whether safety climate or personality traits accounted for more variance in self-report safety performance scores. Overall, personality as a construct (inclusive of all five personality traits), displayed significant incremental and unique variance above safety climate in all four safety performance measures (ΔR^2 = 0.13–24). This suggests that personality variables provide unique value over safety climate in understanding a number of safety performance indicators. Although only between 22 and 31 % of the total variance in safety performance was accounted for at step 2 (R^2 = 0.22–0.31), if safety climate was the sole predictor, we would only be able to explain 5–14 % of the variance (R^2 = 0.05–0.14).

Individually, the amount of variance accounted for by each of the five personality traits is similar to that of safety climate for several of the safety performance indicators. Safety Compliance, for example, is explained uniquely by 4 % by Safety Climate, 4 % by Agreeableness, 4 % Conscientiousness, 3 % Extroversion, and 1 % Emotional Stability. The exception is within Safety Knowledge, where safety climate does not account for any unique variance when personality traits are introduced. Each safety performance indicator will be discussed further below. The results of this unique study suggest that both personality and safety climate are important predictors of safety behavior, although, further longitudinal research would be required to make a claim as bold as this.

4.1 Safety Performance Personality and Safety Climate

The personality traits and safety climate measures related to the four safety performance outcomes in both similar and unique patterns. In general, all of the relationships were significant and in the expected direction, with the exception of Safety Climate in relation to Safety Knowledge. Additionally, the personality trait Openness was not significantly related to any of the safety performance indicators.

Safety Motivation had the least variance accounted for (R^2 = 0.22), with each predictor accounting for only 2–3 % of the unique variance. Motivation to work safely may be captured better by other predictors that were not included in this study. Safety Knowledge (R^2 = 0.28) was predicted by Conscientiousness (5 %), Extroversion (4 %) and Agreeableness (3 %), but not Safety Climate. This finding suggests that the influence of Safety Climate becomes negligible when in the presence of personality traits; that is, the level of Conscientiousness, Extroversion, and Agreeableness are more influential in understanding an employees' knowledge of safety practices than is safety climate. Conversely, Safety Climate was the top predictor in Safety Participation (R^2 = 0.25), accounting for 6 % of unique variance. However, Extroversion (5 %) and Conscientiousness (3 %) were also uniquely important to understanding Safety Participation. And to reiterate from before, Safety Compliance (R^2 = 0.31) unique variance was relatively evenly distributed among the predictors.

4.2 Limitations

This study is not without its limitations. The use of a relatively small sample of railway transportation workers may not necessarily generalize across organization or industry. Additional studies may be required across industry and organization to validate these findings. Additionally, the data may be influenced by common method variance, that is, responses from a single method can inflate or attenuate the strength to the relationship [29]. Further, objective and longitudinal safety performance measures would strengthen the claims, and we recommend that future research explores more objective safety performance indicators over multiple time points.

4.3 Implications

The results of this study support a systems approach towards safety. Each factor plays a role in understanding the current safety system. The results of this study bring to light the importance of broadening the scope of safety research to include variables and predictors that are critical, even if they have acquired a negative stigma in the past. Ignoring the influence of individual differences and focusing solely on organizational factors should not be an acceptable practice. There may be merit in considering personality factors during recruitment of employees for safety critical roles.

Although the TSB has made several recommendations on how to avoid a reoccurrence of an accident similar to Lac-Mégantic disaster, they are policy based, broad in scope and open to interpretation. Given that each of the 18-causal factors can have multiple underlying antecedents, the solutions and interventions for railway organizations will need to be more tailored and precise, as it is counter-productive to use a broadsword when trying to conduct brain surgery.

References

1. Transportation Safety Board of Canada, Minster of Public Works and Government Services.: Lac-Megantic runaway train and derailment investigation summary, http://www.tsb.gc.ca/eng/rapports-reports/rail/2013/r13d0054/r13d0054-r-es.asp
2. Rundmo, T.: Associations between risk perception and safety. Saf. Sci. **24**, 197–209 (1996)
3. Mearns, K., Flin, R.: Risk perceptions and attitudes to safety by personnel in the offshore oil and gas industry. J. Loss Prev. Process Ind. **8**, 299–305 (1995)
4. Neal, A., Griffin, M.: Safety climate and safety at work. In: Barling, J. (ed.) The Psychology of Workplace Safety, pp. 15–34. APA, Washington, DC (2004)
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**, 1103–1127 (2009)
6. Reason, J.: Human error: models and management. BMJ **320**, 768–770 (2000)
7. Greenwood, M., Woods, H.M.: The incidence of industrial accidents upon individuals with special reference to multiple accidents. Industrial Fatigue Research Board, Medical Research Committee, Report No. 4. Her Majesty's Stationery Office, London (1919)
8. Wong, W.A., Hobbs, B.A.: Personal factors in industrial accidents: a study of accident proneness in an industrial group. Ind. Med. Surg. **18**, 291–294 (1949)
9. Zohar, D., Luria, G.: A multilevel model of safety climate: cross-level relationships between organization and group-level climates. J. Appl. Psychol. **90**, 616–628 (2005)
10. Griffin, M., Neal, A.: Perceptions of safety at work: a framework for linking safety climate to safety performance, knowledge, and motivation. J. Occup. Health Psychol. **5**, 347–358 (2000)
11. Clarke, S.: The relationship between safety climate and safety performance: a meta-analytic review. J. Occup. Health Psychol. **11**, 315–327 (2006)
12. Johns, G., Saks, A.L.: Organizational Behaviour Understanding and Managing Life at Work. Pearson Prentice Hall, Toronto (2008)
13. Goldberg, L.R.: A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In: Mervielde, I., Deary, I., De Fruyt, F., Ostendorf, F. (eds.) Personality Psychology in Europe, vol. 7, pp. 7–28. Tilburg University Press, Tilburg (1999)
14. Costa Jr., P.T., McCrae, R.R.: Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO FFI) Professional Manual. Psychological Assessment Resources, Odessa, FL (1992)
15. Costa, P.T.: Work and personality: use of the NEO-PI-R in industrial-organizational psychology. Appl. Psychol. Int. Rev. **45**, 225–241 (1996)
16. Clarke, S., Robertson, I.T.: 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)
17. Arthur Jr., W., Graziano, W.G.: The five-factor model, conscientiousness, and driving accident involvement. J. Pers. **64**, 593–618 (1996)
18. Glendon, A.I., Clarke, S.G., McKenna, E.: Human Safety and Risk Management, 2nd edn. CRC Press, New York (2006)
19. Hansen, C.P.: A causal model of the relationship among accidents, bio data, personality, and cognitive factors. J. Appl. Psychol. **74**, 81–90 (1989)
20. Iverson, R.D., Erwin, P.J.: Predicting occupational injury: the role of affectivity. J. Occup. Organ. Psychol. **70**, 113–128 (1997)
21. Eysenck, H.J., Eysenck, M.W.: Personality and Individual Differences: A Natural Science Approach. Plenum, New York (1985)
22. Jonah, B.A.: Sensation seeking and risky driving: a review and synthesis of the literature. Accid. Anal. Prev. **29**, 651–665 (1997)
23. Cellar, D., Nelson, Z., Yorke, C., Bauer, C.: The five-factor model and safety in the workplace: investigating the relationships between personality and accident involvement. Prev. Interv. Commun. **22**, 43–52 (2001)

24. Barrick, M.R., Mount, M.K.: The big five personality dimensions and job performance: a meta-analysis. Pers. Psychol. **44**, 1–26 (1991)
25. Burke, M.J., Sarpy, S.A., Tesluk, P.E., Smith-Crowe, K.: General safety performance: a test of a grounded theoretical model. Pers. Psychol. **55**, 429–457 (2002)
26. Slappendel, C., Laird, I., Kawachi, I., Marchall, S., Cryer, C.: Factors affecting work-related injury among forestry workers: a review. J. Saf. Res. **24**, 19–32 (1993)
27. Neal, A., Griffin, M.A.: A study of the lagged relationships among safety climate, safety motivation, and accidents at the individual and group levels. J. Appl. Psychol. **91**, 946–953 (2006)
28. Neal, A., Griffin, M.A., Hart, P.M.: The impact of organizational climate on safety climate and individual behaviour. Saf. Sci. **34**, 99–109 (2000)
29. Podsakoff, P.M., Mackenize, S., B, Podsakoff, N.P.: Common method biases in behavioural research: a critical review of the literature and recommended remedies. J. Appl. Psychol. **88**, 879–903 (2003)

The Use of Modelling Tool in Order to Evaluate the Dwelling Times for Trains

Guillaume Craveur and Olivier Anselmi

Abstract The exchanges between the platform and the trains structure the punctuality of those. Indeed, the positioning of the travelers who expect the train on the platform, their choice of car in the train according to the exit door in their station of arrival, are inter alia elements which condition the dwelling times of trains. In mass transit railway or commuter rail systems, dwelling times are usually long and chaotic, which can lead to tardy trains and a decrease in the system efficiency, especially during peak hours. For more efficient and robust schedules, for an improvement of passengers' comfort, a train operator must take care of the passenger movements in the train and on the platform in order to improve the design of both trains and platform to optimize pedestrian flows.

Keywords Rolling stock · Numerical simulation · Dwelling times

1 Passengers Counting Method

The current economic world is structured around the collection of data of all kinds that they come from social networks, loyalty cards, surveys, or private sales systems. From these data collections, resulting statistical studies to better target customers, their expectations, their needs with the aim of constant continuous improvement of services offered, which can sometimes emerge on the creation of new services, or conversely deleting other. As we will see later, the field of passenger transport has not escaped this trend, especially in the railway world. Indeed, in a holding where the reservation instead—a typical example of "reserved seat" being the French TGV model—is not required and where unlimited monthly

G. Craveur (✉) · O. Anselmi
SNCF, Rolling Stock Department, Rolling Stock Engineering Centre (CIM), 4 allée des Gémeaux, 72100 Le Mans, France
e-mail: guillaume.craveur@sncf.fr

O. Anselmi
e-mail: olivier.anselmi@sncf.fr

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_5

Fig. 1 Illustration for passengers counting for a stop

subscriptions prevail, then it becomes difficult or impossible to know precisely the number of passengers on a given route. That is how appears historically the concept of “Passengers Counting” to identify the occupancy rate of the train and also to determine fraud. A focus is done on this article for the data which served to improve the exchange between passengers at the platform.

The necessary data are the following:

- “Direction”: Knowing of the number of passenger per direction: got in/got down
- “Door”: Knowing of the number of passenger per direction: got in/got down for each door
- “For each stop”: Affect the number of got in/got down traveler at a given stop (Fig. 1).

Historically, the first passengers counting systems were made by the board train control agents, via handwritten manual tally sheets. But often, this mission is just one of the side missions of controllers, it was not uncommon that the quality of this system does not accurately reflect the reality of things and do not meet the requirements listed above, always more complex.

Nowadays, since the omnipresence of embedded computing that these manual records were replaced by a dedicated architecture based on the science of networks called system of “Embedded Automatic Passenger Counting”. Many large retail chains, among others, resorted to computerized solutions counting the number of customers entering and leaving their facilities. These solutions have greatly inspired the equipment on board rail vehicles, but specific arrangements have been made as regards existing constraints of the environment of rolling stock. Indeed, the sensors used in the “static” world require implantation at a minimum height in order to detect the passage in one direction or the other. This is not a problem in shopping centers, but in an onboard integration, where cars ceiling height is defined by a standardized template and a highly constrained environment, a specific engineering work was required to obtain the performance at least identical to “static” solutions (Fig. 2).

A specific architecture from “embedded world” is needed. Here is how integrated is a passenger counting system in a railway environment in a concrete way (Fig. 3).

The train architecture presented above is actually itself a subsystem of the final operation that is made. Indeed, in ever more connected world, rail vehicles are today, as well as watches and other smartphones enrolled in universe says “communicating”.

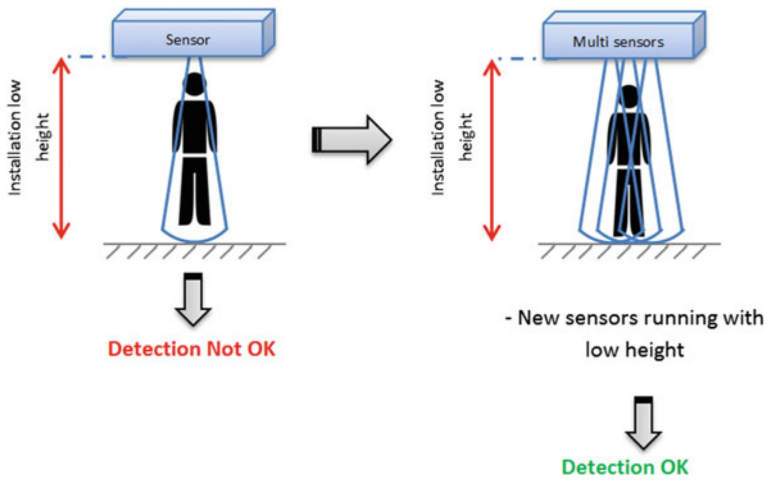


Fig. 2 The rail solution for detection

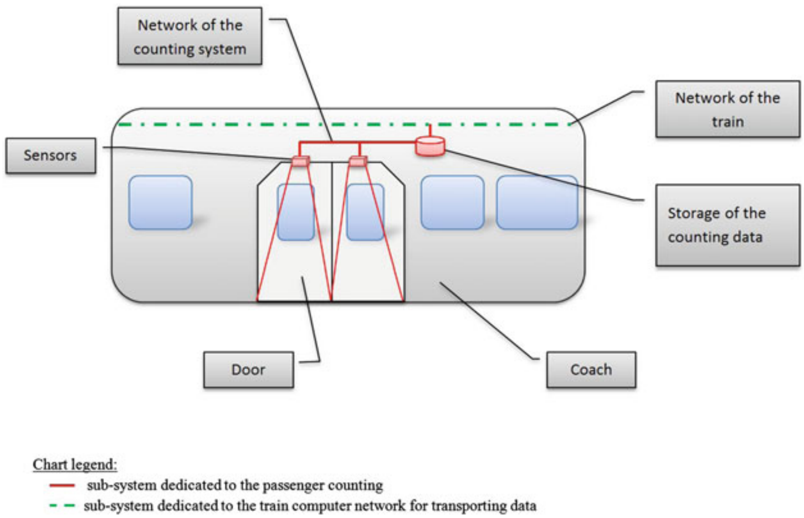


Fig. 3 Train network: embedded computing

Steps described in Fig. 4:

- *A*: The train ("Onboard") is connected via radio between its embedded telecommunication unit (3G, 4G, Wi-Fi, …) with the telecom operators at ground via the relay antenna. An ongoing exchange:
  - Data from the train to the ground
  - Data from the railway operators to the train

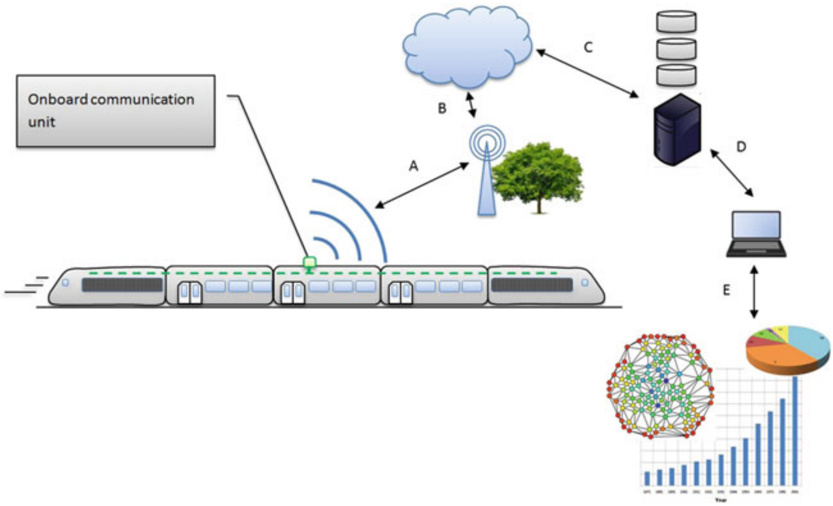


Fig. 4 A complete architecture: the "Onboard/Ground"

- *B*: Internal Routing between the network operator and the network "cloud" specific to the operator
- *C*: Data Storage:
    - From the train
    - From the ground waiting to be send to the train concerned
- *D*: Receipt of data on the operator's information system
- *E*: Analysis, processing and formatting data from the "Onboard".

The analysis, processing and formatting data is sort of purpose expected by the railway operator. Indeed, it is this part on the analysis of data from the train which allows it to modify certain operating rules to improve traveler comfort, more fluid traffic management, or to adapt the transport plan. All this is possible thanks to the knowledge of the real needs of the traveler, from the study of his movements and behavior. Here are some real examples that illustrate the practical use of passengers counting data in a complex railway environment.

- Door by traveler counting allows to study the location of some boarding platform, or evacuation platform. We know that in peak period (beginning or end of working day), in big cities, that passengers walk to the front of the train shortly before the train arrived at the terminus station in order to be closer to exit doors. It then becomes relevant to achieving the amenities and traveler access infrastructure necessary to absorb this important flow of travelers within a few minutes.

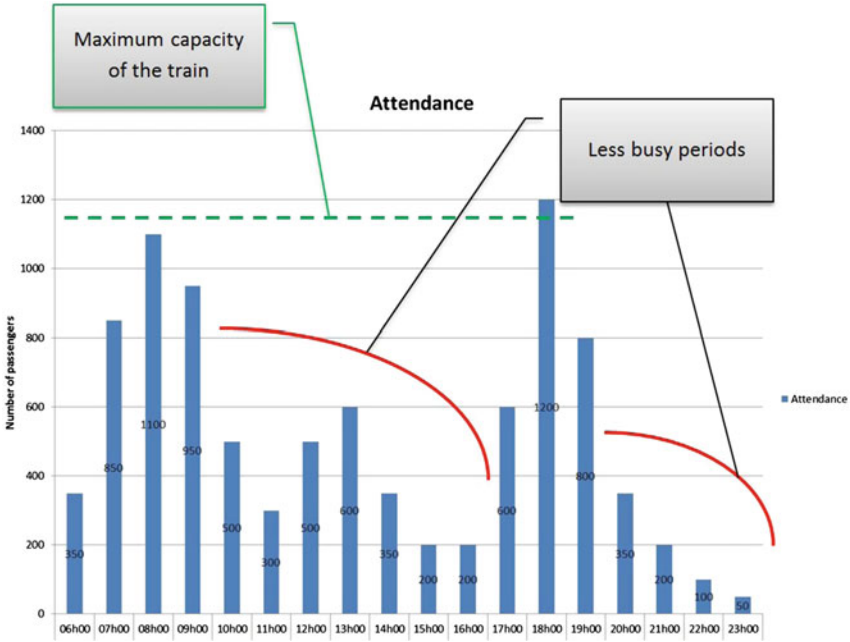


Fig. 5 Example of number of passengers in function of the timetable

- Interior doors can also be equipped by sensors in order to know the passengers flux inside the train.
- The transportation plan can also be adapted depending on attendance. Indeed, the latter varies the rhythm of work days or special events (concerts, sport events…). Thanks to statistical surveys of passenger counting data, it is possible to adapt the mobilization of a rowing fleet. For example, it is relevant to mobilize the maximum of trains during peak periods (morning or evening) by coupling trains them to increase the capacity of travelers, but also by increasing the frequency passage in railway station. Rather, it takes advantage of the less busy periods to reduce frequency shift and thus free up time for equipment maintenance operations (Fig. 5).

Finally, we conclude these series of concrete operational applications of metering data, by the illustration of one of the crucial components in the design of downtime of train's station: "Exchange Time". This variable is the difference of time between the first event counting (input or output) and the last event (input or output) at a given stop. This time is widely used especially in the design flow of travelers, as explained in the following section.

2 Simulations Done with BuildingEXODUS

The software computes the trajectory of each person according to the path which they chose to borrow at each step time. The geometry and the grid defined in building EXODUS are well adapted to the real coach geometry. In buildingEXODUS, the entire space of the geometry is covered with a grid of nodes which are generally spaced by intervals of 0.5 m. At any time, only one person can occupy one node. It implicates a conflict management based on several parameters like the traveling time, the parameter "drive" which measures the capacity for a person to win a node against another, etc.

There are different types of nodes: "free space nodes" which allow unhindered movement and represent unobstructed horizontal terrain, "seat nodes" which represent seating area and either force occupant to engage in hindered movement, "stair nodes", "external exit nodes" [1] for those which interest us. A person can pass from one node to one of the neighborhood nodes connected by arcs. The arcs have a physical length, i.e. the distances between all the nodes are the real distances of the geometry. Moreover, arcs have a specific parameter called "obstacle" which defines the difficulty of passing from one node to the other. The walking speed on an arc is then defined by the value Walk/Obstacle where Walk is the speed of standard walk of a person. Therefore, more the value of the parameter obstacle is high (it is the case for arcs between seats) more the time to pass from one node to the other is important. The localization of the various nodes has to be representative of the possible ways followed in the trains. The units of passage are respected in all the sites of the train. Among a large number of modifiable parameters, only the gender, the age, the weight and the size of each person were fixed with respect to the real people. All the others parameters are randomly defined according to the previous parameters and the associated databases. An illustration of buildingEXODUS is on Fig. 6.

Fig. 6 Illustration of buildingEXODUS applied on a typical train coach

3 Results Obtained with BuildingEXODUS

The results of two cases of application are presented in this article.

The first study concerns the rolling stock called NAT (Nouvelle Automotrice Transilien). This is a new range of trains which enter into service in Parisian suburb in 2010. This rolling stock can have eight coaches for a normal capacity of 922 passengers whose 472 seated. Each coach has lateral access doors with an opening width dimension of 1950 mm which corresponds to three units of passage.

For this first study, the scenario is the following: 17 people are waiting for the train on the platform and will get on the train and 12 people are in the train and will get off the train. So, we have here an exchange between 29 passengers. The output data is the allocated exchange time which is 38 s.

In order to represent as closely as possible this case of application, we had to analyze the incoming data. For example, it is necessary to know the initial position of each passenger in order to represent this configuration of recording. It is very important to keep in mind that the results obtained are only for this specific application, that is to say this train, on this platform with this number of passengers.

The results being able to vary from a simulation to another because of the randomness of many parameters, twenty simulations were carried out. The average exchange time for these 20 simulations is 41 s. For recall, the exchange time of the real test is 38 s. Thus, the average variation with the real test is nearly 8 %, the maximum variation with the real test is 13 % and the minimum variation with the real test is 2 %. Figure 7 shows these results.

The standard deviation is used to measure the dispersion of a data set. The weaker it is, the more the values are gathered around the average. Rather than there are only 20 simulations performed, this indicator is interesting because it permits to show the robustness of the software. The results obtained with buildingEXODUS have a standard deviation equal to 1.41. This point shows that the "node-to-node"

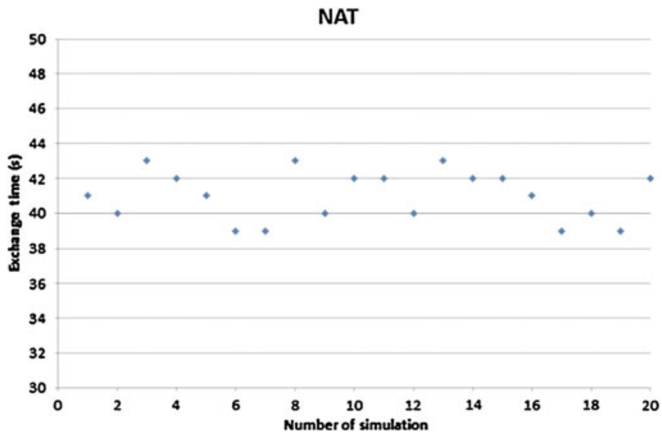


Fig. 7 Exchange times for the NAT

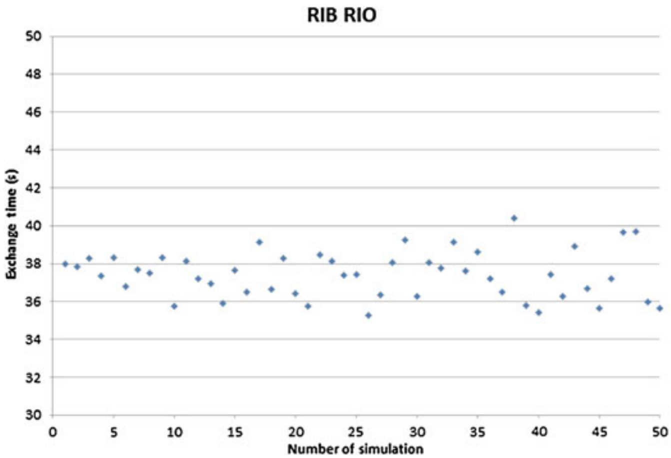


Fig. 8 Exchange times for the RIB RIO

system of buildingEXODUS permits to have a very good reproducibility and repeatability. Another point which is very interesting to have in mind is the very short calculation time. Indeed, the average calculation time for each of the 20 simulations made with buildingEXODUS is around 4 s.

The second study concerns the rolling stock called RIB RIO (Rame Inox de Banlieue–Rame Inox Omnibus). This rolling stock can have seven coaches for a capacity of 1263 passengers whose 796 seated. Each coach has lateral access doors with an opening dimension of 1300 mm which corresponds to two units of passage.

For this second study, the scenario is the following:

- 41 people are waiting for the train on the platform and will get on the train,
- 4 people are waiting on the platform and will stay on it,
- 40 people are in the train and will get off the train,
- 40 people are in the train and will stay in it.

So, we have here an exchange between 125 passengers. The output data is the allocated exchange time which is 38 s.

The results being able to vary from a simulation to another because of the randomness of many parameters, fifty simulations were carried out. The average exchange time for these 50 simulations is 37.4 s. For recall, the exchange time of the test is 38 s. Thus, the average variation with the real test is nearly 1.5 %, the maximum variation with the real test is 7 % and the minimum variation with the real test is 0 % (which corresponds to an exchange time of 38 s). The Fig. 8 shows these results.

The standard deviation is used to measure the dispersion of a data set. The weaker it is, the more the values are gathered around the average. Rather than there are only 50 simulations performed, this indicator is interesting because it permits to show the robustness of the software. The results obtained with buildingEXODUS have a standard deviation equal to 1.23. This point shows that the "node-to-node" system of buildingEXODUS permits to have a very good reproducibility and repeatability. Another point which is very interesting to have in mind is the calculation time. The average calculation time for each of the 50 simulations made with buildingEXODUS is around 4 s.

4 Conclusions/Prospects

To conclude, the simple empirical concept of the passengers counting turns implementing complex in a global railway environment. However, exploitation of these data can significantly improve passenger comfort; orient the station according to his need, his behavior, but also to adapt the transportation plan on account of variations in the rate of attendance of the trains. The management of passenger flows is a major issue of mobility called "smart mobility".

The results obtained from the numerical study achieved by the SNCF Rolling Stock Engineering Centre show that a very good agreement was obtained with the buildingEXODUS software. Moreover the exchange times performed by simulations are very close to the time monitored during tests. It seems that for this application, buildingEXODUS could be used.

Finally, this study shows advantages of exchange modelling which allows realizing a lot of different simulations by changing the incoming data and which is less expensive than a real scale test (it does not need the immobilization of a train, hundreds of volunteers and a platform of station) which does not give any information about average time and standard deviation. This tool can also be used to optimize platforms.

To note that in the case of a call for a new train supply, the numerical simulation makes possible to test upstream and to determine if the specifications are fulfilled. Thus, it will be possible to compare diagrams during the design phase of the project. Moreover, with these tools, it is possible to compare train's architecture, to compare the exchange time in function of the population, to simulate the evacuation with or without luggage, to simulate panic movement…. All those items have an influence which obviously can not be really tested but which can be computed, which is a main advantage of numerical simulation. Specialized tools exist nowadays and have reached an important level of reliability. Some of them are of general purpose so are used in various fields (aeronautics, naval industry, civil engineering and building industry…). Very high level trainings are now available for engineers already familiar with numerical simulation, especially in the field of computational fluid dynamics and finite element simulation for mechanical structures.

The work to do now is to achieve others tests (with others populations, trains, platforms) and compare them to simulations.

Acknowledgments We would like to thanks our colleagues at SNCF Rolling Stock Engineering Centre for the help they gave us to do this work.

Guillaume Craveur also would like to thank Prof Ed Galea and senior research fellow Lazaros Filippidis from the Fire Safety Engineering Group of the University of Greenwich who taught me the human behavior and the practical use of buildingEXODUS for movement and evacuation. Finally Guillaume would like to thank Prof Jean-Charles Craveur from ISMANS Engineering School for his lectures and the Mastère Spécialisé in Fire Safety Engineering he developed at ISMANS.

Olivier Anselmi would like to take this opportunity to thank Sandra Cornille who have contributed in some way to this article and especially for the given translation work.

Reference

1. buildingEXODUS (version 5)—user's guide (2011)

An Overview of the Factors Associated with Driver Distraction and Inattention Within the South African Railway Industry

Inga Dambuza

Abstract Driver performance and the attention the driver pays to the primary task of operating trains safely have been shown to decrease because of driver distraction and inattention. This paper presents the factors associated with driver distraction and inattention within the South African railway environment, current interventions utilized by South African railways and the impact these interventions have on driver distraction. This paper also stresses the need to conduct future research into driver distraction in order to mitigate the contribution of driver distraction and inattention to railway occurrences experienced in South Africa.

Keywords Human factors · Driver behaviour · Train driver distraction · Inattention · South african railway industry

1 Introduction

Driver distraction and inattention has been found to be a contributing factor to accidents and plays a role in railway occurrences that have occurred in the South African railway industry.

The factors determined to contribute to driver distraction and inattention range from the technologies introduced to improve driver performance, secondary tasks believed to increase safety and aid in the safe operation of the rolling stock, as well as interventions thought to maintain driver alertness. Examples of some contributory factors include excessive noise from the cab design, the lack of toilets on the trains, the use of the dead man's switch and of mobile phones, listening to music and conversing with a passenger, as well as the use of a train driver assistant. Although the nature and application of the technologies may differ, the collective

I. Dambuza (✉)
Railway Safety Regulator, Lake Buena Vista Building no. 1 Gordon Hood Avenue, Centurion, South Africa
e-mail: ing.ad@rsr.org

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_6

objective of said technologies introduced is to prevent driver distraction, but often times these interventions further add to the inattention from the primary task of driving trains.

While technology has the potential to improve safety while driving, its acceptability and adequacy should be investigated thoroughly before implementation in the railway industry, taking the driver's needs and functional abilities into account [1].

Despite the complexities associated with driving, it is typical to see drivers executing other activities while driving, such as talking to passengers or listening to music [1] often times without understanding the consequences related to attempting to multitask while driving or understanding the level at which their safety is compromised.

While literature has provided abounding research on the contributing factors to driver distraction, many of the theories on the measurement of driver distraction are conflicting. The findings from studies have also been inconclusive and some of the interventions put forward may not be practical in different contexts or environments.

The majority of the literature available on driver distraction and inattention focuses on road vehicle drivers and a few have been studies conducted in the US, UK, Australian and New Zealand railway industries. Very little to no research has been conducted within the railway environment in South African and the effects of the current risk factors and interventions that exist in South Africa have not been adequately researched. Research into the factors affecting driver distraction is critical to prevent accidents and increase the railway safety [2]. Driver distraction can lead to the train driver missing vital information such as signals, approaching trains, road vehicles or pedestrians at level crossings [3].

While the contribution of driver distraction and inattention to South African railway occurrences has yet to be quantified, the need for in-depth investigation into the matter is required in order to provide practical and realistic interventions and is crucial in the pursuit of achieving zero railway occurrences.

1.1 Aims and Objectives

The objective of this paper is to highlight the factors associated with driver distraction and inattention within the South African railway environment and to contextualize how the existing interventions introduced in South African railway can affect driver distraction and inattention. This paper also aims to provide a baseline in order for the South African railway industry to continue and further the research in an attempt to quantify and understand how driver distraction and inattention contributes to railway occurrences and how it can be mitigated in order to decrease the number railway occurrences in which driver distraction play a role.

2 Driver Distraction and Inattention

Different definitions for driver distraction and inattention exist within the literature, with conflicting meanings and interpretations depending on the context in which they are used. Driver distraction is defined by Naweed [4] as “the diversion of attention away from activities critical for safe driving towards a competing activity”. The National Highway Transportation Safety Administration (NHTSA) has defined distracted driving as “an activity that could divert a person’s attention away from the primary task of driving” [5]. Other literature presented by Freund et al. [6] defines driver distraction as a loss of attention. What the definitions seem to suggest is that distraction diverts the person’s attention away from something and that diversion interrupts the ability to concentrate on something else [7].

The definition that is perhaps relevant for the context of this paper is that of the American Automobile Association Foundation for Traffic Safety, who define driver distraction as occurring “when a driver is delayed in recognizing information that is required to safely complete the driving task due to some event, activity, object or person within or outside the vehicle compelled to induce the driver’s shifting attention away from the driving task” [1].

There are three categories in which distracted driving is classified in:

1. Manual distraction, where the driver removes their hands form the steering wheel;
2. Visual distraction, where the driver removes their eyes from the road; and
3. Cognitive distraction, where the driver’s mind is not entirely focused on driving due to talking to other passengers, using cellular phones or merely thinking.

While the above categories are applicable within the railway industry, Young et al. (2003) presents a fourth category, auditory distraction. Auditory distraction is stipulated to occur when a driver momentarily or continuously focuses their attention on sounds or auditory signals instead of focusing on the road [1]. Auditory distraction typically occurs when the driver is listening to the radio or holds a conversation with a passenger, but is most pronounced when using their mobile phone [1].

There is a currently a lack of understanding of the different types of driver distraction and inattention and the mechanisms that give to distraction in the South African railway environment. These mechanisms have different implications in terms of the types of interventions used and the likely effectiveness of such interventions [7].

2.1 *Factors Affecting Driver Distraction in South African Railways*

Operating trains requires a high level of concentration, attention and alertness [9]. Train drivers are affected by visual, cognitive, auditory, as well as psychological and physical disturbances [2]. The driver's predisposition to various disturbances and risks that impact distraction increase the longer the driver must sustain attention [6]. The complex interaction of various factors such as tediousness, physical and mental workload, environmental stressors, the quality and quantity of sleep and circadian effects result in distraction and inattention [6].

Research has demonstrated that loss of attention typically occurs when the driver works for long periods, receives inadequate sleep or works during times when circadian rhythms are at their lowest [6]. Further research on task-related distraction in the rail industry is required in order to fully understand the other contributory factors to distraction and inattention [10].

While many factors exist that distract train drivers, the focus of this paper will be on the use of mobile phones, locomotive design and the physiological and psychological conditions as these factors are the most prevalent within South African railways.

The Use of Mobile Phones. The research on the use of mobile phones is extensive within the road transportation industry; however the contribution of mobile phones to driver distraction in the South African railway industry is somewhat lacking. Studies on driver distraction conducted by the Rail Safety and Standards Board (RSSB) have found the use of mobile phones to be of particular concern. One particular study conducted by the RSSB found that the use of mobile phones lead to reduced situational awareness, poor speed control, slower reaction, reduced decision-making and less attention paid to checking for hazards [3].

Research has shown that while completing different cognitive tasks, two different areas of the brain are utilized, which can result in performance problems while attempting to perform those tasks [11]. According to Wickens [11], this could help to explain why using mobile phone can have an effect on what the driver sees as the tasks compete for the brain's information processing resources, which in turn limits mental workload.

The use of mobile phones while driving has been shown to cause physical and cognitive distraction and has also been shown to significantly diminish the visual search patterns employed by the driver, the reaction times, processes used to make decisions and the ability of the driver to maintain speed [1].

In the South African context, the use of mobile phones while operating trains is somewhat of a vexation. The primary means of communication in South African railways is the trunk radio. Most locomotives are fitted with a trunk radio and the driver is required to use the trunk radio when communicating with other railway employees. The problem that arises is that most of the trunk radios have been found through audits and inspections to be consistently defective in most locomotives. This forces the train drivers to utilize their private or company issued mobile phones to communicate with the other employees. The distraction is further compounded by the fact that in certain areas of South Africa, the cellphone reception/signal is poor. The lack of or breakdown in communication has been found to be a contributing factor to many railway incidents, but the contribution of the distraction arising from the use of mobile phones is not absolutely clear.

While some of the South African railways have developed circulars forbidding the use of mobile phones while operating the train in certain conditions, the enforcement thereof is minimal and is difficult to enforce during train journeys. The RSSB has highlighted the difficulties associated with enforcement and recommend an education framework that will ensure that train drivers fully comprehend the potential risks and implications related to the use of mobile phones and key decision-making skills, so that driver are effectively able to assess the conditions in which it is safe to use the mobile phone [3]. The education framework could be of vital benefit to South African railways if utilized in a contextual manner.

Locomotive Design. The cab of the locomotive can be a contributing factor in cases where the seating is uncomfortable, where there is excessive noise [9] or where the equipment in locomotives is not functioning (e.g. faulty trunk radios). The above-mentioned contributing factors can be found in most of the older and outdated locomotives used in South African railway companies.

While new and modernized locomotives have been purchased, the older locomotives have still not been phased out and therefore some drivers are still exposed to such conditions. Diesel locomotives are used in various parts of the rail network where there is no overhead power and these locomotives have been found to expose the drivers to higher levels of noise. Compounded with the resistance to wear hearing protection to decrease the exposure to noise and the distraction associated therewith, the distraction provided by excessive noise still remains a significant problem for South African train drivers.

Most of the older locomotives in South Africa do not have toilet facilities, which can add to the driver's distraction as it means that the driver must wait to relieve himself or herself. This is not particularly a problem for train trips of shorter lengths, but poses a greater distraction for longer train journeys.

The use of older locomotives raises a need to investigate the effects of cab design on driver distraction and railway incidents further as the effect of the distraction as a result of the locomotive design has also not been adequately investigated. Therefore, the mitigation of said effects through the procurement and use of new locomotives will be difficult to evaluate.

Physiological and Psychological Factors. Certain medical or psychological factors can adversely affect the driver's alertness [6]. According to the National Safety Council [3], research has identified "reaction-time switching costs", which is the time the brain spends switching its attention and focus from one task to another. It has also been discovered that spending even small amounts of time switching from one task to another can lead to adverse risks with regards to delayed reaction and braking time [3].

During driving, the brain must extract information from shared and limited resources in order to fulfil the other tasks, which places constraints on the mental resources available for the primary task [12, 13].

While certain physiological and psychological conditions predispose the train driver to distraction and inattention, no research has been conducted in South African railways to determine the extent of the distraction that arises from these conditions.

2.2 Interventions to Mitigate Driver Distraction

"Train driver tasks are without a doubt psychologically demanding" [9]. In South Africa, the train driver is bombarded with different signals to observe, road vehicles to look out for at level crossings and other operational duties, such as communicating with the train control officer and checking that the load is still intact.

Railway companies employ different interventions in an attempt to mitigate driver distraction and in an attempt to improve operational safety. While introducing new technologies may increase the attention of the driver, the technologies must be investigated and the behavior of the driver must be addressed before implementation in order to determine whether the technologies will have any adverse effects on the driver [14]. Research has demonstrated that some of these interventions have offered little guarantee, while others have been criticized. The use of these technologies differs from country to country and from operator to operator.

Within South Africa, there are a number of strategies used, including the dead-man's switch (vigilance switch) and the driver assistant.

Dead-Man's Switch. South African railways utilize the dead-man's switch, or vigilance switch, as a means of checking the alertness of the train driver. This switch is a method employed to check the alertness of the driver in that it requires a response from the driver. In most long haul railway industries, the switch is installed on the floor of the cab, in front of the driver's seat. The switch sets off a noise at set intervals and the driver is required to use their foot to press the switch, failing which the device is used to apply the brakes automatically and the train comes to a stop [15].

One of the problems that arises with the use of the dead man's switch is that the drivers hear the noise emitted frequently and press the button without actively thinking about their actions [16]. This becomes automatic and if the driver is distracted or preoccupied with other tasks, the automatic pressing of the button occurs without thinking about the train speed or location [16].

The implementation of the dead man's switch in South Africa is a bit of a sore point, particularly the positioning of the dead man's switch. The main issue found with the positioning of the switch is that the driver's continue to drive the train with the foot on the switch continuously instead of pressing the switch when required.

This defeats the purpose of the switch as drivers continue to be distracted with other activities and factors, thereby nullifying the objective of the switch.

Other positions where the switch can be installed were not explored extensively before implementation in South Africa. The most effective position in which to place the dead man's switch must be investigated, as the switch should provide the required input from the driver while ensuring that the alertness of the driver is maintained and verified as desired.

Driver Assistant. The use of a driver assistant is an intervention that is unique to South Africa. The functions of the driver assistant vary from operator to operator but the primary function of the driver assistant is to assist the train driver in the movement of the train. Some auxiliary functions of the driver assistant include assisting the train driver to observe signals, reminding the train driver about speed restrictions, assisting with checking the train load and verifying the information communicated to the driver by the train control officer.

While conversing with passengers may not seem to be a great risk as it is an activity that is considered to be of low risk, Pauzié et al. [1] reported that passengers can be a form of distraction to the drivers under certain situations.

Some investigations into railway occurrences have revealed the driver assistant to be a distraction to the train driver at times, particularly through the conversations held by the driver and driver assistant. Conversations that have resulted in the driver missing vital information, such as signals at danger or speed restriction boards erected; leading to signals passed at danger or derailments due to the non-adherence to speed restrictions.

The employment of a driver assistant to assist the train driver and the effects thereof have not been investigated in South Africa. Future research conducted in South Africa should include an assessment of the driver assistant and the role the driver assistant plays on driver distraction and railways incidents.

3 Future Research to Determine Driver Distraction

Within the railway environment, there are currently a number of methods utilized to determine driver distraction. Some of the more common methods employed globally include driver response measures, vision related measures and manual-related measures [2].

Eye-Tracking Studies. Eye-tracking is a common method used to determine driver drowsiness and alertness. The movement of the eyes typically shows where the driver's attention is focused. This is the reason why studies have utilized eye glance behavior to determine driver attention [2].

Eye-tracking studies have been conducted in the UK rail industry. One particular study focused on tracking the movement of the driver's eyes while operating a train. The study revealed that drivers spend about 50 % of their time scanning the visual scene when approaching signals. The remainder of the time was spent focusing on railway signs, infrastructure and locations next to the track and signals [17].

Eye-tracking studies can help South African railways to determine how much time the driver spends focused on the primary task and can assist in determining how much time other activities and factors contribute to driver distraction.

Questionnaires and Other Subjective Methods. No verified questionnaire has been developed for use in railways yet, but questionnaires do exist for use in the road industries. These questionnaires can be modified so that they are suitable for the railway environment, but it will be important to ensure that they are specific to the environment in which they are being used. The questionnaires also need to consider all the contributing factors to driver distraction and need to effectively quantify the self-reporting and self-assessment of driver distraction.

One of the most used subjective assessment is the NASA TLX. This tool is easily available and can be administered effortlessly for train drivers to determine the workload associated with the primary and secondary tasks. Another tool that has been developed that has not gained the same amount of popularity as the NASA TLX is the Driving Activity Load Index (DALI). The DALI is similar to the NASA TLX as it has a scale rating. The main difference between the two tools is that the DALI is directed towards driving tasks. This tool has been effectively used for the road transportation industry but the application in the railway industry has yet to be investigated [4].

According to Freund et al. [6], some important questions to ask in the investigation of driver distraction and inattention are:

- How to define driver distraction and inattention within the South African context;
- When does it happen and as a result of the contribution or combination of which factors;
- How much does the driver distraction affect the deterioration of operator performance; and
- How can it be mitigated?

No studies have been conducted in South Africa resulting in a void in the information available in this regard. Future research in this field is required in an attempt to quantify the contribution of driver distraction to railway occurrences.

Future research can take from studies conducted in US, UK, Australian and New Zealand railway industries; however, the application must be specific to the South African environment.

4 Discussion/Conclusion

There is extensive literature available regarding driver distraction and inattention in the road transportation industry. The current literature available for the railway industry mainly exists within the US, UK, Australian and New Zealand railway environments. While this information is crucial to understanding driver distraction and inattention, the application to the South African environment must be investigated as the interventions and technologies utilized could have different consequences. The need for future research into this matter is greatly required in South Africa in order to determine the level at which driver distraction and inattention contributes to South African railway occurrences.

References

1. Pauzié, A., Manzano, J., Dapzol, J.: Driver's behavior and workload assessment for new in-vehicle technologies design. In: Proceedings of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design (2007)
2. Jain, J.J., Busso, C.: Assessment of driver's distraction using perceptual evaluations, self assessments and multimodal feature analysis. In: 5th Biennial Workshop on DSP for In-Vehicle Systems, Kiel, Germany (2011)
3. National Safety Council.: Understanding the distraction brain: why driving while using hands-free cell phones is risky behaviour. White Paper (2010)
4. Naweed, A.: Psychological factors for driver distraction and inattention in the Australian and New Zealand rail industry. Accid. Anal. Prev. **30**, 193–204 (2013)
5. Craye, C., Fahri, K.: Driver distraction detection and recognition using RGB-D sensor. In: Proceedings of the International Conference on Image Analysis and Recognition (2014)
6. Freund, D.M., Knipling, R.R., Landsburg, A.C., Simmons, R.R., Thomas, G.R.: A holistic approach to operator alertness research. U.S. Department of Transportation. Paper No. 950636. Paper presented at the Transportation Research Board 74th Annual Meeting January 22–28, 1995 Washington, DC (1995)
7. Regan, M.A., Hallett, C., Gordon, C.P.: Driver distraction and driver inattention: definition, relationship and taxonomy. Accid. Anal. Prev. **43**, 1771–1781 (2011)
8. Young, K., Regan, M., Hammer, M.: Driver distraction: a review of the literature. Report No: 206. Monash University Accident Research Centre (2003)
9. Srinivas, T., Vidyagasar, M.: The human intervention in train safety on Indian Railways. Ind. J. Res. **4**, 366–372 (2015)
10. Naweed, A.: Hurry up and wait: danger signals in the rail environment. Ergon. Aust. **3**, 1 (2013)
11. Wickens, C.D.: Multiple resources and mental workload. Hum. Factors **50**, 449–455 (2008)
12. Just, M.A., Carpenter, P.A., Keller, T.A., Emery, L., Zajac, H., Thulborn, K.R.: Interdependence of nonoverlapping cortical systems in dual cognitive tasks. NeuroImage **14**, 417–426 (2001)
13. Dux, P.E., Ivanoff, J., Asplund, C.L., Marois, R.: Isolation of a central bottleneck of information processing with time-resolved fMRI. Neuron **52**, 1109–1120 (2006)
14. Larue, G.S., Kim, I., Rakotonirainy, A., Ferreira, L., Haworth, A.: Integrating driving and traffic simulators to study railway level crossing safety interventions—A methodology. In: Computers in Railways XIII: Computer System Design and Operation in the Railway and Other Transit Systems, pp. 719–732. WIT Press, Southampton (2012)
15. Sai shenagha, D., Saranya, K., Vandana, I., Vijay Murugan, S.: Vigilance control system using Mcf5235 processor. Int. J. Res. Eng. Technol. **3**, 736–740 (2014)
16. Kumar, A., Sinha, P.K.: Human error control in railways. Jordan J. Mech. Ind. Eng. **2**, 183–190 (2008)
17. Luke, T., Brook-Carter, N., Parkes, A., Grimes, E., Mills, A.: An investigation of train driver visual strategies. Cogn. Tech. Work **8**, 15–29 (2006)

Investigating the Potential to Mitigate Crowding Issues on Trains by Providing Improved Information to Passengers

James Pritchard and John Preston

Abstract Crowded trains can adversely affect the experience of rail passengers and can cause practical issues for train operators. It is thought that some of these issues can be mitigated by providing better information to passengers and encouraging them to make different travel choices as a result. A pilot study was undertaken, in which rail passengers took part in a stated preference survey concerning the provision of information about crowding levels. It was found that some passengers would consider choosing a less crowded train, giving weight to the hypothesis underpinning the research.

Keywords Information provision · Rail travel · Behavioral change · Crowding

1 Introduction

Crowded trains can adversely affect the experience of rail passengers. They can also cause practical issues for train operators, especially if slow boarding and alighting at stations makes it hard to maintain tight dwell times. It is thought that some of these issues can be mitigated by providing better information to passengers and encouraging them to make different travel choices as a result.

On routes where the timetabled services are relatively frequent, it may be possible to spread passenger flow more evenly by encouraging passengers to choose a less busy train. This could be achieved by making historical crowding data available at the time of booking, allowing prospective passengers to take into account the typical crowding levels when planning their journey. Additionally, the use of more real-time information, made available via Customer Information

J. Pritchard (✉) · J. Preston
Transportation Research Group, Room 4001, Building 176, University of Southampton
Boldrewood Campus, Burgess Road, Southampton SO16 7QF, UK
e-mail: J.A.Pritchard@soton.ac.uk

J. Preston
e-mail: J.M.Preston@soton.ac.uk

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_7

Systems and mobile applications, could be used to encourage passengers to alter their travel choices nearer the time of departure.

The spread of passengers within a particular train can also contribute to overcrowding, because some carriages may fill up faster than others. The provision of real-time information (again via Customer Information Systems and mobile applications) could be used to mitigate this, by encouraging passengers to move to more lightly loaded carriages. If such information were made available to waiting passengers before boarding, it may be particularly beneficial for maintaining tight dwell times.

Previous work undertaken in this area includes studies of the effects of posters displaying (likely) train crowding information. The results suggest that there had been some change in passenger loading patterns as a result of the posters. However, it is thought that further benefits could be obtained by use of more real-time data and other methods of communicating with passengers. An increasing number of data sources for monitoring train loading are becoming available, including on-train systems which can estimate the number of passengers in a given carriage. Similarly, methods of communicating with passengers are also becoming more sophisticated, with an increased use of real-time information systems and a high proportion of smartphone usage. Hence it is now possible to provide passengers with better crowding information in order to influence their travel choices, and such systems are already operative on some suburban railway lines in Tokyo.

In order to ascertain whether it is worth investing in improved information systems, a pilot study was undertaken on the Gatwick Express, in the UK, which provides a regular service for both airport passengers and commuters. Train passengers were surveyed during their journey, using a stated preference methodology.

2 Methodology

With the full co-operation of Govia Thameslink Railway (GTR), who operate Gatwick Express services between London Victoria, Gatwick Airport and Brighton, on-train surveys were conducted on February 11th and 12th 2015. The surveys were conducted on ASUS T100TA tablet PCs, using software provided by Snap Surveys (www.snapsurveys.com). The surveys comprised five main sections:

1. Participant consent and basic demographics.
2. The participant's journey details and their perceptions of comfort, crowding and reliability.
3. A stated preference exercise concerning information about train crowding levels, which may be provided before the journey commences (for example, via a smartphone app or ticket booking website).
4. A stated preference exercise concerning information about train crowding levels, which may be provided at the station before boarding.

5. A stated preference exercise concerning information about carriage crowding levels, which may be provided at the station before boarding.

The survey software automatically skipped questions which were not relevant; for example, if a participant selected that they did not use any sources of information to plan a journey in advance, they were not shown the first stated preference exercise. To help eliminate the effects of fatigue (when participants stop paying attention to the options they are selecting), the order stated preference questions was randomized each time.

The first stated preference exercise, henceforth referred to as SP1, presented the participant with nine sets of three options; in each case the participant could choose from a train which matched their chosen timings, a train which departed and arrived a number of minutes earlier, and a train which departed and arrived a number of minutes later. In each of the nine cases, the crowding levels (either “no seats available”, “40 % of seats available” or “90 % of seats available”) and the relative departure times of the earlier and later trains were varied. The train which didn’t require any changes in travel timings was always shown as having no seats available.

The second stated preference exercise, henceforth referred to as SP2, followed a similar principle, although in this case there were only two options each time; it was assumed that a passenger arriving at a station would not be in a position to catch an earlier train (or that if they were, their decision would not necessarily be based around crowding issues). Hence there were six repeated observations, asking participants to choose between a train leaving immediately with no available seats and one leaving a given number of minutes later with some available seats.

The third stated preference exercise, concerning the choice of carriage on a given train, rather than the choice of train, is not discussed in this paper.

Two versions of the survey were used during the pilot study: a ‘Standard’ version in which estimated train crowding levels were shown graphically using colors (Fig. 1), and an ‘Alternative’ version in which estimated crowding levels were stated (Fig. 2). This was so that an insight into the effect of information presentation could be obtained. Some of the tablets were set up with the ‘Standard’ version, whilst the remainder were set up with the ‘Alternative’ version.

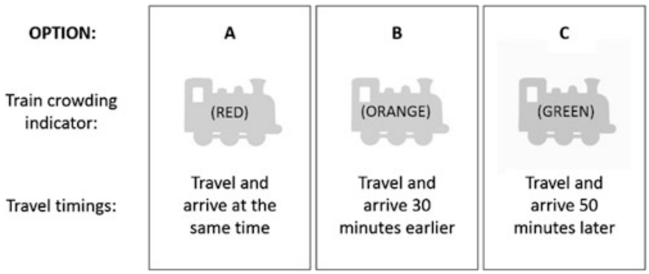


Fig. 1 An example set of stated preference options from the ‘standard’ survey (modified to suit black and white print)

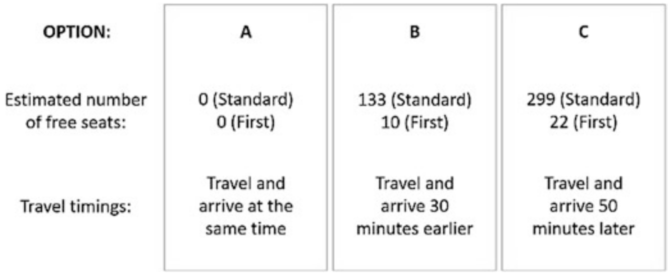


Fig. 2 An example set of stated preference options from the 'alternative' survey

3 A Summary of the Data Collected

Data were collected on a number of different Gatwick Express journeys over a 24 h period. The collected data were separated into four journey segments: between London Victoria (VIC) and Gatwick Airport (GTW) and vice versa and between Gatwick Airport and Brighton (BRI) and vice versa. The services on which data were collected are summarized in Table 1.

The survey software logged the start and end times of each response, which were matched with actual train running data (taken from www.realtimetrains.co.uk) and linked accordingly with one of the unique Segment IDs. A total of 115 responses were received by the Snap Survey hub during the pilot, although some of these were from initial testing of the tablets, and some of them were incomplete (largely because the participant did not give their consent to proceed). A valid response was defined as one which the participant gave their consent to proceed with the survey,

Table 1 A summary of the train services on which the surveys were conducted

Segment ID	Journey segment	Dep. time	Delay (min)	Arr. time	Delay (min)	Observed crowding
1	GTW-VIC	13:45	2	14:15	1	Busy
2	VIC-GTW	14:30	0	15:00	0	Busy
3	GTW-BRI	15:02	0	15:24	-1	Quiet
4	BRI-GTW	15:48	0	16:12	0	Moderate
5	GTW-VIC	16:15	0	16:45	6	Moderate
6	VIC-GTW	17:30	−1	18:07	0	Busy
7	GTW-VIC	07:49	6	08:20	13	Busy
8	VIC-GTW	08:45	−1	09:16	-4	Moderate
9	GTW-VIC	09:20	4	09:56	2	Busy
10	VIC-GTW	10:30	2	11:01	17	Moderate
11	GTW-BRI	11:02	22	11:24	22	Quiet
12	BRI-GTW	11:48	3	12:12	8	Quiet

and one which was successfully matched to one of the segments in Table 1. On this basis, 90 valid responses were collected, of which 53 were from participants who completed the ‘Standard’ survey and 37 were from participants who completed the ‘Alternative’ survey.

Figures 3 and 4 show the spread of ages of the participants and a breakdown of the data by gender respectively. Although there appears to be a bias towards male

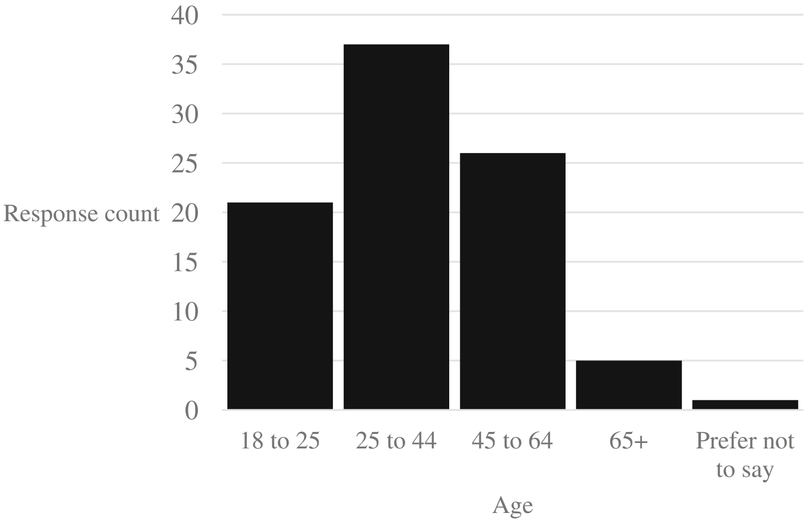


Fig. 3 The spread of the ages of the participants

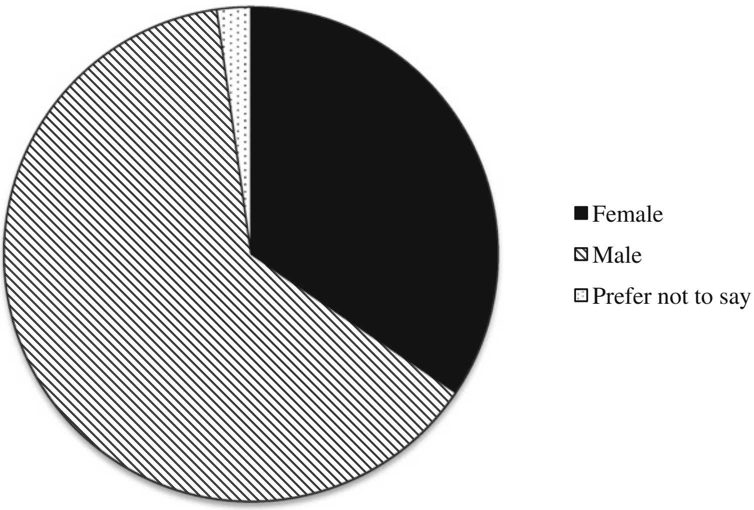


Fig. 4 A breakdown of the participants by gender

participants, Figs. 3 and 4 fit with general trends observed amongst UK rail passengers [1] and the data obtained could be said to be a representative sample demographically. However, amongst the participants, 61 % said that they were travelling for leisure, and 39 % for work. This is the opposite of the findings in the 2014 National Travel Survey, in which 61 % of rail journeys were said to be made for business or commuting. The fact that the Gatwick Express is an airport service is likely to be a factor here—59 % of respondents were either on their way to catch a flight or had just arrived on one, and 32 % of respondents said that they weren't UK residents.

The high number of non-UK residents could be used to infer a level of unfamiliarity with the system, which has benefits when investigating the clarity and effectiveness of information provision. However, the fact that the balance of trip purposes is atypical means that assumptions about the applicability of the results to UK rail travel as a whole ought to be made with caution.

A very high proportion (93 %) of participants were smartphone users. Of these, 81 % said that they use them for journey planning or travel updates. These figures fit with observed trends [2] and confirm that—in theory—disseminating train crowding information via mobile friendly websites or smartphone apps is plausible.

A five-point scale was used to determine each participant's perception of comfort on the train they were on: "Very uncomfortable", "Uncomfortable", "Neither comfortable nor uncomfortable", "Comfortable", "Very comfortable". During analysis, integer values were assigned to the responses from 1 ("Very uncomfortable") through to 5 ("Very comfortable"). Figure 5 shows how these perceptions compare with the loose observations made of crowding levels on each train (Table 1), and adds weight to the theory that crowding can adversely affect the experience of rail passengers.

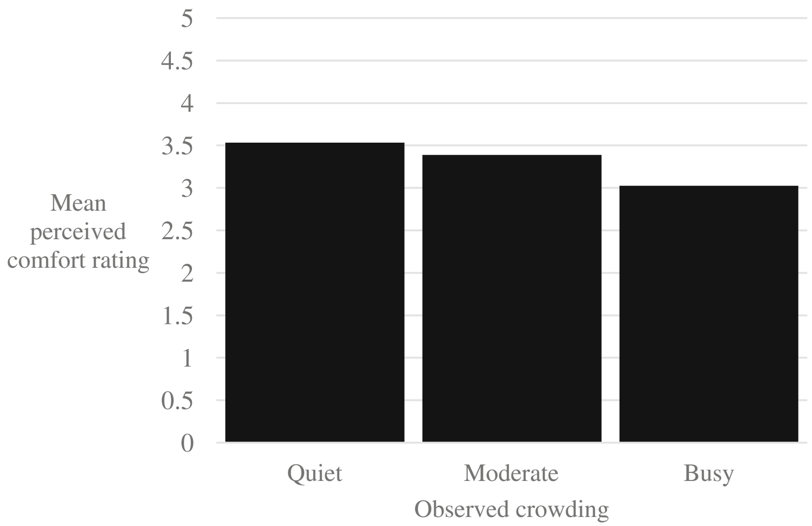


Fig. 5 Perceived comfort vs. observed crowding levels

4 Analysis of the Stated Preference Results

For each of the stated preference tests, an alternative-specific conditional logit model was developed, in order to estimate the probabilities of a passenger making a particular choice given information about crowding.

Microsoft SQL-Server was used to store the results and encode the data into an appropriate format, and Stata was used to fit a logit model in each case (http://www.stata.com/manuals13/rasclogit.pdf). Jack-knife resampling was specified to account for the fact that, because the data comprised several observations per participant, there may have been some correlation.

A stated preference exercise was marked as 'complete' if no choices were skipped; in the 90 responses analyzed here, there were 75 complete responses to the first stated preference exercise (SP1), and 80 complete responses to the second (SP2).

4.1 SP1—Information Provided Before Travelling

Of the 75 complete responses to this exercise, 44 were from the 'Standard' survey, in which the three crowding levels were represented graphically (Fig. 1), and 31 were from the 'Alternative' survey, in which crowding levels were represented textually, in terms of number of available seats on the train (Fig. 2).

The time offset for the earlier and later trains relative to the train with "ideal" timings was taken to be continuous and given in minutes, using an absolute value (i.e. it was always positive, whether the train was leaving earlier or later).

Crowding was described as the estimated number of available seats remaining on a train (coded as a percentage of the total number of seats on the train), and was also taken to be continuous.

A Boolean variable indicating whether the observation was made in the 'Standard' survey (=0) or the 'Alternative' survey (=1) was also specified in Stata as a case variable.

The probability of a user selecting one of the three trains, T, is given by:

$$P(T) = \frac{e^{U_T}}{\sum e^{U_i}}. \quad (1)$$

where Ui is the utility function of train i, given by:

$$U_i = \beta_{seats} s + \beta_{offset} t + \beta_{survey\_i} x + \beta_i. \quad (2)$$

s is the number of available seats on the train (as a proportion of the total)
t is the time offset in minutes
x represents whether the participant was using the 'Alternative' survey (= 1) or the 'Standard' survey (= 0)

β_{seats} is the crowding coefficient, constant for all trains
β_{offset} is the time coefficient, constant for all trains
β_{survey_i} is the survey coefficient, different for each train
β_i is a unique constant for each train.

Data was used to estimate the parameters for this model, and the output is given in Table 2.

It can be seen that the coefficients relating to the survey are not statistically significant, but that everything else is. This suggests that, in this case, the format of the crowding information was not important, but that crowding levels and relative times of departure and arrival are.

β_{offset} is negative, such that the utility function decreases with time offset; this is expected given that any deviation in travel plans is likely to be seen as a cost rather than a benefit. Similarly β_{seats} is positive, which is expected given that an increased number of available seats on a train is likely to be seen as a benefit.

$\beta_{LaterTrain}$ is negative, suggesting that participants have a general preference for not catching a later train; this is not unexpected, especially on an airport service.

$\beta_{EarlierTrain}$ is positive, suggesting that participants view catching an earlier train as inherently positive. This can be seen by considering the results of the first test; where all three trains are equally full and around 30 % of participants would still have chosen to travel and arrive an hour earlier. There is an implication, therefore, that they were not making their choice entirely based on crowding levels, and some revision of the survey may be necessary. For example, the use of the term '…arrive earlier' may have been read as a benefit whilst not taking in the implication that this would also mean leaving earlier. It is also thought that participants may have been biased if their train was delayed—those with a flight to catch may have wished that they had left earlier, for example. Anecdotally, comments received also suggest that if all services are known to be crowded, it becomes preferable to travel on an earlier service because crowding is perceived to contribute to delays.

Table 2 Parameters for the asc logit model for the first stated preference exercise

	Value	Std. err.	t statistic	$P > \|t\|$	95 % conf. intervals	
β_{offset}	−0.036	0.004	−8.68	0.000	−0.044	−0.028
β_{seats}	0.018	0.002	8.87	0.000	0.014	0.022
$\beta_{survey\_EarlierTrain}$	−0.246	0.197	−1.25	0.213	−0.634	0.141
$\beta_{EarlierTrain}$	1.123	0.191	5.87	0.000	0.747	1.499
$\beta_{survey\_LaterTrain}$	−0.364	0.260	−1.40	0.163	−0.874	0.147
$\beta_{LaterTrain}$	−0.467	0.219	−2.14	0.033	−0.897	−0.038
$\beta_{Survey\_SameTrain}$	0	(base alternative)				
$\beta_{SameTrain}$	0					

Table 3 Parameters for the asc logit model for the second stated preference exercise

	Value	Std. err.	t statistic	*P* > |t|	95 % conf. intervals	
β_{offset}	−0.096	0.011	−8.98	0.000	−0.118	−0.0754
β_{seats}	0.002	0.004	0.44	0.661	−0.007	0.0108
β_{survey}	0.069	0.232	0.30	0.768	−0.388	0.5247
$\beta_{LaterTrain}$	1.747	0.406	4.30	0.000	0.949	2.5443
$\beta_{SameTrain}$	0	(base alternative)				

4.2 SP2—Information Provided at the Station

The analysis of the second stated preference exercise followed a similar principle to the analysis of the first one, although in this case there were only two choices; it was assumed that a passenger arriving at a station would not be in a position to catch an earlier train (or that if they were, their decision would not necessarily be based around crowding issues). It was also assumed that a passenger would only wait for a later train if it was less crowded than the first one. The values for the logit model fitted to the data are given in Table 3.

As with the results of the first stated-preference survey, β_{survey} is not statistically significant, and β_{offset} is negative—which is again to be expected.

β_{seats} is not statistically significant, which implies that the tendency to wait for an emptier train is not especially dependent on how empty that train is. It is noted that, in this stated preference exercise, the later train was always less crowded and hence the positive value of $\beta_{LaterTrain}$ is likely to encapsulate some of the associated perceived benefits. Further analysis could be done in Stata, including generation of a model without a constant term.

Because this exercise only presented participants with two options, the model is binary and can be shown graphically in two dimensions (Fig. 6). Table 4 compares the model with the observed data, and it appears to be a good fit. It can be seen that the reported level of crowding on the later train does indeed make little difference, assuming that it is indeed less than the current train. It can also be seen that where there is a high service frequency, a high proportion of passengers might be prepared to wait for a less crowded train, although whether the stated preference is borne out in reality is likely to be dependent on a number of things—including whether the information provided is perceived to be reliable. It is noted that the surveys were conducted when participants were on a train, and not when they were actually waiting on a platform when their perceptions may have been different.

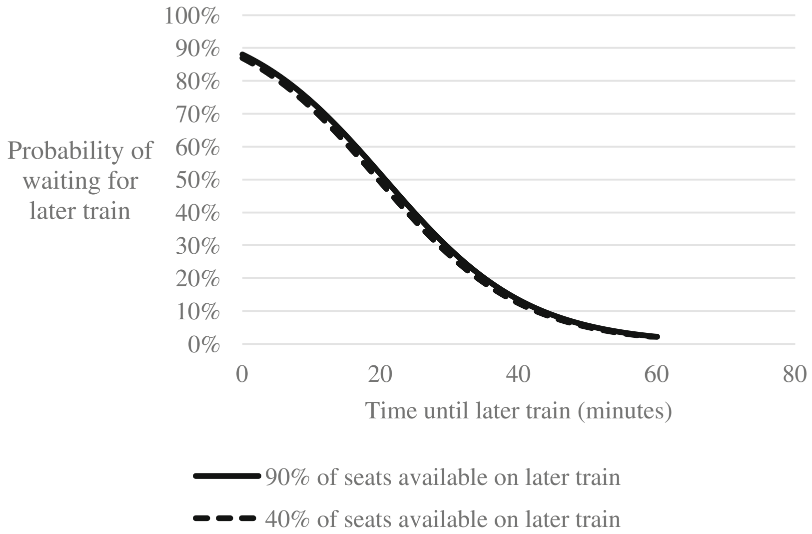


Fig. 6 Predicted probabilities of a passenger at the station waiting for a less crowded train if the current train is full

Table 4 Observed data compared with the predictions from the model

Time to later train (min)	Proportion of seats available on later train (%)	Observed % of respondents who would wait for later train ('Standard' survey).	Observed % of respondents who would wait for later train ('Alternative' survey).	Predicted proportion of people who would wait for later train
15	90	67	58	63
	40	59	61	61
30	90	29	29	29
	40	22	23	27
45	90	6	6	9
	40	6	19	8

5 Conclusions

It was suggested at the beginning of this paper that rail passengers disliked overcrowding, and might change their travel behavior if better information about levels of crowding were provided. The pilot study discussed here supports those suggestions, with the findings showing that perceived comfort decreased with increased crowding levels, and a large proportion of respondents stating that they might choose an earlier or later train if presented with information about crowding.

Two different methods of presenting crowding information were trialed, and although there did not appear to be any statistical significance between the responses received further research would need to be undertaken. Concerns were raised in both cases that some of the information presented may not have been clear, and the tendency to choose an earlier train in all circumstances implies that crowding concerns may not have been the main factor when making a choice; this casts doubt on whether the stated preferences would be borne out in reality. Having completed the pilot survey, a further detailed survey is to be undertaken, with a redesigned questionnaire to alleviate some of the concerns discussed here. Further analysis will also be undertaken to ascertain the types of passenger most likely to change their travel habits.

Acknowledgments The author would like to thank RSSB and Association of Train Operating Companies (ATOC) for funding the research, and Govia Thameslink Railway (GTR) for their full co-operation in allowing survey work to take place.

References

1. Rail Executive.: Rail users factsheet 2014/15 (2015)
2. Office of Rail Regulation.: Rail passenger experience report (2014)

Part II
Aviation—Human Factors in Aviation

Task Demand Variation in Air Traffic Control: Implications for Workload, Fatigue, and Performance

Tamsyn Edwards, Cynthia Gabets, Joey Mercer and Nancy Bienert

Abstract In air traffic control, task demand and workload have important implications for the safety and efficiency of air traffic, and remain dominant considerations. Within air traffic control, task demand is dynamic. However, research on demand transitions and associated controller perception and performance is limited. This study used an air traffic control simulation to investigate the effect of task demand transitions, and the direction of those transitions, on workload, fatigue and efficiency performance. A change in task demand appeared to affect both workload and fatigue ratings, although not necessarily performance. In addition, participants' workload and fatigue ratings in equivalent task demand periods appeared to change depending on the demand period preceding the time of the current ratings. Further research is needed to enhance understanding of demand transition and workload history effects on operator experience and performance, in both air traffic control and other safety-critical domains.

Keywords Workload transitions · Workload history · Air traffic control · Fatigue · Time based metering · Task demand transitions

T. Edwards (✉) · C. Gabets · N. Bienert
San Jose State University, NASA Ames Research Center,
Moffett Field, San Jose, CA 94035, USA
e-mail: tamsyn.e.edwards@nasa.gov

C. Gabets
e-mail: cynthia.gabets@nasa.gov

N. Bienert
e-mail: nancy.bienert@nasa.gov

J. Mercer
NASA Ames Research Center, Moffett Field, Moffett Field, San Jose, CA 94035, USA
e-mail: joey.mercer@nasa.gov

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_8

1 Introduction

Within the safety critical domain of air traffic control (ATC), workload “is still considered one of the most important single factors influencing operators’ performance” [1, p. 639]. Workload has been defined within the ATC domain as the result of an interaction between task demand and the controllers’ selected strategy [2]. The association of workload and controller performance has important implications for the safety and efficiency of air traffic (e.g. [3, 4]). Workload therefore remains a dominant consideration.

In ATC, as with many other safety critical environments, task demand and workload are dynamic. Air traffic controllers (ATCOs) frequently experience changes in traffic load and the complexity of the traffic situation, potentially resulting in the experience of transitions between high and low workload. These transitions can be expected by the controller, such as when traffic load changes based on the time of day or known activities in surrounding sectors, or unexpected, for example, through increased complexity resulting from an emergency situation. Transitions may also be gradual or sudden [5].

Research on demand transitions, and the effect on both performance-influencing covariate factors (such as workload and fatigue) and task performance is limited however, with studies frequently utilizing a constant task demand or workload [6]. Of the research available, there appears to be conflicting findings. Some (e.g. [7]) have reported that overall performance efficiency on a vigilance task was not affected by task demand transitions, regardless of whether the transition was expected or unexpected. However, others (e.g. [8]) have found that performance on vigilance tasks was influenced by a low to high demand transition or high-to-low demand transition (e.g. [5]). Task demand and workload transition research specific to an ATC environment is particularly underrepresented. Consequently, there is limited understanding of the influence of workload transitions on performance in an air traffic environment.

The aim of this study was to investigate the influence of expected and gradual task demand transitions (high-low-high and low-high-low) on workload, fatigue and performance, within a high fidelity ATC simulation environment. Due to the quantity of measures and data generated from this study, only a subset of the measures and findings that are most relevant to the research aim are presented.

2 Method

2.1 Design

An en-route ATC human in the loop (HITL) simulation was utilized to investigate task demand variation on workload, fatigue, and performance. Efficiency-related performance was inferred from aircraft delay (in seconds) at a specific point in the arrival sequence (three nautical miles before the meter fix). Participants were eight ex-ATCOs who had previously worked in enroute airspace in Oakland Air Route Traffic Control Center (ARTCC). Pseudo pilots were paired with controllers, and completed standard pilot tasks such as controlling the aircraft in accordance with controller instructions and communicating with controllers. Participants operated a combined low and high altitude sector, and were assigned to meter aircraft into the northeast corner of the Phoenix Terminal Radar Approach Control (TRACON). This airspace was selected for the complex mix of arrivals and overflights.

The study used a within-measures design. The direction of the task demand transition was manipulated to create two scenarios. Scenario 1 followed a high-low-high task demand pattern and scenario two followed a low-high-low task demand pattern. The creation of three task demand periods was implemented in order to better reflect the multiple task demand transitions that can be experienced within an operational environment. In addition, this permitted an extension of previous studies that had focused on the comparison of workload and performance between one transition period (e.g. [5]). Each simulation session lasted for 90 min and consisted of three, 20 min [9] periods of stable task demand which alternated between high and low, interspersed with a total of three, 10 min transition phases. Task demand was created by the number of aircraft under control [10] as well as the ratio of arrival aircraft and overflights. Arrival aircraft create complexity in the task, which influences task demand. Task demand phases for equivalent stable task demand periods (i.e. high demand regardless of which scenario the high demand was positioned in) were created using the same aircraft counts and number of arrival aircraft, permitting comparability between demand variation scenarios. Scenarios followed a counterbalanced presentation. Participants were required to complete all control actions and meter aircraft to arrive at a meter fix at a scheduled time. Participants were provided with a 1 h briefing prior to the start of the study, and six training runs (four 90 min training runs and two 45 min training runs).

2.2 Participants

A total of eight male ex-controllers took part in the simulation. Age ranged from 50 to 64 years. Participants responded to grouped age ranges and so an average age could not be calculated. Participants had worked as en-route controllers in the Oakland, California, ARTCC. Participants' years of experience as active ATCOs (excluding training) ranged from 22 to 31 years (M = 26.56, SD = 3.90).

2.3 Measures and Apparatus

Covariate factors were measured using subjective, self-report scales. Mental workload was measured using the uni-dimensional Instantaneous Self-Assessment

(ISA) scale which measures workload from 1-6 [11]. Every three minutes, participants were presented with the ISA rating scale at the top of the radar scope and asked to click on the workload rating. Fatigue was measured using the Samn-Perelli scale, which ranges from 1 to 7 with behavioral anchors at each point on the scale. [12]. Fatigue measures were taken three minutes into, and three minutes prior to the end of, each 20 min task demand phase and six minutes into each 10 min transition phase. This periodicity was selected to capture data across each stable task demand period, and refined based on results from three pilot studies. Performance was assessed by aircraft delay at three nautical miles from the meter fix point as a measure of participants' efficiency-related performance. An aircraft that is on-time, i.e. without delay, suggests optimal performance. An efficiency-related performance measure was selected for analysis as opposed to a safety-related performance measure as previous research suggests that controllers can maintain safety-related performance without significant observed changes even under high periods of demand by applying workload management strategies [13]. Changes in performance are frequently first observed in efficiency-related tasks (e.g. [13]). An efficiency-related task was therefore potentially more sensitive to changes in performance.

The software used was the Multi-Aircraft Control System (MACS) [14]. Participant workstations were configured with a BARCO large-format display and keyboard/trackball combination that emulates what is currently used in en-route air traffic control facilities. Voice communications between ATCOs (the participant and a non-participant controller controlling neighboring sectors) and the pilot were enabled via a custom, stand-alone system. Datalink communications were also available. Data were collected continuously through MACS's data collection processes.

3 Results

3.1 Analysis Approach

Due to the quantity of analyses and findings, only the data trends most relevant to the research aims are presented in this paper. To address the research aim, comparisons of the three 20 min task demand periods per scenario are presented in the following sections.

For each workload, fatigue and the dependent variable of aircraft delay in arrival, descriptive statistics were first reviewed, followed by further exploration through the application of two repeated measures ANOVAs—one for each task demand transition scenario (scenario 1: high-low-high demand; scenario 2: low-high-low demand). The decision to apply separate repeated measures one-way ANVOAs was made based on a review of previous research analysis approaches to similar experimental designs (e.g. [5]) and research aims. The research aim of this study

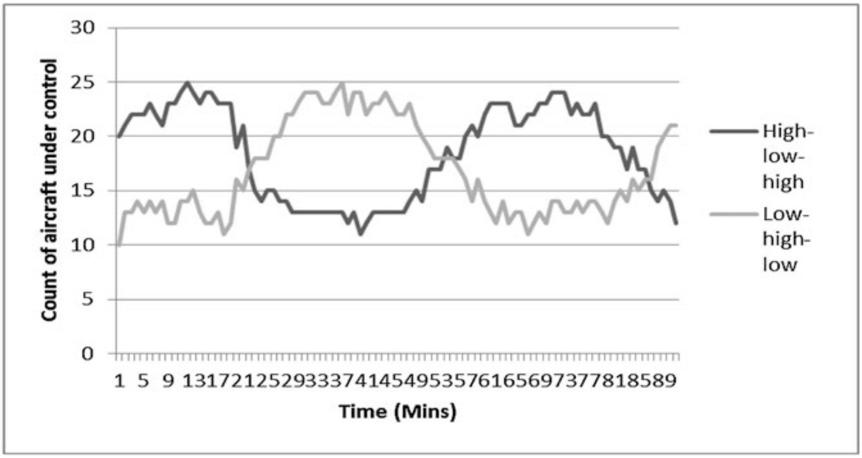


Fig. 1 Count of aircraft under control by minute for scenario 1 (*high-low-high demand*) and scenario 2 (*low-high-low demand*)

focused on investigating the effect of task demand on covariate and performance variables, including the direction of the task demand. One way ANOVAs permitted the exploration of changes within each task demand scenario. Prior to all inferential statistics, data were checked for normality and sphericity violations. Unless otherwise reported, all data met these assumptions.

3.2 Task Demand Variation Manipulation Check

A review of the descriptive statistics suggests that task demand did vary in the intended direction (Fig. 1). Figure 1 confirms that the number of aircraft in the controller's sector were similar between equivalent task demand periods regardless of scenario (high-low-high demand or low-high-low demand). The number of arriving aircraft was also similar.

3.3 Task Demand and Subjectively Experienced Factors

Task Demand and Workload

Workload ratings were averaged across the 20 min periods of stable task demand for analysis to facilitate comparison between the separate task demand periods. A review of the descriptive statistics (Table 1) suggest that workload in both

Table 1 Mean and standard deviation for workload (as rated by ISA) in both scenario 1 and scenario 2, averaged across 20 min task demand periods

Workload (ISA)	Task demand period 1 (0–20 min)		Task demand period 2 (31–50 min)		Task demand period 3 (61–80 min)	
	M	SD	M	SD	M	SD
Scenario 1 workload (High-low-high)	3.67	0.77	2.87	0.61	3.85	0.62
Scenario 2 workload (Low-high-low)	2.78	0.64	4.06	0.71	3.33	0.61

scenarios varied as expected with task demand. In scenario 1 (high-low-high demand) workload appears to be rated slightly higher in the third task demand period (high demand) compared to the first task demand period (high demand). In scenario 2 (low-high-low demand), workload was rated highest in the high demand, second task demand phase. However, on average, participants rated perceived workload to increase in the third task demand period (low demand) compared to the first low demand period. Comparing between scenario 1 and 2, the high demand period is perceived to generate the most workload for participants in the low-high-low demand scenario, although the high demand periods were objectively equivalent between scenarios. Comparing across low demand periods between conditions, workload is rated similarly in the first period of scenario 2 and the middle period of scenario 1. However, the low demand period in the third period of scenario 2 is rated as higher workload than either of the other low demand periods. To further examine the changes in perceived workload, a one-way repeated measures analysis of variance (ANOVA) was conducted for each scenario [5]. A one-way ANOVA was applied to each scenario, to explore differences within-scenarios. In relation to scenario 1 (high-low-high demand) a significant main effect of task demand period was found on self-reported workload $F(2,14) = 44.23$, $p < 0.001$. Pairwise comparisons revealed that workload was significantly lower in task demand period 2 (low demand) than high task demand period one ($p < 0.005$) and three ($p < 0.001$). Workload was not rated significantly differently between high demand period 1 and high demand period 3 ($p = 0.68$). In scenario 2 (low-high-low demand) a significant main effect of task demand period was found on self-reported workload $F(2,14) = 32.72$, $p < 0.001$. Pairwise comparisons revealed that workload was rated significantly higher in the high demand period than the first low demand period ($p < 0.001$) and second low demand period ($p < 0.005$). It was also identified that the workload ratings in the second low demand period were significantly higher than the first low demand period ($p < 0.05$).

Task Demand and Fatigue

A review of the means of reported fatigue for each task demand period in scenario 1 (high-low-high demand) revealed that ratings of fatigue appeared similar between high demand period one (M = 2.23, SD = 0.71) and low demand period one (M = 2.15, SD = 0.77) (Fig. 2). Fatigue ratings were slightly higher in the third demand period, high demand period two (M = 2.70, SD = 1.08). Conversely, in

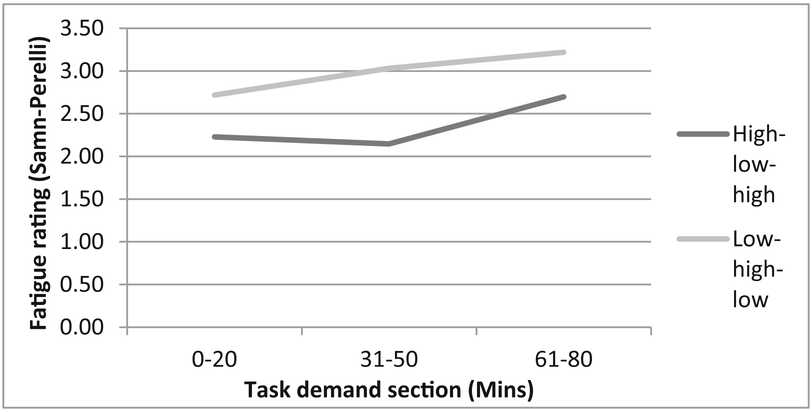


Fig. 2 Fatigue ratings (as measured by Samn-Perelli scale) averaged across 20 min task demand periods for scenario 1 (*high-low-high demand*) and scenario 2 (*low-high-low demand*)

scenario 2 (low-high-low demand) fatigue ratings appeared to increase across each task demand period (first low task demand period: M = 2.71, SD = 1.01; first high task demand period: M = 3.03, SD = 1.42; second low task demand period: M = 3.22, SD = 1.54) (Fig. 2).

A one way ANOVA was utilized to explore the effect of task demand on fatigue ratings for both scenarios. In scenario 1 (high-low-high demand) Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 9.44$, $p < 0.01$. When considering this main effect, therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (E = 0.56). No significant differences between fatigue ratings were identified $F(1.12, 7.81) = 2.48$, $p > 0.05$. Differences between fatigue ratings in scenario 2 (low-high-low demand) approached significance, $F(2,14) = 3.40$, $p < 0.1$. A further review of the descriptive data revealed that averaging across the two fatigue measures per task demand period (one three minutes into the period, and one three minutes before the end of the period) may be masking the effect of task demand on fatigue. Participants' fatigue rating was frequently lower for the first measurement compared to the second measurement of the task demand period. Therefore, ANOVAs were repeated on two fatigue measurements per workload period. In scenario 1 (high-low-high demand) Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(14) = 26.82$, $p < 0.05$. When considering this main effect, therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (E = 0.44). No significant differences between fatigue ratings were identified $F(2.18, 15.22) = 2.82$, $p > 0.05$. The ANOVA applied to scenario 2 revealed a main effect of task demand on fatigue ratings $F(5,35) = 2.69$, $p < 0.05$. Pairwise comparisons revealed that fatigue ratings were significantly lower at the

Table 2 Mean and standard deviation for arrival aircraft delay (in seconds) in both scenario 1 and scenario 2, averaged across 20 min task demand periods

Arrival aircraft delay (secs)	Task demand period 1 (0-20 min)		Task demand period 2 (31-50 min)		Task demand period 3 (61-80 min)	
	M	SD	M	SD	M	SD
Scenario 1 Aircraft delay (High-low-high)	13.88	5.32	7.70	3.6	−1.71	6.92
Scenario 2 Aircraft delay (Low-high-low)	10.48	3.07	9.93	2.54	7.50	4.86

first fatigue measurement of the first low task demand period (M = 2.63, SD = 1.06) compared to fatigue ratings in the second low task demand period (first fatigue measurement M = 3.13, SD = 1.46, $p = 0.05$; second fatigue measurement M = 3.31, SD = 1.65), ($p < 0.05$). No other differences were significant.

Task Demand and Performance
A review of the average delay across 20 min task demand periods in scenario 1 (high-low-high demand) (Table 2) suggests that participants reduced average aircraft delay across the task demand periods until aircraft were arriving early in the final task demand period (Table 2). The same pattern was seen in scenario 2 (low-high-low demand), although smaller reductions in delay are observed (Table 2). However, in both scenarios, performance variability appears to increase in the final task demand period, indicated by comparatively large standard deviations (Table 2). Data in scenario 1 (high-low-high demand) were further examined with a repeated measures ANOVA. A significant main effect of task demand period was found on arrival delay $F(2,14) = 12.84$, $p < 0.005$. Pairwise comparisons revealed that aircraft delay was significantly longer in the first high demand period than the first low demand period ($p < 0.05$) and the second high demand ($p < 0.01$). Delay was also significantly longer in the first low demand period than the second high demand period ($p < 0.05$). Data in scenario 2 (low-high-low demand) were also further examined with a repeated measures ANOVA. No significant differences in arrival aircraft delay were identified $F(2,14) = 3.04$, $p > 0.05$.

4 Discussion

A within-measures design was used to investigate task demand variation on workload, fatigue, and performance. The direction of the task demand transition was manipulated to create two scenarios: scenario 1—high-low-high demand; scenario 2—low-high-low demand. Results showed that task demand varied as intended. Descriptive statistics confirmed that equivalent demand periods, regardless of scenario or position, were composed very similarly in terms of controlled aircraft count and arrival aircraft count. This suggests that changes in the covariates or dependent variable are unlikely to be attributed to demand differences between the created scenarios.

As expected, a change in task demand appears to affect both workload and fatigue ratings. Significantly different workload and fatigue ratings were reported within scenario, across task demand periods. However, a key finding of interest is that perception of workload and fatigue appear to differ depending on the demand period preceding the current ratings, in line with previous findings [5]. This finding is observed in the average workload ratings for each task demand period within scenarios (Table 1). In the first scenario (high-low-high task demand), workload is not perceived significantly differently between the first and second high task demand periods. Workload is rated as significantly lower during the low demand period compared to the high demand periods, however. Comparatively, in scenario 2 (low-high-low demand) workload is perceived to be significantly greater in the second low demand period than the first, potentially suggesting that workload is perceived to be greater after the high demand period. This increased workload would not be the result of working to resolve delays from the previous period, as any remaining delays were absorbed in the 10 min transition period between the stable demand periods. In addition, it is interesting to note that workload was perceived to be higher in the high demand period of scenario 2 than either of the high task demand periods in scenario 1, suggesting that the preceding low demand may have impacted the perception of workload of the high demand period in scenario 2.

These findings indicate that the workload appears to be perceived differently depending on what precedes the time of rating. More specifically, results suggest that in this ATC task, a demand transition pattern of low-high-low demand may result in operators perceiving subsequent high and low demand periods after the initial low demand period as generating a greater workload than equivalent demand periods in a high-low-high demand transition pattern. A similar pattern of findings was seen in participants' fatigue ratings. In scenario 1 (high-low-high demand), fatigue ratings were not significantly different between demand phases. Fatigue ratings did increase in the final high demand period, although not significantly. In contrast, in scenario 2, participants reported on average that fatigued increase with each subsequent task demand period.

Although there is a lack of common agreement regarding the mechanisms by which task demand transitions may impact covariate factors [15], this collection of workload and fatigue findings may be interpreted in the context of Limited Resource theory [16] and arousal theories. Potentially, in scenario 1, the low demand period may have enabled controllers to use this time to recover resources and prepare for the next high task demand period. [17] has previously documented that this is an active control strategy that controllers use during low demand periods, when it is considered safe to do so. This recovery period may then limit the increase of perceived fatigue in the final high task demand period. Arousal theories may provide some insight into why this effect may not be seen in the low-high-low demand transition pattern. Arousal theories suggest that low workload (or underload) may lead to lower arousal, which may limit attentional resources and create boredom and lack of motivation. If a human operator started a task from this point, it may be that the following demand periods are perceived to be more demanding or fatiguing. By the final low demand period, the operator may find it difficult to pay attention. Attentional resources theories suggest however that if preceded by a higher demand, lower demand periods can be utilized to replenish attentional resources, not necessarily reducing arousal to a level that would create negative effects. The application of these theories therefore potentially account for the disparate findings between the different task demand transition patterns.

Performance did not appear to be negatively affected in relation to task demand variation, with delay times reducing across task demand periods within each scenario. Performance variability did increase however across task demand period, as inferred from increasing standard deviations. This pattern of findings for performance measures has also been documented previously, although for vigilance-based performance [7]. The finding of improved aircraft arrival time may be the result of controllers applying strategies to support performance across the demand periods [18].

Although controller strategies were not a direct focus of this research, this finding highlights an important issue for future research considerations. Although this measure of performance indicates that performance in terms of aircraft arrival time was maintained, and even improved, in scenario 2, controllers also reported greater perception of workload and fatigue. It is therefore possible that controllers may have experienced having to work harder to maintain performance, even though this was not observable in the performance measure itself. This result emphasizes that in order to detect, and prevent, performance declines, further research should focus on measures that are sensitive to the operators' experience, and that can be monitored and utilized to detect potential performance decline prior to a performance related incident.

It is acknowledged that these results are provisional, and results need to be interpreted within context. For example, in an air traffic environment, it is easier for the controller to build a picture of the traffic by ramping up with the traffic rather than just starting a session in a high demand period [17]. However, findings do have important implications for the prediction of controller performance in an operational environment. Findings suggest that high and low demand periods can affect controller perception of covariate factors such as workload and fatigue differentially depending on what has happened prior to the current situation. Thus, supervisors may need to pay close attention to the number and direction of transitions that a controller experiences per session to most effectively support controller performance.

Future research should further explore the relationship between previous task demands and the relationship on present controller experience, including the exploration of sudden, and unexpected, transitions. Better predictions are needed to identify and prevent potential performance declines and associated performance-related incidents. Such predictions may be particularly relevant for adaptive automation technologies that support operator performance.

5 Conclusion

The effect of task demand transitions on covariate factors of workload and fatigue and one efficiency related performance measure was investigated within the context of an air traffic control task. Initial findings suggest that task demand variations affected participants' perceptions of workload and fatigue, although the effect appeared to be influenced by the direction of the previous demand periods. Performance appeared to be maintained across the control session. Previous research has infrequently considered transitions of task demand in an applied environment. Findings are consistent with the description of workload history effects [5], and that equivalent task demand periods can elicit different experiences for a human operator depending on what precedes the time of rating. Attentional resource and arousal theories appear to support interpretation of the results. Further research is required to enhance understanding of demand transition and history effects. Practical applications include guidance for operations room supervisors, and implications for predictions of performance in high and low demand periods, with important implications for identifying and preventing potential performance declines and associated performance-related incidents.

References

1. Di Nocera, F., Fabrozo, R., Terenzi, M., Ferlazzo, F.: Procedural errors in air traffic control: effects of traffic density, expertise, and automation. Aviat. Space Envir. Md. **77**, 639–643 (2006)
2. Djokic, J., Lorenz, B., Fricke, H.: Air traffic control complexity as workload driver. Transport Res. C-Emer. **18**, 930–936 (2010)
3. Redding, R.E.: Analysis of operational errors and workload in air traffic control. In: Human Factors and Ergonomics Society Annual Meeting Proceedings, vol. 36(17), pp. 1321–1325. Human Factors and Ergonomics Society (1992)
4. Schroeder, D.J., Nye, L.G.: An Examination of the Workload Conditions Associated with Operational Errors/Deviations at Air Route Traffic Control Centers. Federal Aviation Administration Office of Aerospace Medicine, Washington, DC (1993)
5. Cox-Fuenzalida, L.E.: Effect of workload history on task performance. Hum. Factors **49**, 277–291 (2007)
6. Hancock, P.A., Williams, G., Manning, C.M.: Influence of task demand characteristics on workload and performance. Int. J. Aviat. Psychol. **5**, 63–86 (1995)
7. Helton, W.S., Shaw, T., Warm, J.S., Matthews, G., Hancock, P.: Effects of warned and unwarned demand transitions on vigilance performance and stress. Anxiety Stress Copin. **21**, 173–184 (2008)

8. Moroney, B.W., Warm, J.S., Dember, W.N.: Effects of demand transitions on vigilance performance and perceived workload. In: Proceedings of the Human Factors And Ergonomics Society Annual Meeting 39, pp. 1375–1379. SAGE Publications, London (1995)
9. Galster, S.M., Duley, J.A., Masalonis, A.J., Parasuraman, R.: Air traffic controller performance and workload under mature free flight: conflict detection and resolution of aircraft self-separation. Int. J. Aviat. Psychol. **11**, 71–93 (2001)
10. Tenney, Y.J., Spector, S.L.: Comparisons of HBR models with human-in-the-loop performance in a simplified air traffic control simulation with and without HLA protocols: task simulation, human data and results. In: Proceedings of the 10th Conference on Computer Generated Forces and Behaviour Representation, Norfolk, VA (2001)
11. Tattersall, A.J., Foord, P.S.: An experimental evaluation of instantaneous self-assessment as a measure of workload. Ergonomics **39**, 740–748 (1996)
12. Samn, S.W., Perelli, L.P.: Estimating aircrew fatigue: a technique with application to airlift operations (No. SAM-TR-82-21). School of Aerospace Medicine Brooks Afb Tx (1982)
13. Edwards, T.: Human performance in air traffic control. PhD dissertation, University of Nottingham (2013)
14. Prevot, T., Lee, P., Callantine, T., Mercer, J., Homola, J., Smith, N., Palmer, E.: Human-in-the-Loop evaluation of nextgen concepts in the airspace operations laboratory. In: Proceedings of the AIAA Modeling and Simulation Technologies Conference, AIAA 2010-7609, Toronto, Canada (2010)
15. Wickens, C.D., Mabor, A.S., McGee, J.P.: Flight to the Future: Human Factors in Air Traffic Control. National Academy Press, Washington DC (1997)
16. Wickens, C.D.: Engineering Psychology and Human Performance. Harper Collins, New York (1992)
17. Edwards, T., Sharples, S., Kirwan, B., Wilson, J.R.: Identifying markers of performance decline in air traffic controllers. In: Di Bucchianico, G., Vallicelli, A., Stanton, N.A., Landry, S.J. (eds.) Human Factors in Transportation: Social and Technological Evolution Across Maritime, Road, Rail, and Aviation Domains (in press)
18. Sperandio, J.C.: Variation of operator's strategies and regulating effects on workload. Ergonomics **14**, 571–577 (1971)

How Important Is Conflict Detection to the Conflict Resolution Task?

Joey Mercer, Cynthia Gabets, Ashley Gomez, Tamsyn Edwards, Nancy Bienert, Lauren Claudatos and Jeffrey Homola

Abstract To determine the capabilities and limitations of human operators and automation in separation assurance roles, the second of three Human-in-the-Loop (HITL) part-task studies investigates air traffic controller's ability to detect and resolve conflicts under varying task sets, traffic densities, and run lengths. Operations remained within a single sector, staffed by a single controller, and explored, among other things, the controller's conflict resolution performance in conditions with or without their involvement in the conflict detection task. Whereas comparisons of conflict resolution performance between these two conditions are available in a prior publication, this paper explores whether or not other subjective measures display a relationship to that data. Analyses of controller workload and situation awareness measures attempt to quantify their contribution to controllers' ability to resolve traffic conflicts.

Keywords Human factors · Air traffic control · Human-in-the-loop simulation · Function allocation · Human-automation interaction

J. Mercer (✉) · J. Homola
NASA Ames Research Center, Moffett Field, San Jose, CA 94035, USA
e-mail: joey.mercer@nasa.gov

J. Homola
e-mail: jeffrey.r.homola@nasa.gov

C. Gabets · A. Gomez · T. Edwards · N. Bienert · L. Claudatos
San Jose State University, NASA Ames Research Center, Moffett Field, San Jose, CA 94035, USA
e-mail: cynthia.gabets@nasa.gov

A. Gomez
e-mail: ashley.n.gomez@nasa.gov

T. Edwards
e-mail: tamsyn.e.edwards@nasa.gov

N. Bienert
e-mail: nancy.bienert@nasa.gov

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_9

1 Introduction

The transition to NextGen will likely include increasing levels of automation to help controllers perform their duties. A progression towards higher levels of automation could enable the controllers' working environment to move from tactical separation management to strategic decision-making. Such automation is envisioned to expand performance beyond today's limits by off-loading workload from controllers onto automated functions for the majority of routine operations [1]. However, the nature of this human-automation team is not well understood. It is still unknown exactly which tasks are best allocated to the human operator as opposed to the automation, and vice versa. In considering this system as a whole, careful and thorough investigation is needed to better understand, not only how each team member performs in such environments, but also any associated human-automation cooperation issues.

2 Background

The motivation behind these investigations is to address a well-known problem: current-day air traffic control techniques are very labor intensive, and are limited to the amount of information controllers can process and keep in their working memory. Function allocation is but one approach to this problem, wherein automation can take responsibility for some tasks, theoretically easing the controller's workload.

The current series of studies fall under NASA's revised function-allocation research plan, which calls for advancing our understanding of the related air-ground and human-automation issues. In particular, the Airspace Operations Laboratory (AOL) focused on the following question: "Which separation assurance functions can air traffic controllers effectively perform in future air traffic management systems?" Understanding the strengths and weaknesses of individual team members is an important aspect in determining how to distribute tasks between team members. As a first step towards gaining such insights into human-automation teaming, our approach has been to conduct part-task HITL simulations that identify the capabilities and limitations of the controller in key separation assurance tasks.

2.1 Function Allocation Research

In March of 2015, the AOL at NASA's Ames Research Center [2] conducted the first in a series of studies that explored the capabilities and limitations of human operators with regard to the separation assurance element of air traffic control. Specifically, the research sought to better understand how best to allocate functions between controllers and automation. A second study conducted in May of 2015 continued that work, but with the conflict resolution task as its main focus. Although the first study differed in that it investigated the conflict detection task, both studies shared the same approach, in which they sought to tease apart the primary task from related secondary tasks. While looking across varying levels of automation, the studies measured the overall impact on the performance of the primary task. Of particular interest to the second study was discovering whether removing controllers' involvement in the detection task would impact their ability to resolve conflicts.

The first study, referred to as the Human-Automation Conflict Detection study (or HACD), and the second study, referred to as the Human-Automation Conflict Resolution study (or HACR), are reported in [3–5]. However, a brief summary of the HACR simulation environment is in order, to provide the appropriate context for the discussions of this paper.

The HACR Simulation. HACR examined controller performance on the conflict resolution task under different task sets, traffic density levels, and run lengths. The group of tasks under the controller's responsibility and those under the automation's responsibility defined a given task set. Traffic density and run length completed the study's set of independent variables. Although the full study featured a 5 × 2 × 2 within-subject repeated-measures design, the scope of this paper and its analyses are limited to two of the study's task sets (*Conflict Resolution* and *Conflict Detection & Resolution*.), both traffic densities (1x current-day traffic levels and 1.2x current-day traffic levels) and one run length (60 min).

Clearly, the key distinction between the two task sets of interest lies in whether or not the controller was responsible for the conflict detection task. The *Conflict Detection & Resolution* condition operated much like current-day air traffic control. The controller kept constant watch over their sector's radar display, observing the progress of air traffic in and around their sector, and issuing control instructions they deemed necessary. In contrast, the *Conflict Resolution* condition went to great lengths to isolate the conflict resolution task, and in doing so, removed the controller from the conflict detection task. The study accomplished such isolation by developing a clever display capability that suppressed all air traffic from the radar display unless the automation (i.e., a trajectory-aided conflict probe) detected a potential conflict. Once the automation detected a conflict, the system would turn off the 'blackout' mode, and displayed all traffic as it normally would, albeit with the aircraft in conflict highlighted (see Fig. 1). At this point, the automation's task of detecting the conflict was complete, and it was then the controller's responsibility to, just as in the *Conflict Detection & Resolution* condition, issue whatever control instructions they deemed appropriate.

The airspace used during the simulation consisted of a single high-altitude sector, with a mix of overflights passing through at level altitudes, and transitioning aircraft descending to or climbing out from area airports. The scenarios progressed through a ramp-up, peak, and ramp-down phase, with each phase lasting

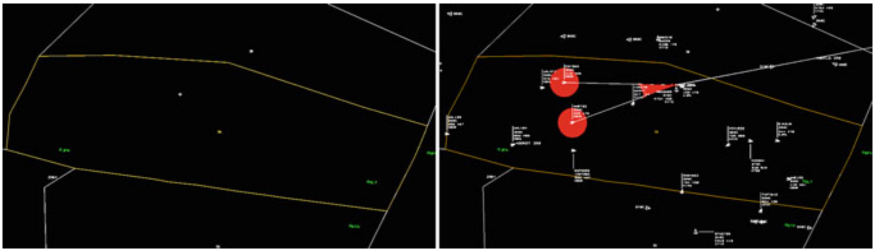


Fig. 1 Screen capture of the controller's radar display in the Conflict Resolution condition before the automation detects a conflict (*left*), and after the automation detects a conflict (*right*)

approximately 20 min. Traffic levels reached 18 aircraft in the sector in the 1x traffic density, and 22 aircraft in the 1.2x density. The simulation's environment also included winds for the area, which were constant-at-altitude with a nominal forecast error. Eight retired FAA en route controllers (with an average of 24.9 years of experience among them) participated in the study, all of which worked the same conditions.

The primary simulation platform used for the study was the Multi Aircraft Control System (MACS) [2], which, for each controller workstation, hosted an En Route Automation Modernization (ERAM) emulation on a large-format monitor. The controller workstation also included a specialized keyboard and trackball, similar to those used in current air traffic control facilities, as well as a custom, stand-alone voice application emulating the fielded communication system. Data recorded and collected at each workstation included aircraft flight states, operator task data and workload, automation states, voice communications, etc.

2.2 *Previous Findings*

The data presented in [4] compared the time at which the controllers issued a clearance to resolve a conflict, with the time of that conflict's detection. In the *Conflict Detection & Resolution* condition, the detection time was marked when the controller made a keyboard entry to signal they believed an aircraft pair to be in conflict. In the *Conflict Resolution* condition, the detection time was marked when the automation identified an aircraft pair to be in conflict (i.e., typically when the 'blackout' mode turned off). The difference between these two event times represents the Resolution Response Time measurement.

The findings showed that the controllers were able to issue resolution maneuvers within 30 s of conflict detection for 49 % of cases in the *Conflict Resolution* condition, but did so for 59 % of cases in the *Conflict Detection & Resolution* condition. Even after accounting for the traffic density variable, this trend held true:

Table 1 Resolution response time mean and standard deviation values (in seconds) from the CR and CD&R task sets in both 1x and 1.2x densities

Resolution response time	CR		CD&R	
	M	SD	M	SD
1×	48.00	17.47	56.88	28.33
1.2x	45.63	14.78	49.25	31.28

the proportion of resolution maneuvers issued within 30 s of conflict detection were 46 and 56 % for the same conditions (respectively) at the 1x traffic density, and 51 and 64 % at the 1.2x density. These results indicate that when removed from the conflict detection task, controllers more often needed more time in order to issue a resolution. Although measurements did not distinguish between solution identification time and solution execution time, when considering the fact that the solution execution methods available were constant across conditions, one can reasonably believe this data reflects an increase in the solution identification time (i.e., the controllers needed more time to determine how to solve the conflict).

Although the concentration of resolution response times showed more noticeable changes within the different comparisons, a repeated-measures analysis of variance (ANOVA) for resolution response time mean values did not provide significant results for the task set or traffic density variables ($p > 0.05$). Table 1 lists the relatively similar descriptive statistics for the four combinations of task sets and traffic densities.

3 Method

This paper explores whether or not other subjective measures display a relationship to the controllers' conflict resolution performance. The current analyses examine workload and situation awareness because prior research identified both as critical factors that frequently and negatively influence controller performance [6]. The results from [4] seem to support an obvious hypothesis: when controllers are not involved in the conflict detection process, they know less about the circumstances surrounding the conflict, and as a result, need more time to assemble a detailed enough picture in order to know what action(s) to take. This paper seeks to validate this idea using the available situation awareness data. Analyses will also compare the conflict resolution performance data against the controller workload data to seek out other hidden relationships between the objective and subjective data.

Although the full study included two treatments of run-length, the analyses in this paper are limited to only the 60-min duration runs. However, the results from [4] collapse the run-length variable, combining data from the 60-min runs with data from the 20-min runs. In order to better align with the analyses presented here, new analyses of the performance data are also included.

3.1 Workload

Workload Assessment Keypads (WAKs) probed controller workload at three-minute intervals during the simulation trials. Controllers responded to the workload probes with Air Traffic Workload Input Technique (ATWIT) [7] ratings along a modified six-point scale (e.g., 1 as low workload, 6 as high workload).

3.2 Situation Awareness

The study collected situation awareness data using the Situation Present Assessment Method (SPAM) [8]. After responding to each of the workload prompts, a small window appeared on the display, presenting participants with a situation awareness question. Developed in collaboration with three retired air traffic controllers who were not participants in the study, the questions used a yes/no response format, implemented as separate response buttons within the question window. After answering the situation awareness question (i.e., after clicking either the 'yes' button or the 'no' button), the window automatically disappeared, allowing the participants to return to their air traffic control duties with minimal interruption. Results included in this paper benefit from two different measures of situation awareness: percentage of questions answered correctly (accuracy), and elapsed time between question presentation and correct answer (response time).

4 Results

The following describes the results from the current data analyses, all sourced exclusively from the 60-minute runs within the *Conflict Resolution* (CR) and *Conflict Detection & Resolution* (CD&R) task sets. The selected metrics are first considered individually, followed by multi-variate examinations that look to identify quantifiable relationships (via a series of Spearman's Correlation tests) between the objective conflict resolution performance data, and the subjective workload and situation awareness data. Other publications provide additional results from the HACR simulation [4, 5].

4.1 Resolution Response Time

Across task sets, the controllers were able to issue resolution maneuvers within 30 s of conflict detection for 51 % of cases in the CR task set, and 53 % of cases in

Table 2 Summary of means and standard deviations of the resolution response times (seconds), workload ratings, situation awareness accuracy, and situation awareness response times (seconds) from the CR and CD&R task sets in both 1x and 1.2x densities

Resolution response time	CR		CD&R	
	M	SD	M	SD
1x	46.75	20.48	69.88	37.78
1.2x	43.13	9.06	66.38	56.36
Workload	CR		CD&R	
	M	SD	M	SD
1x	1.69	0.42	2.48	0.72
1.2x	2.11	0.60	3.12	0.55
Situation awareness accuracy	CR		CD&R	
	M (%)	SD (%)	M (%)	SD (%)
1x	74.50	10.20	77.75	11.34
1.2x	70.25	16.15	74.25	7.34
Situation awareness response Time	CR		CD&R	
	M	SD	M	SD
1x	6.89	1.28	6.23	1.61
1.2x	7.34	1.26	5.89	1.65

CD&R. After further isolating the traffic density variable, data from the trials simulating 1x traffic density indicated that 52 % of resolution maneuvers occurred within 30 s of detection in the CR condition, compared to 51 % in CD&R. These numbers changed to 49 and 59 % (respectively), at the 1.2x traffic density. This data differs from the findings reported in [4] that, at the highest level, associate controller involvement in the conflict detection task with more often needing less time to resolve a conflict. Such distinction is no longer present in this data, now characterized by largely similar distributions.

When looking at the mean values for resolution response time, ANOVA results approached significance for the comparison between task sets ($F(1,7) = 3.928$, $p = 0.088$), where CR (surprisingly) had faster resolution times ($M = 44.938$, $SD = 4.74$) than the CD&R condition ($M = 68.125$, $SD = 10.853$). Traffic density did not have a significant effect on resolution response time. Table 2 lists the relevant mean and standard deviation values.

4.2 Workload

A Kolmogorov-Smirnov test indicated that the workload data violated the assumptions of normality ($p < 0.05$), thus requiring a Friedman's ANOVA for non-parametric data. This test revealed a significant difference between task sets and traffic densities, $\chi^2(3) = 18.600$, $p = 0.01$. Post hoc analyses applied a Bonferroni correction to Wilcoxon signed-rank tests and showed significant differences between task sets, in both the 1x ($Z = -2.100$, $p = 0.036$) and the 1.2x ($Z = -2.521$, $p = 0.012$) densities, with CR reporting lower workload ratings than the CD&R condition. Traffic density had a significant effect on workload, but only in the CD&R task set ($Z = -2.521$, $p = 0.012$), with lower workload ratings coming from the 1x density. There was no significant effect of traffic density within the CR condition ($Z = -1.120$, $p = 0.263$). Descriptive statistics reflect these trends (see Table 2). These workload results appear to support our expectation (and also align with the HACD data reported in [3]), that workload would increase under less automated working environments and during higher levels of traffic.

4.3 Situation Awareness

Situation Awareness Accuracy. Listed in Table 2, the average percentages of correctly-answered situation awareness questions remained fairly stable throughout the four combinations of task sets and traffic densities, with between 70.25 and 77.75 % accuracy. A repeated-measures ANOVA confirmed this with no significant effect of condition or density ($F(1,7) = 0.747$, $p = 0.416$) and ($F(1,7) = 1.268$, $p = 0.297$), respectively. These results contest the expectation that removing controllers from the conflict detection task (i.e., the CR task set) would negatively impact their situation awareness.

Situation Awareness Response Time. A repeated-measures ANOVA showed a significant difference in situation awareness response time as a result of task set, ($F(1,7) = 7.555$, $p < 0.05$), where the CR condition had slower response times ($M = 7.114$, $SD = 0.391$) than the CD&R condition ($M = 6.057$, $SD = 0.525$). Tests also revealed that traffic density had no significant effect on situation awareness response time. In contrast to the situation awareness accuracy data, these results support the notion that the conflict detection task is an important contributor to the controller's understanding of the traffic in their sector. Admittedly, the statistical significance here represents a difference of less than two seconds; therefore such findings may have limited meaning. Descriptive statistics are listed in Table 2.

4.4 Resolution Response Time and Workload

Results from a Spearman's correlation test between resolution response time and workload approached significance in the CD&R-1.2x pairing ($r_s(18) = 0.452$, $p = 0.060$). This shows a positive relationship in that resolution response time and workload increased together. It is also interesting to note that while controller

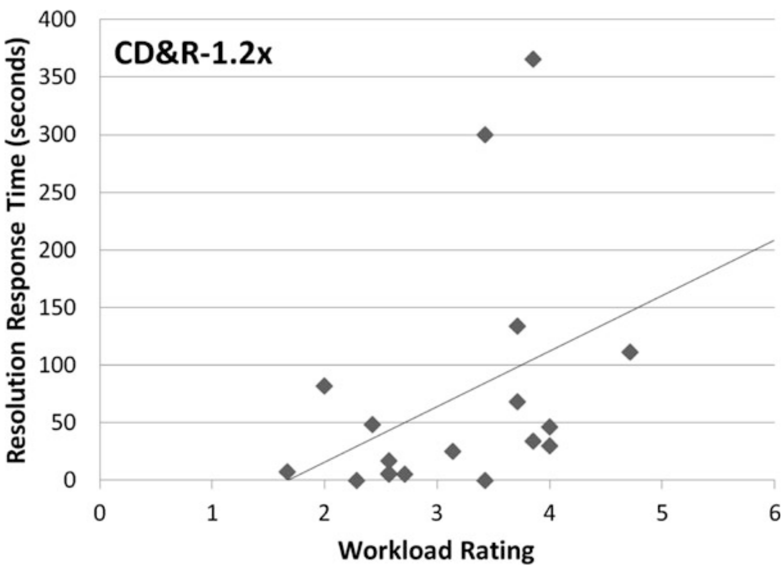


Fig. 2 Scatterplot of average resolution response time against average workload rating, including a linear trend-line, for the CD&R task set in the 1.2x traffic density

workload increased, the variability in resolution response time increased as well (see Fig. 2). There were no significant or near-significant relationships found for any of the other combinations of task set and traffic density.

4.5 Resolution Response Time and Situation Awareness

The correlation between resolution response time and situation awareness accuracy approached significance for two pairings: CR-1x and CD&R-1x (($r_s(24) = 0.374$, $p = 0.072$) and ($r_s(24) = -0.375$, $p = 0.071$), respectively). Although these two correlations are similar in strength, their directions are inverted relative to each other. The CR-1x correlation displayed a positive relationship, with resolution response time and situation awareness accuracy increasing together. Meanwhile, the CD&R-1x correlation revealed a negative relationship, where resolution response time decreased as situation awareness accuracy increased. For reference, these findings are reflected in Figs. 3 and 4. Further tests were unable to find any correlations of significance or near-significance for any of the combinations of task set and traffic density ($p > 0.1$).

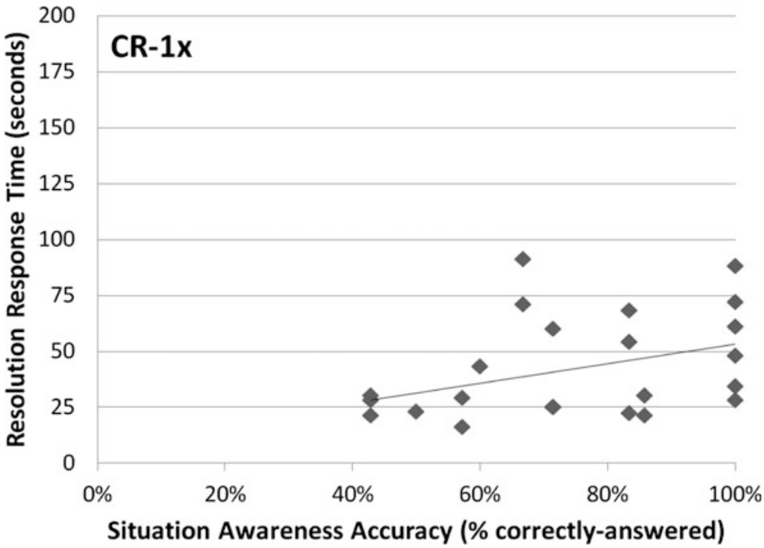


Fig. 3 Scatterplot of average resolution response time against average situation awareness accuracy, including a linear trend-line, for the CR task set in the 1x traffic density

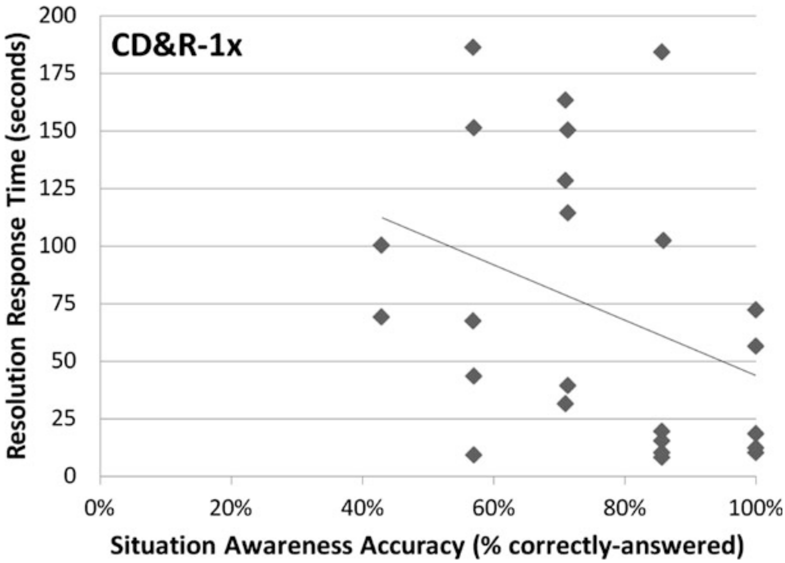


Fig. 4 Scatterplot of average resolution response time against average situation awareness accuracy, including a linear trend-line, for the CD&R task set in the 1x traffic density

5 Discussion

When the confounds of run length and traffic density were present in the comparison of resolution response time data between the CR and CD&R task sets, the distribution of data indicated value in having controllers involved in the conflict detection task. However, the new analyses, which focused just on the 60-min trials and split out the traffic density variable, no longer supported that argument. One possible explanation for this is that in the shorter, 20-min trials controllers were more engaged and experienced less fatigue; perhaps supported by the larger standard deviations. Within the longer, perhaps more tiring trials, the demands of the more manual environment associated with the CD&R task set, seem to have led to the increased mean response times; enough to reach significance. Another element possibly contributing to these differences stems from a simulation artifact: in the CR condition, the trigger for the radar display's blackout mode was simply the absence of any detected conflicts. The following sequence captures an unintended consequence of this implementation: (1) conflict detected, blackout mode disengages; (2) controller issues a heading vector to maintain separation; (3) conflict resolved (i.e., no longer detected), blackout mode engages. The end result of this example is that an aircraft continues along an open vector, with the controller unable to see when to issue the follow-up heading instruction to put the aircraft back on course. During the simulation, controllers received training on how to manually disengage the blackout mode, facilitating the ability to follow-up on an open-ended instruction in order to 'close the loop', at which point they could manually re-engage the blackout mode. Comments from a few participants indicated this process was a bit cumbersome, and could explain a higher proportion of simpler, cruder resolution maneuvers (e.g., altitude instructions) that better supported the ability to more quickly complete the resolution process for a given encounter before moving on to the next thing.

The workload data describes very believable circumstances, where controllers felt less busy during conditions where they had (literally) nothing to look at for part of the time. Additionally, their workload ratings helped validate our traffic scenarios, reporting lower workload in the 1x traffic density. That the same difference between traffic densities was only observed in CD&R is likely because the effects of the CR task set's blackout mode outweighed the impact of traffic density.

A major concern about human-automation interactions is the possible reduction of operator situation awareness. While the situation awareness accuracy metric did not see any effect of task set, the situation awareness response time data did. The results not only highlight the need to examine situation awareness from multiple angles, but suggest that the controllers were able to perform equally well in correctly answering the situation awareness questions across both task sets, but only at the expense of response time. Given that such expenses amounted to less than two seconds of time, the likely implication is that both the level at which situation awareness was degraded and the amount of compensation needed to overcome it, were minimal.

Correlating the resolution response time and workload data uncovered a potential trend describing a positive relationship in which the resolution response times and workload ratings increased together. This relationship only appeared during the CD&R-1.2x condition, which generally speaking, was the most challenging of the analyzed conditions, since it paired the higher traffic density with the more manual task set. Perhaps the more difficult nature of this condition explains why the statistical relationship did not appear anywhere else, suggesting that it was the only condition able to elicit a meaningful range of workload ratings from the participants. Also, the resolution response times appear to disperse more as the workload ratings increase, providing additional evidence of the relationship between the two measures: during the more complex situations likely associated with higher workload ratings, it would be reasonable for controllers to need more time to resolve a conflict.

The correlational analysis between resolution response time and situation awareness accuracy helped uncover a few key aspects of how they influenced each other. During the CR task set, any time spent by the controllers resolving a conflict directly corresponded to the amount of time that the blackout mode was disengaged, and consequently, the amount of time they were able to observe the traffic in their sector. Therefore, any increase in resolution response time brought with it more time for the controllers to observe traffic, and naturally led to better answers to the situation awareness questions. Whereas the resolution response time (and 'screen time', indirectly so) seemingly drives the situation awareness in the CR task set, that simple story may not hold true in the CD&R task set. Rather, it appears as if the situation awareness is driving the resolution response time. As controllers perform their conflict detection duties, they naturally need to observe more things and consider more things, and as a result, may need more time to resolve certain conflicts (note the larger spread in resolution response time data in Fig. 4 vs. Fig. 3). When we consider the situation awareness component, a controller with low situation awareness is likely to take even longer to resolve a conflict. Conversely, a controller with good situation awareness can more likely identify a resolution more quickly.

6 Conclusion

This paper examined the subjective measures of workload and situation awareness within the objective context of conflict resolution response time. Real-world service providers are considering future air traffic management systems that include more automation: automation that will likely work jointly with human operators. It is critical then, to understand the various impacts of human-automation interaction, in order to identify any costs or consequences that could inform good system design. In addition to showing that creating an environment which removes the controller from the detection task is more difficult than one might assume, the results here uncover not only the importance of analyses which co-examine multiple factors, but also offer evidence of the obvious relationship between conflict detection and conflict resolution. It was the situation awareness data that best identified how the detection and resolution tasks influence each other. Findings from the situation awareness accuracy data point to the idea that controllers can likely resolve conflicts with our without first detecting the conflict… but will do so in very different ways. Data here suggest removing the conflict detection task will limit situation awareness and may result in the consideration of only a few factors, producing resolutions of a more simplified nature; whereas detecting a conflict beforehand will add to the controller's situation awareness and may result in the considerations of several factors, producing resolutions of a more *optimized* nature.

References

1. Joint Planning and Development Office.: Concept of operations for the next generation air transportation system. Version 3.2, JPDO, Washington, DC (2010)
2. Prevôt, T., Smith, N., Palmer, E., Callantine, T., Lee, P., Mercer, J., et al.: An overview of current capabilities and research activities in the Airspace Operations Laboratory at NASA Ames Research Center. In: AIAA-2014-2860, American Institute of Aeronautics and Astronautics, Reston, VA (2014)
3. Edwards, T., Homola, J., Mercer, J., Claudatos, L.: Multi-factor interactions and the air traffic controller: the interaction of situation awareness and workload in association with automation. In: Proceedings of the 13th IFAC Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, Kyoto, Japan (2016)
4. Mercer, J., Gomez, A., Gabets, C.: Impact of automation support on the conflict resolution task in a human-in-the-loop air traffic control simulation. In: Proceedings of the 13th IFAC Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, Kyoto, Japan (2016)
5. Mercer, J., Homola, J., Edwards, T., Gomez, A., Bienert, N., Martin, L., et. al. (2016, in press). Assessing Operator Capabilities in a Human-Automation Team Environment during Simulated Air Traffic Control Work. Human Factors
6. Edwards, T., Sharples, S., Wilson, J. R., Kirwan, B.: The need for a multi-factorial approach to safe human performance in air traffic control. In: Proceedings of the 4th AHFE International Conference, San Francisco: USA (2012)
7. Stein, E.: Air traffic controller workload: an examination of workload probe. In: DOT/FAA/CT-TN84/24, Federal Aviation Administration, Washington, DC (1985)
8. Durso, F.T., Truitt, T.R., Hackworth, C., Crutchfield, J., Nikolic, D., Moertl, P., et al.: Expertise and chess: a pilot study comparing situation awareness methodologies. In: Garland, D.J., Endsley, M.R. (eds.) Experimental Analysis and Measurement of Situation Awareness, pp. 295–304. Embry-Riddle Aeronautical University Press, Daytona Beach (1995)

Human Systems Integration and Strategic Planning

Edward Austrian, Michael Sawyer and Katherine Berry

Abstract The National Airspace System (NAS) Enterprise Architecture (EA) describes Next Generation Air Transportation System (NextGen) goals, operational changes, planned infrastructure changes, and guidance materials that are referenced by FAA programs throughout the acquisition process. To strengthen the presence of human factors in NAS infrastructure plans and improvements, the FAA Human Factors Research and Engineering Division executed a redesign of the Human System Integration (HSI) Roadmap. This paper will present the methods utilized to redesign the HSI Roadmap, provide an overview of sample human factors integration opportunities, and key lessons learned.

Keywords Transportation · Aviation · NextGen · Human-systems integration

1 Introduction

The Federal Aviation Administration (FAA) is transforming the National Airspace System (NAS) through the implementation of the Next Generation Air Transportation (NextGen). NextGen aims to improve safety, capacity, and efficiency through the introduction of inter-dependent infrastructure, service, procedure, and policy updates [1]. Serving as a guide to this complex transition is the FAA's NAS

E. Austrian (✉) · M. Sawyer · K. Berry
Fort Hill Group, 660 Pennsylvania Ave, Suit 204, Washington, DC 20003, USA
e-mail: Eddie.Austrian@FortHillGroup.com

M. Sawyer
e-mail: Michael.Sawyer@FortHillGroup.com

K. Berry
e-mail: Katie.Berry@FortHillGroup.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_10

Enterprise Architecture (EA). The NAS EA provides users with a strategic blueprint of top-down service changes, infrastructure changes, and supporting guidance materials. These data elements are communicated through a variety of EA products and are referenced by FAA programs throughout the acquisition process [2].

For approximately eight years, the HSI Roadmap has been an integral part of the NAS EA. Since its inception, the HSI Roadmap has been the only NAS EA product to exclusively document human factors engineering activities and the evolution of user-centered needs. On an annual basis, the HSI Roadmap is updated to reflect the most current human factors needs and how FAA Human Factors intends to respond to them [3].

To strengthen the inclusion of human factors in NAS infrastructure plans and improvements, the FAA Human Factors Research and Engineering Division (ANG-C1) executed a redesign of the HSI Roadmap between October 2014 and June 2015. The intent of the redesign was to ensure compliance with EA guidelines while more clearly documenting the impact of human factors products on NAS infrastructure changes. As such, the renewed focus of the HSI Roadmap moved away from workforce evolution and shifted towards the documentation of human factors integration opportunities to support infrastructure changes and responsive human performance.

1.1 Purpose

One of the many challenges across industry is the timely inclusion of human factors in acquisitions. It is critical that the FAA proactively identify and address user-centered needs as it continues to concurrently develop and implement NextGen changes [4]. From an infrastructure development standpoint, the HSI Roadmap aims to address this need and promote the consistent consideration and inclusion of human factors in infrastructure development plans [5]. This paper presents the development of the 2016 HSI Roadmap, an overview of current human factors and safety needs, sample human factors integration opportunities, and lessons learned throughout the development process.

2 Methods

The HSI Roadmap redesign was executed in two phases—Information Display Structure and Content Development. Each of these phases and related methods are detailed in the following sub-sections.

2.1 Information Display Structure

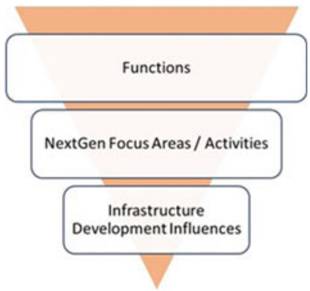
During Phase I of the redesign process, data was gathered from the FAA Human Factors Library [6]. NextGen research reports documenting projects managed by ANG-C1 between 2012 and 2014 were collected. Each report's abstract was evaluated to identify three components—(1) the NAS shortfall/mission gap that a project aimed to address, (2) the project's end product description and proposed benefits, and (3) the NAS infrastructure impact that a project's end product may yield.

As these information elements were identified, they were sorted into three categories—Shortfalls, Benefits, and Infrastructure Impacts. At the completion of this evaluation, each category was further examined to identify independent, trending themes. The Shortfalls category was examined to identify the leading NextGen human factors needs addressed by ANG-C1 between 2012 and 2014. The Benefits category was examined to identify the most frequently proposed benefits for the projects evaluated. The Infrastructure Impacts category was examined to identify the leading impacts to NAS infrastructure. The emerging themes from each of the three categories were applied at an enterprise-level to develop the HSI Roadmap's information display structure.

Utilizing the classified information, three layers of information (shown in Fig. 1) within the HSI Roadmap were developed. From the Shortfalls category the top layer, Functions, are a representation of the on-going FAA assessment to identify targeted, NAS-wide human factors needs. The middle layer, NextGen Focus Areas/Activities, was derived from the Benefits category. NextGen Focus Areas/Activities are a representation of the project-level responses to one or more human factors needs. From the Infrastructure Impacts category, the final layer, Infrastructure Development Influences, was determined. Infrastructure Development Influences are a representation of the direct impact that a specific human factors product may introduce to one or more NAS infrastructure program(s) at a specific point-in-time.

Figure 2 provides an overview of the HSI Roadmap Functions and NextGen Focus Areas/Activities. Specific Infrastructure Development Influences are depicted through product integration lines in the HSI Roadmap diagrams.

Fig. 1 HSI roadmap information layers/hierarchy



Functions	NextGen Focus Areas / Activities
Human-System Performance Risk Reduction *NextGen ATC / Tech Ops Error Management and Complex Systems*	Human Performance and Safety Human Centered Design
Workstation Development *ATC Information Integration and User Comprehension*	ATC Information Presentation & Management ATC Display Convergence
FAA Acquisition Evaluation *Human Factors AMS and SMS Support to Emerging Programs*	FAA Order 9550.8 & AMS Policy 4.7 Compliance
Procedure Development *NextGen Air-Ground Procedure Development*	Procedure Complexity Reduction and Usability
Support FAA Evaluation and Approval of Complex Systems *NextGen Flight Deck Error Management*	Evaluation of Applicant Compliance with Human Factors Related Regulations
Support FAA Evaluation and Approval of NextGen Avionics & Flight Deck Technologies *Flight Crew Interfaces, Installation & Integration Issues, & Ops*	Evaluation of Applicant Compliance with Human Factors Related Regulations

Fig. 2 HSI roadmap functions & NextGen focus areas/activities

2.2 *Content Development*

During Phase II of the redesign process, data was gathered from the NAS EA Portal [2]. Seventeen human factors decision points were collected. Each of the human factors decision points were evaluated based on their expected NAS infrastructure impacts. Next, each decision point was linked to the HSI Roadmap Functions and NextGen Focus Areas/Activities. To identify individual Infrastructure Development Influences, an additional 207 non-human factors decision points/regulatory milestones were collected. Each of the 207 non-human factors decision points were evaluated through one-on-one stakeholder working sessions to identify technical relationships with the 17 human factors decision points. Of the 207 non-human factors decision points, 12 decision points had technical relationships with one or more of the 17 human factors decision points. Each of these relationships were validated by stakeholders and are clearly documented throughout the HSI Roadmap diagrams through product integration lines.

3 Results

Upon completion of the annual NAS EA update cycle, the HSI Roadmap was published on the NAS Systems Engineering Portal in February 2016 [2, 3]. The updated roadmap depicts numerous direct and indirect relationships with other infrastructure roadmaps, such as Automation, Aircraft, Airspace & Procedures, Safety, Surveillance, and Communications. Documentation of these relationships increases the visibility and accountability of human factors. It also drives strategic human factors activities and promotes cross-domain coordination to support the timely influence of human factors products on NAS infrastructure changes.

3.1 Human Factors and Safety Results

Diagram 1 of 4 of the HSI Roadmap (documented in Fig. 3) depicts the Human-System Performance Risk Reduction Function and its related NextGen Focus Areas/Activities—Human Performance and Safety and Human Centered Design. The Human-System Performance Risk Reduction Function represents the on-going identification of enterprise-level human factors safety needs. Human factors safety needs may be identified through the examination of acquisition programs, internal FAA safety assessment outputs, or through the targeted analysis

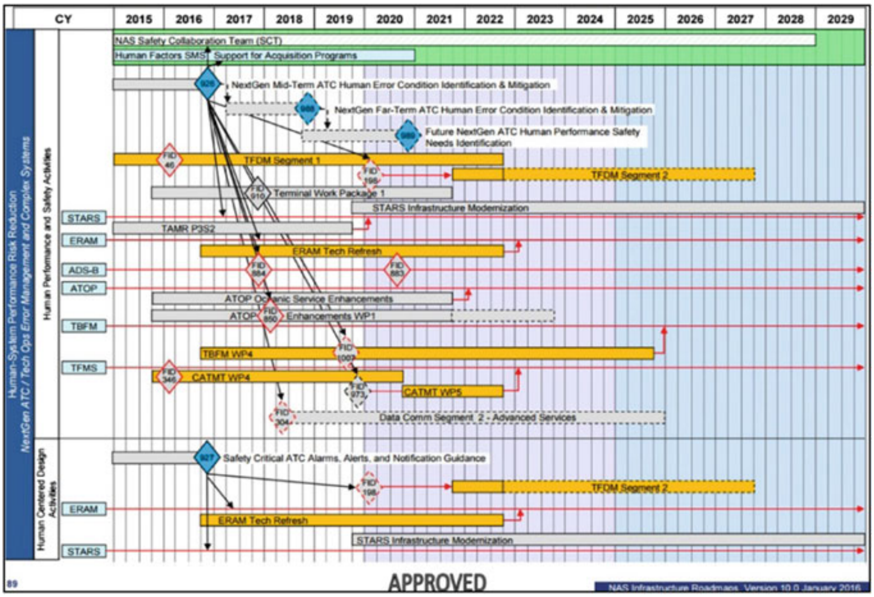


Fig. 3 HSI roadmap page 1 of 4

Table 1 Human factors and safety results overview

Function	NextGen focus areas/activities	Sample infrastructure development influences
Human-system performance risk reduction	Human performance and safety 926 988 989 994 995	• Allocation of functions between users and automation • Mitigation of human error conditions arising from the introduction of complex systems and procedures • Definition of operational human-system performance criteria • Proactive derivation of out-year human performance and safety needs
NextGen ATC/Tech Ops error management and complex systems	Human centered design 927	

of NextGen improvements. This Function drives the execution of its respective NextGen Focus Areas/Activities.

The Human Performance and Safety NextGen Focus Area/Activity aims to proactively identify NextGen human error modes and conditions that may arise from the introduction of complex systems and procedures. These activities also aim to identify air-ground acquisition programs that may benefit from completed or planned Human Performance and Safety products. Programs may apply these products to inform the development of NAS shortfalls, solution alternatives, or required FAA acquisition safety documents (e.g. OSA, SRMD, etc.). Currently there are 5 human factors decisions points that are linked to this NextGen Focus Area/Activity (Table 1).

The Human Centered Design NextGen Focus Area/Activity aims to address emerging air-ground safety-critical design needs. These activities also aim to identify acquisition programs that may benefit from completed or planned Human Centered Design products. Programs may apply these products to inform the development of Requirements Documents, Test & Evaluation (T&E) methods, and Verification & Validation (V&V) criteria. Currently there is 1 human factors decision point linked to this NextGen Focus Area/Activity.

For example, Table 2 summarizes a subset of human factors to non-human factors decision point relationships identified during Phase II of the HSI Roadmap redesign. Each of these relationships represent a specific opportunity (potential infrastructure development influence) for an acquisition program to apply research products related to human factors decision point 926.

In addition to completing the associated research with decision point 926, significant coordination between human factors and mid-term acquisition programs must be conducted to complete this decision. One of the many benefits to the HSI

Table 2 Human factors–non-human factors decision point relationships

Human factors decision point	Non-human decision points	Primary roadmap
926: Decision on the implementation strategy of mid-term human performance safety requirements into the NextGen safety process	198: FID for TFDM Segment 2	Automation
	1007: FID for TBFM Work Package (WP) 4	Automation
	973: FID for CATMT WP 5	Automation
	304: FID for Data Comm Segment 2	Communications
	850: FID for ATOP Enhancements WP 1	Automation
	884: FID for ADS-B In Applications	Surveillance

Roadmap is that it serves as a cross-program coordination tool for human factors. It reduces frequently encountered barriers between programs and supports ANG-C1's effort to transition products from research to reality.

4 Lessons Learned

The HSI Roadmap redesign process was executed over the course of 9 months. During this period, the team encountered multiple technical and logistical challenges. This section will review key lessons from this project.

4.1 Cross-Team Collaboration

During the HSI Roadmap redesign, the team regularly coordinated across multiple domains and lines of business. Engaging in this coordination from the beginning allowed the team seek technical inputs from a diverse pool of non-human factors subject matter experts and gain Agency-wide support during the product's concept development phase. Additionally, this cross-team collaboration served as a risk-reduction mechanism and significantly eased the implementation process.

4.2 Product Flexibility

In Phase I, the team initially employed a bottom-up approach to support information display structure development. In doing this, the team continually found that it was extremely easy to get lost in the details. As a result, the team adjusted their approach to Phase I and used a top-down approach (as detailed in the methods section). This allowed the end product to maintain its enterprise-level focus and have built in flexibility for future growth or document updates.

4.3 Lack of Industry Best Practices

Unlike many other lines of businesses within the FAA, human factors does not acquire infrastructure, it supports it. Due to this, documenting a cohesive, enterprise-level narrative was challenging. While the team found many FAA and Department of Defense Architecture Framework (DoDAF) references helpful, there were no comparable roadmaps, documents, or best practices from industry available [7]. As a result, the team referenced technology roadmaps from other infrastructure and non-infrastructure domains throughout Phase I and II of this effort.

5 Conclusion

For approximately 8 years, the HSI Roadmap has been an integral part of the NAS EA. Since its inception, the HSI Roadmap has been the only NAS EA product to exclusively document human factors engineering efforts and the evolution of user-centered needs. As such, the HSI Roadmap may be used as a tool to develop human factors dependencies and a method to drive the identification of future integration opportunities. Multiple human factors opportunities exist to support the successful delivery of NAS infrastructure and NextGen capabilities throughout the mid- and far-terms. Functionally, the HSI Roadmap may be used a means to support the identification of those opportunities and a tool to promote the timely inclusion of human factors in NAS acquisitions.

Acknowledgments We would like to acknowledge the FAA's Human Factors Research and Engineering Division (ANG-C1) for funding this project and similar work. Additionally, we would like to acknowledge the air traffic control and human factors subject matter experts who provided the valuable insight necessary to developing these results. The results presented herein represent the results of this research project and do not necessarily represent the views of the Federal Aviation Administration.

References

1. FAA.: Destination 2025. Retrieved from http://www.faa.gov/about/plans_reports/media/Destination2025.pdf (2012)
2. FAA.: NAS enterprise architecture. Retrieved from https://sep.faa.gov/ (2016)
3. FAA.: HSI roadmap. Retrieved from https://sep.faa.gov/file/get/IR_BRIEFING_PACKAGE (2016)

4. Zemrowski, K., Sawyer, M.: Impacts of increasing reliance on automation in air traffic control systems. J. Air Traffic **52**(4), 49–55 (2010)
5. FAA.: AMS policy. Retrieved from http://fast.faa.gov/docs/acquisitionManagementPolicy/AcquisitionManagementPolicy4.7.pdf (2015)
6. FAA.: FAA human factors library. Retrieved from http://www.hf.faa.gov/reports.aspx (2015)
7. DoD.: Department of Defense Architecture Framework. Retrieved from http://dodcio.defense.gov/Library/DoDArchitectureFramework.aspx (2016)
8. FAA.: Human factors policy document. Retrieved from https://www.faa.gov/regulations_policies/orders_notices/index.cfm/go/document.information/documentID/12081 (1993)

If It Gets Measured, It Can Be Managed

Sue Burdekin

Abstract The ICAO requirement for aviation operators to adopt a Safety Management System has led to innovative approaches in order to satisfy the measurement and monitoring of performance. One pilot human factors performance evaluation methodology that has gained support in both civil and military operations is Mission Operations Safety Audits (MOSA). MOSA has evolved from an experimental research program which was initially tested in a single pilot FA/18 Hornet simulator. It was then adapted to multi-crewed flight decks and further tested in a European civil airline and an Indian sub-continent regional airline before being reintroduced to military multi-crewed transport operations. Following the success of the trials, a mature MOSA Program was recently rolled out to all of the transport squadrons from 86Wing, Royal Australian Air Force. These included squadrons operating the C-17 Globemaster III, the KC30A MRTT (Multi-Roll Tanker) and the KA350 King Air. MOSA is a structured pilot self-assessment program that collects data across subject matter expert designed categories of behaviour that are tailored to meet the specific requirements of the operation. Each pilot assesses him/herself, their co-pilot and how they perceive the overall performance of their operation as a crew. The anonymity of the self-reporter is protected and the data is submitted by use of electronic tablet technology where it is analysed to produce a system evaluation report and highlight developing issues. Once base-line measures are established, the effectiveness of interventions can be can be measured by subsequent MOSA evaluations. This paper will discuss the development of the MOSA methodology and give examples of the results that it can provide to users.

Keywords Safety management systems · Pilot performance evaluation · Human factors

S. Burdekin (✉)
Australian Defence Force Academy, University of NSW, Northcott Drive,
Canberra, Australia
e-mail: s.burdekin@adfa.edu.au

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_11

1 Introduction

In November 2013, Annex 19 to the Convention on International Civil Aviation came into force becoming the first new Annex to be adopted by the Council of the International Civil Aviation Organisation (ICAO) in over 30 years. Annex 19 consolidated references to Safety Management previously contained in Annexes 1 (Personnel Licensing), 6 (Operation of Aircraft), 8 (Airworthiness of Aircraft), 11 (Air Traffic Services), 13 (Aircraft Accident and Incident Investigation) and 14 (Aerodromes). Annex 19 requires States to develop a State Safety Plan (SSP). In Australia, the Civil Aviation Safety Authority (CASA) developed the SSP in cooperation with the Australian Transport Safety Bureau (ATSB), Air Services Australia, the Department of Transport, the Department of Infrastructure and Regional Development, the Australian Maritime Safety Authority, the Bureau of Meteorology, and the Department of Defence. Where possible, the Australian Defence Force (ADF) aims to harmonize its approach to safety regulations with the civil system. This philosophy provides a relatively seamless interface with joint operations in controlled airspace and multi-user airports. Additionally, as more military practices, such as, aircraft servicing and some transport aircraft leasing arrangements, are outsourced to civilian organisations, common safety regulations ensure high standards of compliance.

There are four components that should be addressed within the State Safety program framework: Safety Policy and Objectives; Safety Risk Management; Safety Assurance; and, Safety Promotion. Under Safety Assurance States are required to maintain safety oversight; collect, analyse and exchange data; and, target the oversight of areas of greater concern or need [1].

The SSP provides governance for the CASA SMS framework, which requires operators to undertake 'Performance Monitoring and Measuring' within the Safety Assurance component (refer to Fig. 1). The regulator does not prescribe how operators are to meet these standards but rather, is open to new and innovative approaches as long as compliance is achieved [2].

One pilot performance human factors evaluation methodology that has gained support in both civil and military operational environments is Mission Operations Safety Audits (MOSA). MOSA was originally designed as a military aviation human factors evaluation tool. The ADF Directorate of Aviation and Air Force Safety (DDAAFS) sought to establish whether there had been a transfer of crew resource management (CRM) training from the classroom to the flight deck. A review of Line Operations Safety Audits (LOSA) found that the behavioural observation information gathered could be utilized within the DDAAFS safety measurement and monitoring program. However, LOSA was originally designed for civil commercial airline application where passenger comfort and on-time performance demonstrate the success of the mission. Military aviation operations are far more diverse, often involving high speed, rapid manoeuvre and terrain following flight. Additionally, many military aviation platforms cannot accommodate an observer and for those transport aircraft that could, the review concluded that LOSA could be expensive,

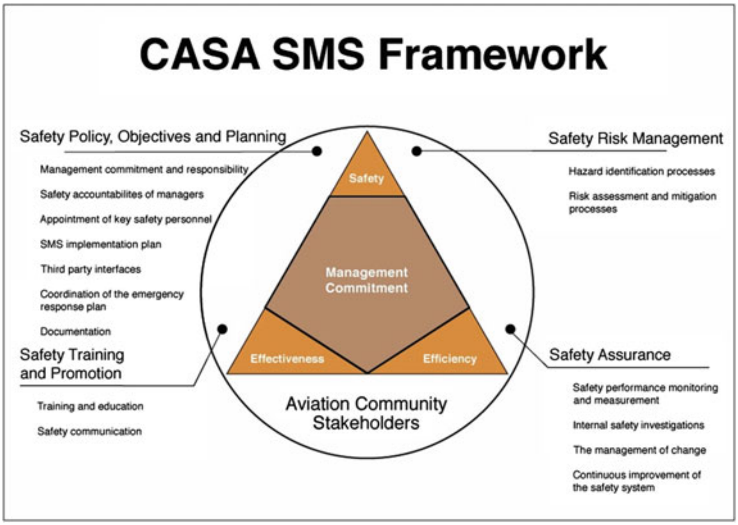


Fig. 1 Australian CASA safety management system framework

labour intensive and intrusive. Experimental research was therefore undertaken to test the validity of structured pilot behavioural self-assessment across a range of pre-determined categories of non-technical skills behaviour.

The original MOSA categories of behaviour were developed by a focus group of subject matter experts (SMEs) comprising military line pilots, instructors and executive (operational management) officers, assisted by the researcher, who is also a commercial pilot. Examples of the categories of behaviour are listed later in this paper. An experimental paradigm was designed utilizing a F/A-18 single seat simulator to test the hypothesis that highly skilled and operationally current professional pilots were able to confidently and accurately self-assess their own performance using the structured categories of behaviour protocols. Thirty RAAF fast jet pilot volunteers flew the same high and medium workload simulator mission whilst two trained observers assessed their performance using the same protocols that the pilots used to self-assess. Prior to conducting the main analysis, an interrater reliability test established that the observers agreed with each other on the assessment of the participant pilots. Those assessments were then compared to the pilot self-assessments and the correlated results supported the hypothesis [3].

Following the success of this initial experiment, the MOSA methodology was then further developed and tested in two multi-crewed civilian airlines. The first was based in Europe and operated A319 aircraft, and the second was a regional airline based in the Indian subcontinent and operated Dash8 aircraft. As the MOSA research was specifically interested in validating the ability of pilots to self-assess, the behavioural categories were tailored, once again by SMEs in each company and the researcher, to target non-technical skills that were relevant in their unique environment. Confidential self-assessment data were collected by captains, first officers and observers, assessing each pilot from the jump seat, during a total of over 100 normal revenue raising flight sectors. Additionally, each pilot was asked to confidentially rate the performance of their co-pilot. Once again, the observer and pilots' correlated results supported the premise that professional pilots are able to self-assess their own performance and that of their co-pilot across a range of predetermined behavioural markers. Furthermore, rather than inflate their ratings, as some literature espoused, pilots were generally more critical of themselves compared to the assessment of the trained observer [4, 5].

So, having successfully tested the MOSA pilot self-assessment methodology in both the simulator and in the field; in military single seat fast jet operations and in multi-crewed civilian commercial aircraft, it was concluded that highly skilled, well trained, professional pilots were able to reliably self-assess their own performance in a confidential, non-jeopardy situation for the purpose of an operational system evaluation. The strength of the MOSA approach, compared to LOSA and NOTECHS is that the protocols can be tailored to the operational needs of the user. MOSA is not a one size fits all process.

Upon completion of the MOSA research program, the Royal Australian Air Force (RAAF) 36 Squadron, which operates the C-17A Globemaster III aircraft, was the first ADF multi-crewed transport squadron to establish baseline human performance measurements by adopting the mature MOSA methodology.

2 MOSA Methodology

In preparation for the MOSA system evaluation, a review of Aviation Safety Occurrence Reports (ASORs) over the past 3 years was completed. Based on the findings from the review, a 36 Squadron Safety Questionnaire was designed by a team of Squadron SMEs, guided by the researcher (refer to Table 1). The results from the questionnaire further enabled the SMEs to highlight issues of operational concern and, therefore, better target the design of the categories of behaviour and the MOSA protocols.

One example of a safety issue identified in both the review of the ASORs and also in the questionnaire was 'fatigue'. Therefore, a decision was taken to incorporate a fatigue scale into the MOSA protocol design. Details including, nomination of the 'duty day', 'commencement of duty', 'end of duty', and a self-reported 'level of fatigue' at the commencement and end of each flight sector (leg) was requested of pilots using the Samn-Perelli 7-pt Sleepiness Scale. This scale is used widely in the aviation sector and was originally developed to estimate aircrew fatigue in USAF airlift operations [6–8].

Following focus group discussions, the MOSA categories of behaviour targeted by the SMEs were resolved to be: briefing; contingency management; monitor/cross check; workload management; situational awareness; automation management; communication; and, problem solving/decision making. An example of a category of behaviour can be seen in Table 2.

Table 1 Pre-MOSA questionnaire

1. In your opinion what are the top 3 aviation safety issues currently affecting ADF aviation?
2. Although we strive to prevent accidents, it is vital that we identify risks that are present in the aviation safety system. With that in mind and with your expert knowledge of the ADF aviation environment can you nominate any circumstances that might contribute to a future aviation incident/accident with the ADF?
3. What do you think would be the best way to prevent this?
4. What process or strategies do you use for making immediate risk assessments as part of your daily tasks?
5. On the scale below please indicate how comfortable you are in raising safety concerns with your chain of command. (1 = uncomfortable for fear of reprisal or not being supported 5 = very confident you will be supported)
6. What is the one thing you remember from your last CRM/HF training, and how have you applied it in your current role?

Table 2 Category of behaviour—automation management

Behavioural category	Descriptor	Grading/word picture (1. Poor; 2. Marginal; 3. Adequate; 4. Very good; 5. Excellent)
Automation management	Interaction between the operator and automated system Automation was properly managed to balance situational and/or workload requirements. *Automation setup was briefed to other members. Effective recovery techniques from automation anomalies*	1. Incorrect crew interaction and management of aircraft automatic systems. Clear errors of competency in automation set-up, mode selection and utilization 2. Basic interaction with aircraft automatic systems. Appropriate mode selection and utilization barely adequate to maintain safe flight profiles 3. Level of automation interaction adequate to maintain prescribed SOP profiles. Mode utilization satisfactory and procedurally correct. Recovery technique from anomalies reflects limited system awareness 4. Automation interaction to a good standard. Effective and timely management of automatic modes. Flight path SOP profiles maintained to a proficient standard. Clear understanding of aircraft automation systems reflected in sound anomaly management 5. Automation management to a high standard. Clear anticipation and use of appropriate modes. All anomalies managed to a highly proficient standard reflecting a deep understanding of the automation system

Table 3 Pilot participant details

Aircrew	Mean age	Mean years since training	Mean total hours	Mean hours on C17	Mean hours in past 30 days
Captains	31	8	2424	1255	25
Co-pilots	29	5.45	1334	196	25

Demographic information was collected from the pilot participants. This consisted of their: aircrew category; age; years since completion of flight training; total hours flight time; hours on the aircraft type (C-17); and recency—flight time in the past 30 days (see Table 3). Each pilot was then assigned a discrete number so that data matching could be achieved.

The design of the MOSA protocol included other variables, such as, 'type of mission' (airborne ops or logistics support), and 'pilot flying' or 'pilot monitoring'.

2.1 MOSA Procedure

As the MOSA program was now a mature system evaluation tool, observers were no longer used. Pilots were asked to confidentially self-assess their own performance, how they perceived their co-pilot performed, and how they perceived they performed as a crew in each of the eight categories of behaviour by using a 5 point Likert scale (as seen in Table 2). This was to be recorded for each sector (leg).

To streamline the data collection, MOSA protocols were adapted for iPad use in the aircraft. The aim was to minimise the impact on each mission and facilitate the efficient transfer of data to a secure server whilst flight crews were deployed.

3 Examples from MOSA Results

The data were collected from 36 Squadron captains and co-pilots during normal operational missions over a four month period. Whilst the evaluation was being conducted, the aircraft was tasked to complete a variety of missions around the globe, including: the delivery of humanitarian aid to countries that had experienced hurricanes, earthquakes and tidal waves; supply missions to the Middle East; and, the repatriation of victims from the MH-17 aircraft disaster in the Ukraine. The final analysis was comprised of data sets from 84 sectors.

The results from the initial MOSA pilot evaluation were able to provide the squadron with a baseline level of performance measures. These included: aggregated mean data of total pilot performance; how captains performed compared to co-pilots, how they perceived each other's performance, and how they performed as a crew. These results could also be broken down by category of behaviour, mission type, experience levels, and issues to do with fatigue. An example of the type of narrative that could be reported is listed below. These behavioural category assessments represent rating 4 on the 5-point Likert scale, and would be particularly relevant to the squadron executives because they themselves were part of the SME group that set the standard and developed them.

- *Briefing*: Effective crew briefing conducted utilizing all squadron/non-squadron information. There was proficient time and workload management with clear interaction and allocation of duties amongst crew.
- *Contingency Management*: Well-established threat management strategies employed. All threats detected early and strategies actively verbalized. All available resources were utilized.
- *Monitor/cross-check*: Consistent and effective monitoring of aircraft systems and crew actions were such that any anomalies were detected and responded to in a timely manner.
- *Workload Management*: All tasks were correctly organized in the manner that makes flight management efficient. Abnormal and emergency situations were quickly resolved to a good outcome.
- *Situational Awareness*: Situationally aware of all significant factors affecting the flight, regularly updated by checking against instruments, ATC, and other crewmembers.
- *Automation Management*: Automation interaction was to a good standard. Effective and timely management of automatic modes. Flight path SOP profiles maintained to a proficient standard. Clear understanding of aircraft automation systems reflected in sound anomaly management.
- *Communication*: Clear and concise communication with other crew members. Use of resources in a manner that improves overall safety.
- *Problem Solving and Decision Making*: Decision making takes into account all essential factors, follows correct procedure, and allows for contingencies.

The collapsed data were also able to show that both captains and co-pilots believed that they performed better as a crew than they individually rated themselves across all categories of behaviour. Therefore, when either the captain (self-report) and/or the co-pilot (self-report) data were marked down, overall the performance as a crew was perceived by each pilot to be of a higher standard than the performance of each pilot individually. This is more pronounced in the co-pilot cohort. Throughout this process both captains and co-pilots were more critical of their own performance than their crew members rated that performance, which was common in the earlier research where the pilots were being rated by an observer and therefore provides evidence that, collectively, these pilots reported accurately.

4 Discussion

A detailed MOSA Report was issued to the squadron executive. The report highlighted both areas of strength and weakness that had been identified by those professionals who perform at the coalface of the operation on a daily basis—the pilots.

The first mature ADF MOSA evaluation achieved a number of goals for 36 Squadron. Firstly, it was used as a safety assurance tool which, for the most part, confirmed that the squadron pilots were applying sound non-technical skills. Additionally, the report articulated the same issues that had been either suspected or had been highlighted from other empirical data collection. Therefore, management had the evidence to substantiate changes to specific areas of training and crew development. Finally, the initial MOSA provided a baseline for annual evaluation comparison. The philosophy of the squadron executive was clearly articulated by the final comment, 'if it gets measured, it can be managed'.

5 Conclusion

The strength, accuracy and practicality of the MOSA methodology were demonstrated by the successful outcome of the 36 Squadron MOSA. The dissemination of the MOSA report attracted the attention of other force element groups within the RAAF. This resulted in 86 Wing further extending the MOSA methodology to the remaining two transport squadrons it controls within Air Mobility Group, that is, 33 Squadron which, operates the Airbus Military KC-30A Multi-Role Tanker providing air-to-air refuelling and strategic transport capability; and 38 Squadron currently providing conversion training on the King Air, light transport capability and Royal Australian Army 3rd Brigade support. Following the steps outlined in this paper each squadron nominated their own SMEs and set about customizing the MOSA protocols to their own unique operational requirements. Some changes including, the addition of a text box for comments and a new category of behaviour, 'Planning', accompanied by the relevant word pictures, were incorporated into the design in expectation of achieving even greater granularity.

These two squadrons are currently in the process of collecting MOSA data. Additionally, 36 Squadron has undertaken its second MOSA evaluation. Analysis of that data will be able to provide a clear indication of the impact that implemented change has made.

The MOSA approach to pilot performance evaluation, as a component of the SMS safety performance and measurement requirement, has made a positive contribution to aviation safety within the RAAF. This is evidenced by the decision of 86 Wing to commit to conducting an annual MOSA for each of its transport squadrons over the next three years.

References

1. ICAO/Annex 19 to the Convention on International Civil Aviation.: Safety Management, 1 edn. International Civil Aviation Organisation, Montreal (2013)
2. Commonwealth of Australia: Australia's State Aviation Safety Program, April 2012. Department of Infrastructure and Transport, Canberra (2012)
3. Burdekin, S.G.: Mission Operations Safety Audits (MOSA): Measurements of Behavioural Performance or Non-technical Skills from Military Aircrew. Directorate of Flying Safety—Spotlight 4/03, Canberra (2003)
4. Burdekin, S.: Evaluating professional aircrew safety related behavior in-flight. In: Proceedings from the 4th International Conference on Applied Human Factors and Ergonomics, San Francisco, 21–25 July (2012a)
5. Burdekin, S.: Non-technical skills assessment to maximize aircrew performance and contribute to the safety management feedback loop. In: Proceedings from the Australian Aviation Psychology Association Conference, Sydney, 19–23 Nov (2012b)
6. Samn, S., Perelli, L.: Estimating Aircrew Fatigue: A Technique with Application to Airlift Operations. School of Aerospace Medicine, Brooks AFB, Texas (1982)
7. Petrilli, R., Roach, G., Dawson, D., Lamond, N.: The sleep, subjective fatigue, and sustained attention of commercial airline pilots during an international pattern. Chronobiol. Int. **23**(6), 1347–1362 (2006)
8. Dorrian, J., Baulk, S., Dawson, D.: Work hours, workload, sleep and fatigue in Australian Rail Industry employees. Appl Ergonomics **42**(2), 202–209 (2011)

New Approach to Determination of Main Solution Taking Dominant of Air Traffic Controller During Flight Level Norms Violation

Oleksii Reva, Sergii Borsuk, Bala Mirzayev Mushgyul-Ogli and Peyman Mukhtarov Shirin-Ogli

Abstract Mutual influence of ICAO flight safety main concept components is grounded from the perspective of main factor—“aviation personnel attitude to safe actions and conditions” taking into account influence of human factor on flight safety. This attitude is found with help of building and analysis of estimate usefulness functions for continuums of aircraft flight norms based on air traffic controllers solutions of closed decision taking tasks. Herewith main solution taking dominant that defines air traffic controller attitude to flight level norms violation (tending, indifferent, non-tending to risk) is commonly found with help of “risk premium” criterion that involves only one point of estimate usefulness function. Improved criterion that includes in calculation all characteristic points of usefulness function is proposed. It was found that under these circumstances efficiency of main solution taking dominant determination is increased in 20 %.

Keywords Air traffic controllers attitude to flight norms violations · Closed decision taking task · Estimate usefulness function characteristic points · Risk premium · Efficiency of main solution taking dominant determination

O. Reva · S. Borsuk (✉)
National Aviation University, Kosmonavta Komarova Av. 1, Kiev 03058, Ukraine
e-mail: grey1s@yandex.ru

O. Reva
e-mail: ran54@meta.ua

B.M. Mushgyul-Ogli · P.M. Shirin-Ogli
Main Center of United Air Traffic Control System of State Owned Enterprise AZANS, Azadlig Av. 1, 1005 Baku, Azerbaijan
e-mail: BalaMirzayev@azans.az

P.M. Shirin-Ogli
e-mail: Peyman.Mukhtarov@gmail.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_12

1 Introduction

It is generally known that front line air operators, that directly influence flight safety (in positive and negative way), professional activity may be considered as continuous chain of decisions processed and implemented in apparent and hidden forms under influence of various factors (objective/subjective, external/internal), especially different kinds of stochastic and deterministic risks [1, 2]. Since human factor is fairly considered by aviation institutions, specialists and researchers as main part among measures used to guarantee required level of flight safety [2–7 and others] correspondent researches should take into account peculiarities of technologies and procedures of aviation operator decision taking.

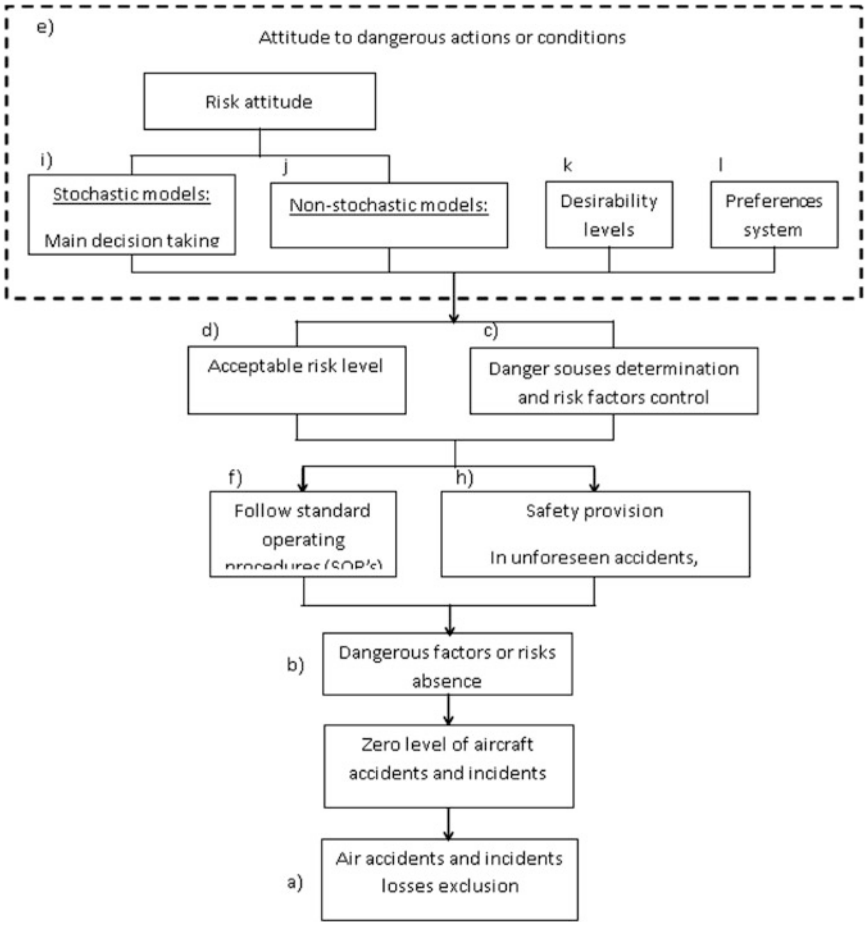


Fig. 1 Determination of ICAO safety concept components mutual influence from the human factor point of view

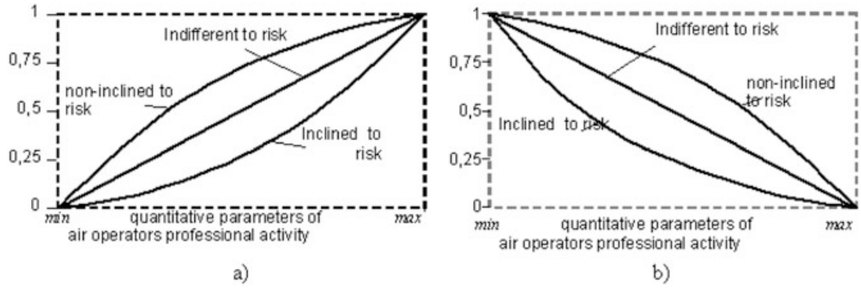


Fig. 2 General form of air operators professional activity estimate usefulness functions characteristics: **a** for functions with ascending argument; **b** for functions with descending argument

Basing on stated above it is possible to present interaction of safety concept components from position of their influence on human factor, namely "aviation personnel attitude to risky actions or conditions" part, as it is presented on Fig. 1.

From this figure one can see that mentioned "attitude" is grounded on several components among which in context of this paper attention should be paid to block (i), that shows basic decision taking dominants (inclined to risk, not inclined to risk, indifferent to risk) (Fig. 2). Besides one should state that these dominants determination is deeply proactive [2, 8–12] that absolutely corresponds to ICAO recommendations about implementation of proactive human factor strategies and alike.

Main decision taking dominants determine motivation for gaining success (risk inclination) or avoiding failures (risk non-inclination) and may be found from estimate usefulness functions parameters along with air operators professional activity characteristics that are built by limited number of key points received from special lotteries with open decision taking tasks. One should take into account that correspondent procedures were used mostly in economical researches and they were brought and developed in air operators scientific area by professor O.M. Reva along with his scientific school representatives.

2 Previous Researches

Peculiarity of proposed approach is construction of estimate usefulness functions upon physically tangible and well known indicators of air operators professional activity. Current paper examines usefulness of aircraft flight norms continuums taken from air traffic controllers opinions.

Five key points with usefulness 0, 0.25, 0.5, 0.75, 1 $(S_0, S_{0.25}, S_{0.5}, S_{0.75}, S_1)$, are used to build estimate usefulness function for mentioned continuum S.

As it was mentioned, special lotteries were used for that purpose (Fig. 3). To solve them respondents must state three determined lotteries equivalents with usefulness $S_{0.25}, S_{0.5}, S_{0.75}$.

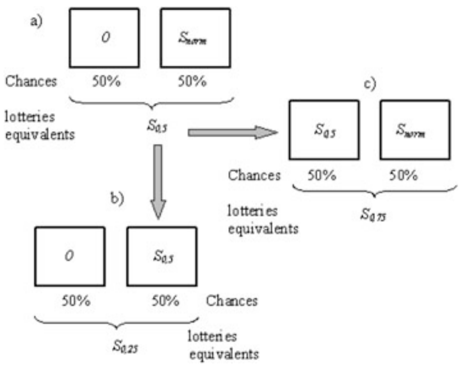


Fig. 3 Lotteries example

By determined lottery equivalent basing on classic sources [13, 14] in context of current research we consider such distance between aircrafts in limits of certain flight level norm that makes air traffic controller indifferent in choice between its value or 50 %–50 % lottery between maximal and proposed minimal values (absolutely acceptable/satisfactory value and absolutely unacceptable/unsatisfactory value).

Research involved 70 professional air traffic controllers of main center of united air traffic control system of state owned enterprise AZANS (Republic of Azerbaijan) and 132 air traffic control students from National aviation university and Kirovohrad flight academy (Ukraine). Researches were carried out according to single method for flight norms set by ICAO for horizontal plane, namely:

- distance between aircrafts $S = 20$ km, during longitudinal separation under IFR (Instrument Flight Rules) procedure with continuous radar monitoring on airways at the same level in ACC (Area Control Center) CTA (Control Area) and APP (Approach Control) TMA (Terminal Control Area);
- distance between aircrafts $S = 10$ km, during longitudinal separation under IFR procedure with continuous radar monitoring when crossing the same direction level occupied by another aircraft in approach area APP (TMA) using ATC automated system at the moment of crossing on conditions that no tracks converging.

Current paper shows results of professional ATC polling given in Table 1.

To determine ATC attitude to risk, i.e. main solution taking dominant, special risk premium is used. It is determined from ratio between taken flight norm median and determined lottery equivalent with usefulness $0.5(S_{0.5})$ and shows main solution taking dominant along all norm interval.

$$RR_{S_{0.5}} = \bar{S} - S_{0.5} = \begin{cases} > 0 & \text{---}\textit{noninclined to risk} \\ < 0 & \text{---}\textit{inclined to risk} \\ = 0 & \text{---}\textit{indifferent to risk} \end{cases}, \tag{1}$$

Table 1 Results of professional ATC building of personal estimate usefulness functions for flight norm S = 20 km **c**ontinuum (fragment)

ATC_i	Estimate usefulness function key points					MDTD	Proposed new method	
	S_0	$S_{0.25}$	$S_{0.5}$	$S_{0.75}$	S_1		$\sum S_i$	MDTD
1	2	3	4	5	6	7	8	9
ATC_1	0	10	14	17	20	T	61	T
ATC_2	0	8	12	16	20	T	56	T
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
ATC_9	0	8	10	16	20	I	54	T
ATC_{10}	0	12	16	17	20	T	65	T
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
ATC_{14}	0	10	15	17	20	T	62	T
ATC_{15}	0	5	10	15	20	I	50	I
ATC_{16}	0	14	15	18	20	T	67	T
ATC_{17}	0	15	17	19	20	T	71	T
ATC_{18}	0	5	10	15	20	I	50	I
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
ATC_{25}	0	13	15	17	20	T	65	T
ATC_{26}	0	5	10	15	20	I	50	I
ATC_{27}	0	5	10	15	20	I	50	I
ATC_{28}	0	7	12	16	20	T	55	T
ATC_{29}	0	5	7	10	20	N	42	N
ATC_{30}	0	5	10	15	20	I	50	I
ATC_{31}	0	5	10	15	20	I	50	I
ATC_{32}	0	5	15	17	20	T	57	T
ATC_{33}	0	10	15	17	20	T	62	T
ATC_{34}	0	5	10	13	20	I	48	N
ATC_{35}	0	5	10	15	20	I	50	I
ATC_{36}	0	5	10	15	20	I	50	I
ATC_{37}	0	3	6	10	20	N	39	N
ATC_{38}	0	5	10	12	20	I	47	N
ATC_{39}	0	10	15	17	20	T	62	T
ATC_{40}	0	7	12	15	20	T	54	T
ATC_{41}	0	5	10	15	20	I	50	I
ATC_{42}	0	5	10	15	20	I	50	I
ATC_{43}	0	5	10	15	20	I	50	I
ATC_{44}	0	5	10	13	20	I	48	N
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
ATC_{55}	0	16	18	19	20	T	73	T
ATC_{56}	0	5	10	13	20	I	48	N
ATC_{57}	0	10	15	17	20	T	62	T
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
ATC_{61}	0	5	10	15	20	I	50	I

(continued)

Table 1 (continued)

ATC_i	Estimate usefulness function key points					MDTD	Proposed new method	
	S_0	$S_{0.25}$	$S_{0.5}$	$S_{0.75}$	S_1		$\sum S_i$	MDTD
1	2	3	4	5	6	7	8	9
ATC_{62}	0	6	9	15	20	N	50	I
ATC_{63}	0	5	10	13	20	I	48	N
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
ATC_{68}	0	7	13	15	20	T	55	T
ATC_{69}	0	5	9	14	20	N	48	N
ATC_{70}	0	9	11	14	20	T	54	T

where $\bar{S}$—average lottery win, given at Fig. 3a:

$$\bar{S} = 0.5 \cdot S_0 + 0.5 \cdot S_1 = 0.5 \cdot (S_0 + S_1)$$

Desire for playing the lottery to receive best possible distance value between aircrafts characterize tendency to risk i.e. motivation to reach the success. At the same time when respondent shows lack of such desire he wants to avoid risk i.e. he is not inclined to it. Risk-indifferent respondents are considered to be "objective" since they has linear estimate usefulness function (Fig. 2).

Using formula (1) and data from Table 1 we may find main solution taking dominant for all air traffic controllers respondents (column 7 in Table 1) and receive following proportion of persons that tends (T), not tends (N) and are indifferent (I) to risk:

$$N{:}I{:}T \Leftrightarrow 5{:}30{:}35 \Leftrightarrow 1{:}6{:}7 \Leftrightarrow 7.1\,\%{:}42.9\,\%{:}50\,\%. \tag{2}$$

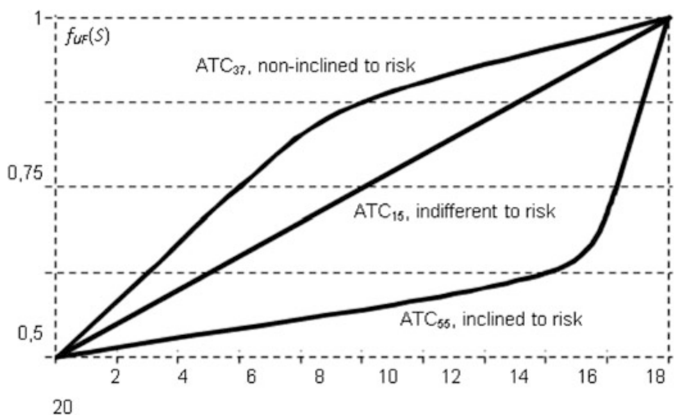


Fig. 4 Typical empiric individual estimate usefulness functions for flight norm continuum in 20 km

Though further use of received results is limited by cases when general tendency in estimation of researched flight norm continuum usefulness can't be determined by single point of lottery equivalent with usefulness $0.5(S_{0.5})$. Really, construction, approximation and analysis of individual estimate usefulness functions for ATC with any attitude to risk makes no problem (Fig. 4). At the same time approximation of expert information received from ATC may witness about different attitude.

3 Problem Statement

From all stated above the goal of current research is development of enhanced criterion for main solution taking dominant determination.

4 Results of Research

Proposed new method of ATC main solution taking dominant determination upon flight norm continuum is oriented on deeper ATC estimate usefulness functions analysis, namely with help of summary index of key points $S_0, S_{0.25}, S_{0.5}, S_{0.75}, S_1$ projection to the X axis research (Fig. 5).

Then as a result of linear estimate usefulness function nature summary index of projections will be equal to:

$$\begin{aligned} L &= S_0 + S_{0.25} + S_{0.5} + S_{0.75} + S_1 = S_{norm} \\ &= 0 + \frac{1}{4}S_{norm} + \frac{2}{4}S_{norm} + \frac{3}{4}S_{norm} + S_{norm} = 2.5S_{norm}. \end{aligned} \tag{3}$$

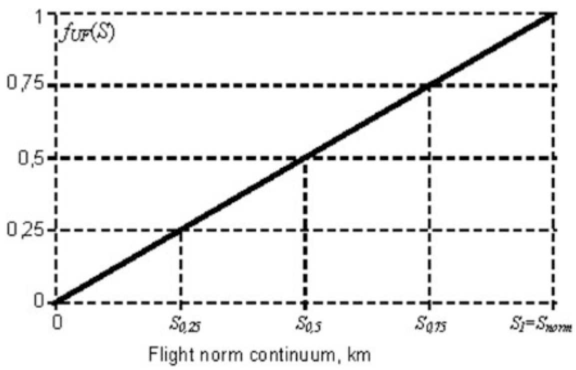


Fig. 5 Risk indifferent ATC estimate usefulness function projection analysis

Hence for ATC indifferent to risk individual summary index of key point projection into X axis should be equal to:

$$L_{ind.} > L \Leftrightarrow L_{ind.} > 2.5S_{norm}. \quad (4)$$

Then it is obvious (Figs. 2a, 4) that for ATC that is inclined to risk given index will be greater:

$$L_{inc.} > L \Leftrightarrow L_{inc.} > 2.5S_{norm}. \quad (5)$$

At last for ATC that is not inclined to risk that index will be equal to:

$$L_{non-inc.} < L \Leftrightarrow L_{non-inc.} < 2.5S_{НЕПС}. \quad (6)$$

In general that gives following risk premium value RR_L:

$$RR_L = L - L_{BDTD} = 2.5S_{norm} - L_{BDTD} = \begin{cases} > 0 & \text{---}non\text{−inclined to risk} \\ < 0 & \text{---}inclined\ to\ risk \\ = 0 & \text{---}indimmerent\ to\ risk \end{cases} \quad (7)$$

Applying criterion 7 to expert information given in Table 1 brought up rectified characteristics about ATC attitude to risk (columns 8 and 9 from Table 1) that may be summarized in following proportion

$$N{:}I{:}T \Leftrightarrow 9{:}25{:}36 \Leftrightarrow 1{:}2.8{:}4 \Leftrightarrow 12.9\,\%{:}35.7\,\%{:}51.4\,\%. \quad (8)$$

One can see, that according to new approach seven ATC received rectified main dominants that allows to state that implementation of such approach increased precision of ATC attitude by 10 %. Table 2 shows main dominants redistribution.

Table 2 Main decision taking dominants redistribution for flight norm continuum S = 20 km

Values found according to $RR_{S_{0.5}}$		Values found according to RR_L	
1	2	3	4
T	35 (50 %)	T	35 (100 %)
		I	–
		N	–
I	5 (7.1 %)	T	–
		I	4 (80 %)
		N	1 (20 %)
N	30 (42.9 %)	T	1 (3.3 %)
		I	5 (16.7 %)
		N	24 (80 %)

Table 3 Main decision taking dominants redistribution for flight norm continuum S = 10 km

Values found according to $RR_{S_{0.5}}$		Values found according to RR_L	
1	2	3	4
T	38 (54.3 %)	T	34 (89.5 %)
		N	–
		I	4 (10.5 %)
N	2 (2.8 %)	T	–
		N	2 (100 %)
		I	–
I	30 (42.9 %)	T	9 (30 %)
		N	1 (3.3 %)
		I	20 (66.7 %)

As it comes from Table 2 rectification of main dominant happened mostly due to main "inclined to risk" ATC dominants redistribution. For corresponding nonlinear estimate functions referred attitude tendencies are kept for all continuum of researched distances between aircrafts. That is confirmed by 100 % of respondents that are "inclined to risk" which has kept their dominants after its determination criterion rectification. Same effect is observed for respondents "non-inclined to risk" since 80 % of them has kept same dominant after rectification.

In the same way researches of ATC attitude to risk for flight norm of 10 km. were carried out. Redistribution of main dominants is shown in Table 3.

As a result proportions of ATC respondents that tends, indifferent and not tends to risk may be presented:

- by criterion (1)

$$N{:}I{:}T \Leftrightarrow 2{:}30{:}38 \Leftrightarrow 1{:}15{:}19 \Leftrightarrow 2.8\,\%{:}42.9\,\%{:}54.3\,\%;$$

- by criterion (7)

$$N{:}I{:}T \Leftrightarrow 3{:}24{:}43 \Leftrightarrow 1{:}8{:}14.3 \Leftrightarrow 4.3\,\%{:}34.3\,\%{:}61.4\,\%.$$

So with help of rectified criteria (7) risk premium determination results for 14 respondents (20 %) were changed that witness about efficient and reliable criterion. Redistribution of main dominants concerned mostly ATC "indifferent to risk" (Table 3). For 89.5 % of respondents that were "inclined to risk" that dominant stayed the stable. Same is observed for ATC "non-inclined to risk" since 100 % of them kept their dominant stable.

5 Conclusions

Basing on received and presented new scientific results about proactive air operators main decision taking dominants determination it is expedient to point out the following main achievements.

New method for main decision taking dominants determination that is based on risk premium for five key points of estimate usefulness function was proposed and successfully tested. Method efficiency is defined by 20 % increase in main decision taking dominants precision.

Further researches should be help in direction of:

- spreading proposed method into all variety of flight norms recommended by ICAO;
- implementing of received results into ATC educational process;
- development of intellectual solution taking support module for ATC.

References

1. Edwards, W.: Solution taking/human factor. In: Activity Modeling, Professional Education and Operators Selection. Part I—Psychological Activity Models, vol. 3–6. Mir, Moscow (1991)
2. Reva, A.N., Tymishev, K.M., Bekmuhambetov, A.A.: Human Factor and Flight Safety: (Proactive Research of Influences): Monograph. Almaty (2006)
3. Karapetyan, G.S., Mihaylik, N.F., Pichko, S.P., Prokofyev, A.I.: Unfavorable Events Caused by Crew Action Inflight Provision. Mir, Moscow (1989)
4. Human Factors Digest No 11.: Human Factors in CNS/ATM Systems. Circular 249-AN/149. ICAO, Montreal (1994)
5. Doc 9806 AN/763.: Human Factors Guidelines for Safety Audits Manual. ICAO, Montreal (2002)
6. Leychenko, S.D., Malyshevsky, A.V., Mihaylik, M.F.: Human Factor in Aviation: Monograph in Two Books. SPb, Kirovohrad (2006)
7. Doc 9859 AN/474 ICAO.: Safety Management Manual (SMM). ICAO, Montreal (2013)
8. Reva, O.M.: Human factor: paradox of pilot activity psychological dominant in stochastic risk conditions. Airnavigation Problems, vol. 3, pp. 40–49 (1997)
9. Bekmuhambetov, A.A.: Operator Activity Improvement on the Basis of Risk Control Theory and Practice During Flight Safety Provision. SPb (2005)
10. Muhtarov, P.S.: Main dominants in air traffic controller solution taking at estimating usefulness-safety of flight air space level norm. In: Aviation-space Technique and Technology, vol. 116, pp. 143–150 (2014)
11. Reva, O.M., Muhtarov, P.S., Mirzoev, B.M.: Main solution taking dominant stability in conditions of risk. In: Aviation-Space Technique and Technology, vol. 117, pp. 147–153 (2014)
12. Borsuk, S.P.: Air traffic controllers main behavior dominant determination in conditions of flight level norm violation. In: Science-Based Technologies, vol. 27, pp. 261–265 (2015)

13. Kini, R.L., Raypha, H.: Taking Solutions at Many Preferences and Substitution Criteria. Radio I svyaz, Moscow (1981)
14. Reliability and efficiency in technique. In: Handbook in 10 volumes. Efficiency of Technical Systems, vol. 3. Mashinostroenie, Moscow, (1988)

Part III
Road and Rail—Pedestrians and Intersections

Quantitative Evaluation of Orientation Performance of Tactile Walking Surface Indicators for the Blind

Shinji Takahashi, Tatsuki Ishibashi, Katsuya Sato, Shin-ichi Ito and Shoichiro Fujisawa

Abstract Tactile walking surface indicators (TWSIs) are installed on roads to support independent travel for the blind. There are two types of TWSIs, attention patterns and guiding patterns. The attention pattern is usually installed at the crosswalk entrances. The direction of the crossing can be acquired by the row of the projection of the attention pattern through the soles of the shoes. In addition, truncated domes or cones of the attention pattern were arranged in a square grid, parallel or diagonal at 45° to the principal direction of travel. However, the international standard organization (ISO) allows a wide-ranging size. In this research, the direction indicating performance was compared at the same intervals for the five diameters specified by the international standard.

Keywords Tactile walking surface indicators (TWSIs) · Attention patterns · Blind · Orientation · Crosswalk

1 Introduction

Tactile walking surface indicators (TWSIs) were invented in Japan in 1965. They are now used around the world to help visually impaired persons. There are two types of TWSIs, attention patterns and guiding patterns [1]. TWSIs are perceived

S. Takahashi · T. Ishibashi · K. Sato · S. Ito · S. Fujisawa (✉)
Institute of Technology and Science, Tokushima University, 2-1 Minamijosan-Jima, Tokushima, Tokushima 770-8506, Japan
e-mail: sfujisawa@tokushima-u.ac.jp

S. Takahashi
e-mail: c501045003@tokushima-u.ac.jp

T. Ishibashi
e-mail: c501532026@tokushima-u.ac.jp

K. Sato
e-mail: katsuyas@tokushima-u.ac.jp

S. Ito
e-mail: s.ito@tokushima-u.ac.jp

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_13

Table 1 Top diameter and corresponding spacing of truncated domes or cones

Top diameter of truncated domes or cones (mm)	Spacing (mm)
12	42–61
15	45–63
18	48–65
20	50–68
25	55–70

using a long white cane or through the soles of the shoes for the blind. The TWSI design of the attention pattern only calls attention to a hazard or to hazards and decision points. This attention pattern is usually constructed at the crosswalk entrance. The direction of the crossing can be acquired from the row of the projection of the attention pattern detected through the soles of the shoes [2–4]. TWSIs must have high detection sensitivity and a high ability to distinguish objects as well as excellent direction detection. An international standard for TWSIs was enacted in 2012 [5]. The current result of the research is reflected for this ISO [6–9]. However, the international standard organization (ISO) allows a widely-ranging size. The background that allowed the size with a wide-ranging ISO standard was the adoption of various sizes by each country. For instance, the top diameter of attention patterns was specified as 12–25 mm. The spacing was specified as 42–70 mm (Table 1). In addition, truncated domes or cones of the attention pattern were arranged in a square grid, parallel or diagonal at 45° to the principal direction of travel (Fig. 1). This diagonal will not indicate the direction of the crossing. This research, compared the direction indicating performance at the same intervals for the five diameters specified by the international standard.

2 Experiment Method

2.1 Experiment Outline

Truncated domes were arranged on the disc to secure omnidirectional performance. The subject enters the disc from a random direction and determines the direction of the row through the soles of their shoes. Five discs with different diameters were selected at random. Figure 2 shows the disc used in the experiment. The spacing of the truncated domes was assumed to be 60 mm as adopted by the national standard of Japan (JIS). The experimenter guides the subject on the disc. The subject faces in the direction of the row of dots sensed through the soles of his/her shoes. The angle of the gap with a positive direction (0°) was measured. The maximum gap angle is assumed to be 45°.

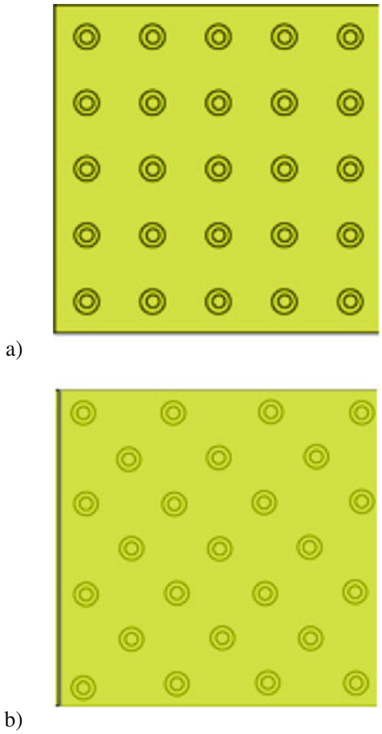


Fig. 1 Spacing and dimensions of truncated domes or cones. **a** Parallel to the principal direction of travel, **b** Diagonal at 45° to the principal direction of travel

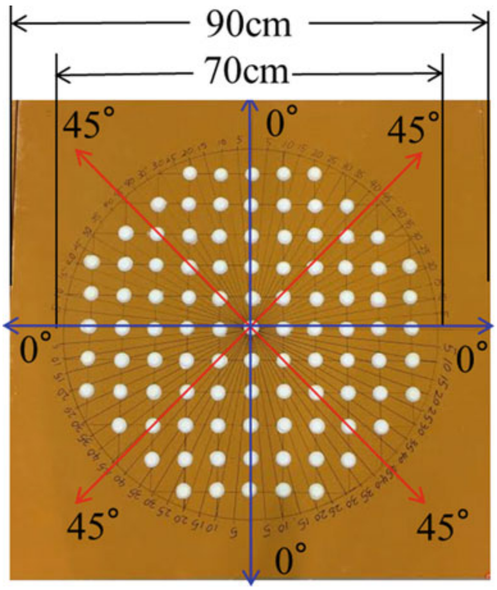


Fig. 2 Disc used in experiment

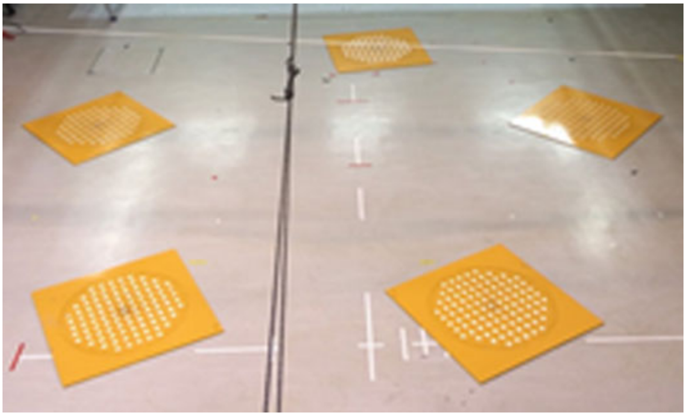


Fig. 3 Disc arrangement

2.2 Experiment Procedure

Five discs were arranged in the laboratory (Fig. 3). Each subject was arranged five discs at random. Five discs were measured ten times as a couple. One subject performed 50 times measurements. Figure 4 shows the subject turning in the direction of the row of the point at soles of the shoes. These discs were rotated at random every cycle. The experiment was performed by 20 blind people (11 male, 9 female).

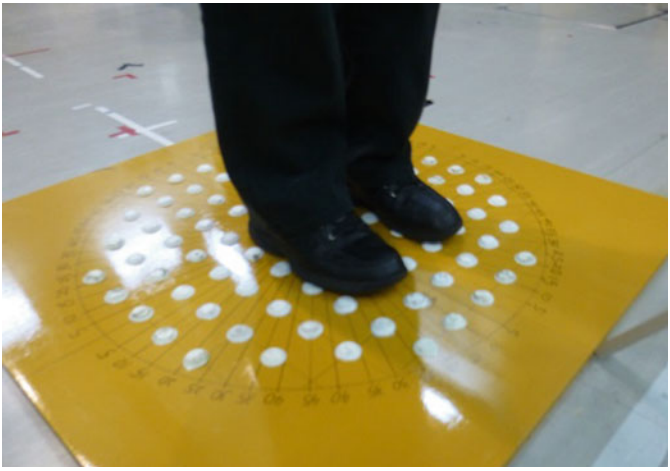


Fig. 4 Experiment

3 Experiment Result

The data of each disc obtained from the experiment are presented and summarized in Fig. 5 in one graph. There were 200 samples for every disc. The 12 mm diameter (JIS standard) increases when the graph is seen and the ratio of 0° has increased. For the five TWSLs, the direction was determined as the direction of the row at the rate from 43 to 54 %. The average was 51 %. Other angles were all less than 7 % at a resolution of 3°. The average was 5 %. Figure 6 shows the average and the standard deviation of 0 degree and other degrees for 12 mm diameter discs. Other diameters were similar to those in Figs. 6 and 7 plots the frequency distribution

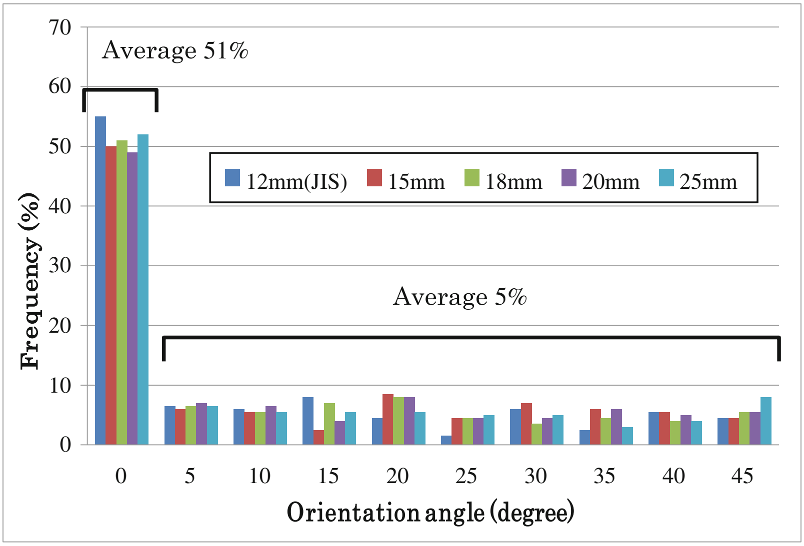


Fig. 5 Frequency distribution of each angle

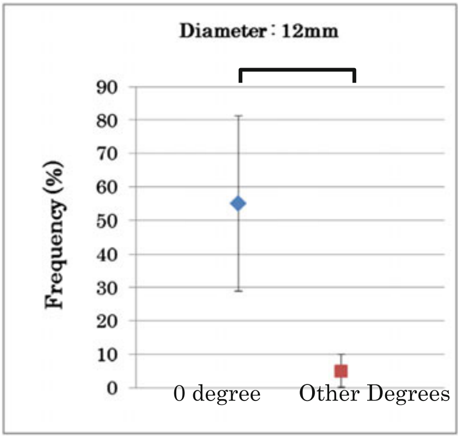


Fig. 6 Average and standard deviation of 0° and other degrees for 12 mm in diameter

when the seven subjects with a slow sense in the soles of the shoes are excluded. Figure 5 and 7 are compared, and the value of 0° increases from 51 to 64 %. Moreover, averages other than 0° have fallen from 5 to 4 %. When a person loses sight due to diabetic retinopathy, the sensitivity in the sole of the shoes decreases. The data obtained from the experiment are analyzed according to the angle. A significant difference was confirmed between 0° and other angles though there

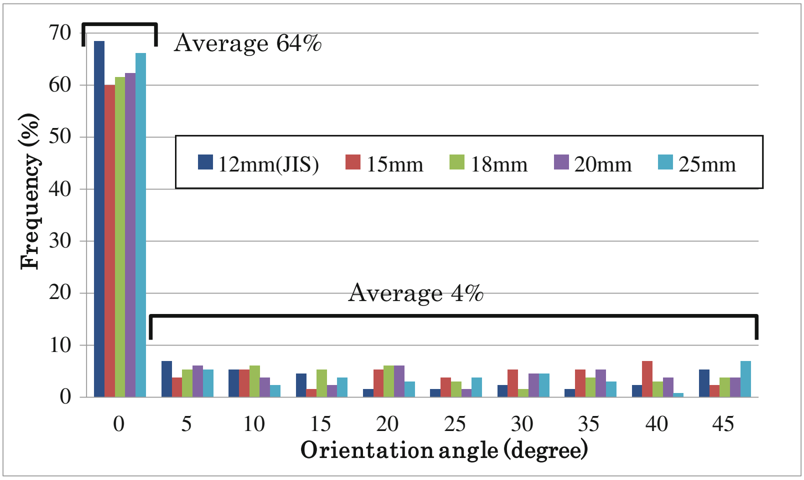


Fig. 7 Frequency distribution when seven subjects with slow sense in sole of the shoes are excluded

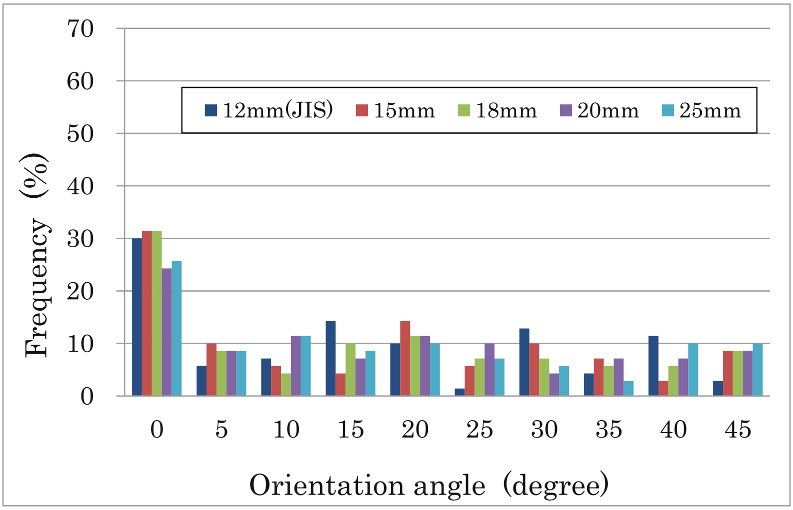


Fig. 8 Frequency distribution only of subjects with slow sense in sole of the shoes

was no significant difference in the difference of each diameter. Figure 8 shows the frequency distribution only of subjects with slow sense in sole of the shoes. It is understood not to be able to orientate from this figure.

4 Conclusions

It has been understood to have the orientation performance of all five kinds of attention patterns with a different diameter. The attention pattern of JIS was the highest value. This experiment does not verify the orientation performance of the difference in spacing, the attention pattern of the diagonal means the performance is not used. The row of 45° forms the row at the spacing of twice the square root of two. It was a result similar as for this row. The attention pattern has the orientation performance in a positive direction (0°), and the orientation performance is not observed at other angles. Therefore, the attention block of the diagonal array of the ISO standard does not provide orientation in the front direction. A blind pedestrian may be misoriented to an angle of 45° if this block is installed at a crosswalk entrance. The JIS arrangement uses a square grid only. This research, quantitatively compares the orientation performance for differences in diameter. The purpose of this research is to acquire a scientific basis for improving the ISO standard. This work was supported by JSPS KAKENHI Grant Number 15K01458.

References

1. JIS T 9251:2001.: Dimensions and patterns of raised parts of tactile ground surface indicators for blind persons (2001)
2. JIS TR T 0006:1999.: Methods for estimating probability and easiness of the visual impaired's recognition of bumps in tactile tiles through their soles of shoes (1999)
3. National Institute of Technology and Evaluation, Japan, Report of Fundamental Research on Standardization of Tactile Tiles for Guiding the Visually Impaired—Standardizing Patterns (Version 1.0) (1998)
4. National Institute of Technology and Evaluation, Japan, Report of Fundamental Research on Standardization of Tactile Tiles for Guiding the Visually Impaired—Targeting Standardizing Patterns (Version 1.0) (2000)
5. ISO 23599:2012.: Assistive products for blind and vision impaired persons—Tactile walking surface indicators (2012)
6. Mitani, S., Fujisawa, S., Sueda, O.: Detection and recognition of tactile walking surface indicators by white cane. In: Proceedings of the 12th International Mobility Conference, No. C2-2, Hong Kong, pp. 1–5 (2006)
7. Jacobson, R.D.: Cognitive mapping without sight: four preliminary studies of spatial learning. J. Environ. Psychol. **18**, 289–305 (1998)

8. Loomis, J.M., Klatzky, R.L., Golledge, R.G., Cicinelli, J.G., Pellegrino, J.W., Fry, P.A.: Nonvisual navigation by blind and sighted: assessment of path integration ability. J. Environ. Psychol. Gen. **122**, 73-91 (1993)
9. Crandall, W., Brabyn, J., Bentzen, B.L., Myers, L.: Remote infrared signage evaluation for transit stations and intersections. J. Rehabil. Res. Dev. **36**, 341–355 (1999)

Proof Experiment of LED Block Equipped with Projections to Locate Travel Direction for Blind and Vision Impaired Persons

Hideaki Nagahama, Tomoyuki Inagaki, Norihiro Ikeda, Kazuya Takahashi, Kiyohito Takeuchi, Hiroshi Ogino, Katsuya Sato, Sin-Ichi Ito, Motohiro Seiyama and Shoichiro Fujisawa

Abstract Crossing crosswalks is one of the most dangerous situations for visually impaired persons. Crosswalk entrances are located on the boundary with the roadway and are among the most dangerous areas for visually-impaired persons. Tactile walking surface indicators (TWSIs) are installed on the road to support

H. Nagahama · K. Sato · S.-I. Ito · M. Seiyama · S. Fujisawa (✉)
Institute of Technology and Science, Tokushima University,
2-1 Minamijosan-jima, Tokushima 770-8506, Japan
e-mail: sfujisawa@tokushima-u.ac.jp

H. Nagahama
e-mail: c501202107@tokushima-u.ac.jp

K. Sato
e-mail: katsuyas@tokushima-u.ac.jp

S.-I. Ito
e-mail: s.ito@tokushima-u.ac.jp

M. Seiyama
e-mail: c5014320379@tokushima-u.ac.jp

T. Inagaki
Nipon University, 7-24-1 Narashinodai, Funabashi, Chiba 274-8501, Japan
e-mail: inagaki.tomoyuki@nihon-u.ac.jp

N. Ikeda · K. Takeuchi · H. Ogino
Kictec Inc., 150 Umegaoka, Agui, Chita, Aichi 470-2295, Japan
e-mail: ikeda@kictec.co.jp

K. Takeuchi
e-mail: ki_takeuchi@kictec.co.jp

H. Ogino
e-mail: h_ogino@kictec.co.jp

K. Takahashi
General Support Co. for the Visually Impaired, 2-37-10 Kamiogi, Suginami, Tokyo 167-0043, Japan
e-mail: saraseina@y3.dion.ne.jp

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_14

independent travel for blind and low-vision individuals. We developed an LED block equipped with projections to strengthen the support of the crosswalk entrance. The blind person can sense this block through the soles of the shoes. A low-vision person can find it through the use of residual vision. Blind and visually impaired persons found that the LED block equipped with projections to locate the travel direction effectively supported crossing. The effectiveness of this block is described based on measurements and questionnaire results.

Keywords Person with low visual capacity (LV) · LED block · Blind person · Crosswalk · Orientation · Veering distance · Tactile walking surface indicators (TWSIs)

1 Introduction

The vision of a person with low visual capacity ("LV") might decrease remarkably under low-light conditions such as the morning, evening, and nighttime, decreasing their opportunities to go out. Crosswalk entrances are located on the boundary with the roadway and are among the most dangerous areas for visually impaired persons. The guidance of visually impaired person has the result of a lot of researches [1–10]. An LV person cannot easily find the entrance of a crosswalk at nighttime. However, LV persons can detect the developed LED blocks with their residual vision. Moreover, for a blind person to cross the crosswalk safely, it is very important to acquire the orientation of the soles of the shoes. The authors have verified the effectiveness of the developed block for visually impaired person using laboratory experiments [11, 12]. In this experiment, both visually impaired persons and blind persons found the developed block to effectively support crossing. Moreover, the time required for locating the LED block was short. Orientation using the LED block required less time. The LED block was observed to have an orientation accuracy equivalent to that of TWSIs. Moreover, subjects expressed opinions such as "It was easy to locate" and "There is a sense of security when going straight" in interviews after the experiment. Installing LED blocks at crosswalk entrances was found to be useful for guiding visually-impaired persons. Moreover, the shape of the projection indicating the direction is easily verified through the soles of the blind person's shoes. A LED block experiment was conducted to determine its actual efficiency. In this research, the visibility of the LED block was verified for different luminescence and illuminance environments.

The results clarified how to provide illumination for a visually impaired person. Based on this research, the authors developed a LED block with projections to determine the travel direction for a visually impaired person. The developed block can be detected with an LV's residual vision. Blind persons can also detect this block through the soles of their shoes. This block was installed at the entrance to a crosswalk of an actual intersection, and a proof experiment was conducted. Twenty persons with weak sight and 20 totally blind people verified the effectiveness of this block. In particular, the effectiveness of the experiment on the LED block for the LVs was verified by conducting the experiment after sunset. This experiment measured the visibility of the LED block on the pavement side and the opposite side. In the experiments with blind persons, the time to acquire the direction of the crossing from the direction location block and the veering distance at the crossing was measured to verify its effectiveness. Moreover, low-vision persons and totally blind persons answered a questionnaire concerning the effectiveness of the block. The effectiveness of the block was determined based on measurement data and questionnaire results.

2 Experiment Conditions

2.1 Developed LED Block Equipped with Projections

To increase the support of the crosswalk entrance, we developed an LED block equipped with projections. Figure 1 presents the specifications for this LED block. The LED block utilizes a solar power supply. A low-vision person can find it through the use of residual vision. The surface has two projections of 5 mm high, flat-topped elongated bars. The height of the flat-topped elongated bars is 5 mm. The top width of the flat-topped elongated bars is 7 mm. The spacing between the axes of the flat-topped elongated bars is 47 mm. The length of the top of the flat-topped elongated bars is 262 mm. A blind person can find this block through the soles of the shoes. Figure 2 is a photo of the LED block installed at a crosswalk entrance.

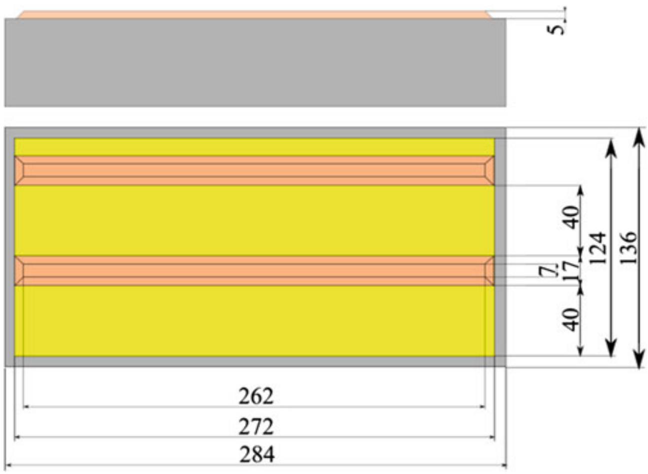


Fig. 1 Spacing and dimensions of LED block

Fig. 2 LED block installed at crosswalk entrance

Fig. 3 The LED block was installed at an actual crosswalk entrance

2.2 *Experiment Environment*

Figure 3 is a photograph showing where the LED block was installed at an actual crosswalk entrance. To verify the effect of the LED block, the LED block is installed only on the other hand. The experiment using blind persons was conducted during the day. The experiment using low-vision persons was conducted at sunset. Figure 4 illustrates the size of the crosswalk. The crossing distance is about 10 m, and the crosswalk is 3.6 m wide.

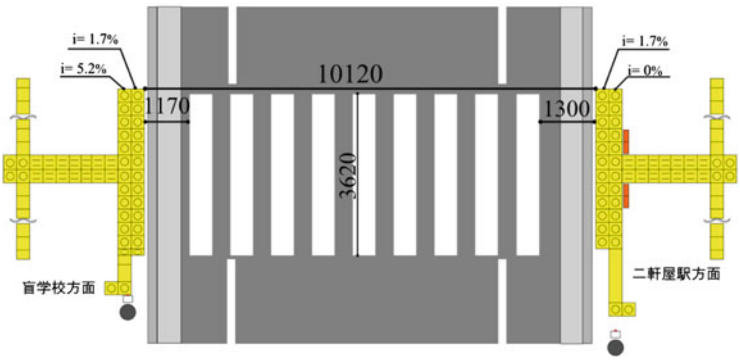


Fig. 4 Size of experimented actual crosswalk (mm)

3 Experiment Procedure

3.1 Blind Experiment

Blind persons can detect the developed block through the soles of their shoes. Seventeen blind people verified the effectiveness of this block. Figure 5 indicates where the experiment using a blind person is started. The time required to travel from this start position to the crosswalk entrance is measured. The time required until this LED block is discovered is measured at the crosswalk entrance. The subject travels over to the opposite side after he or she arrives at the crosswalk entrance. The time required to cross and the veering distance are measured. The veering distance is measured by the number of TWSIs. Round-trip crossing experiments were conducted three times.

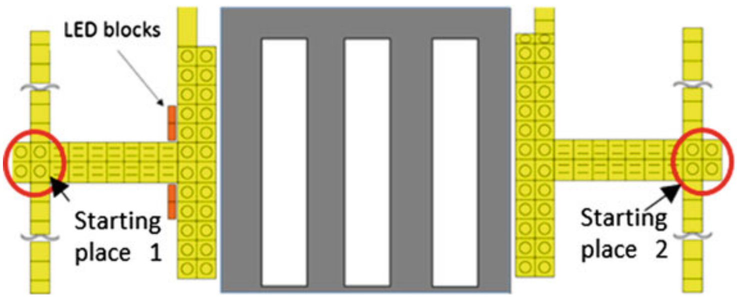


Fig. 5 The decision point is the position where the experiment using a blind person is started

3.2 Low-Vision Experiment

The developed block can be detected with an LV's residual vision. Twenty LV people verified the effectiveness of this block. The first experiment measured the visual distance at which the crosswalk entrance was found from the pavement. The next experiment measured the time required until crossing from the crosswalk entrance to the opposite side. This experiment measured the visual distance for discovering the opposite side while crossing the crosswalk. Round-trip crossing experiments were conducted three times.

4 Experiment Result

4.1 Blind Experiment

Table 1 summarizes the time required for each movement. Much time is needed to find the LED block when discovering the crosswalk entrance because this block was installed on both sides of the guiding pattern block leading to the crosswalk entrance. The LED block did not influence the time required to cross. Figure 6 shows a histogram of the shift in the lateral direction. The horizontal axis is the number of shifted TWSIs. The TWSI is 300 mm wide. The vertical axis is frequency. Figure 7 plots the average and standard deviation of the lateral shift. There was a test of significance revealed a significant tendency. Subjects were able to go straight in the crosswalk by stepping on the two bars with the soles of the shoes. Figure 8 presents the result of a questionnaire administered after the experiments. The questions were as follows.

1. Was it easy to find the LED blocks?
2. Was it easy to confirm the two bars?
3. Was it easy to determine the crossing direction?
4. May I install the LED block?

Table 1 Time required of each movement

	From the start position to the crosswalk entrance (s)	Crossing (s)
With support of LED Block	29.5	11.4
Without support of LED Block	8.9	11.1

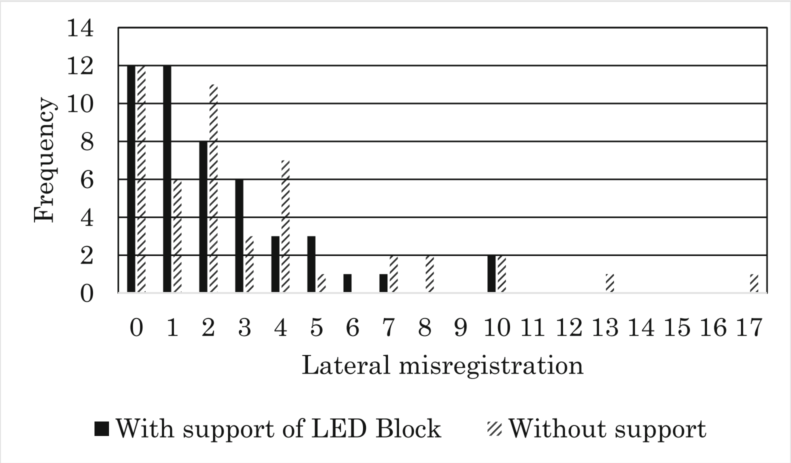


Fig. 6 Histogram of the shift in the horizontal direction of the crossing

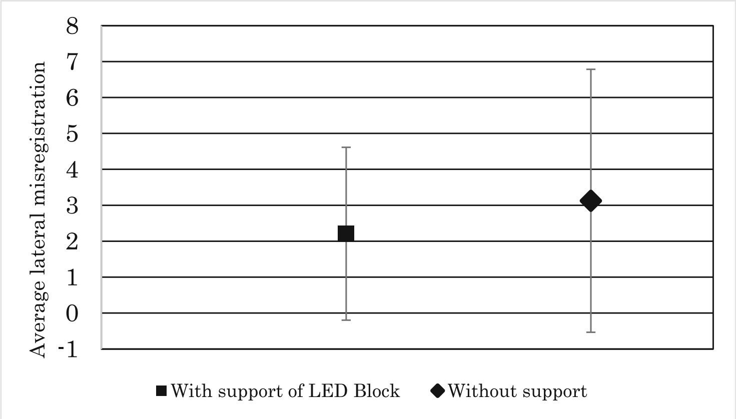
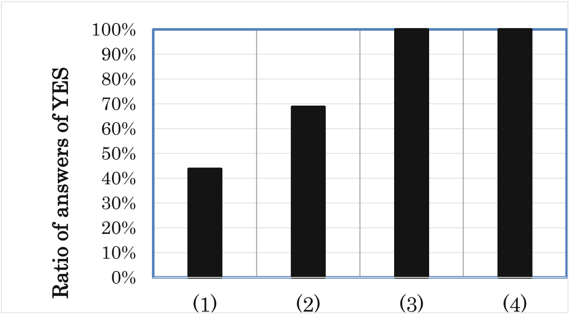


Fig. 7 Average and standard deviation of the shift in the lateral direction

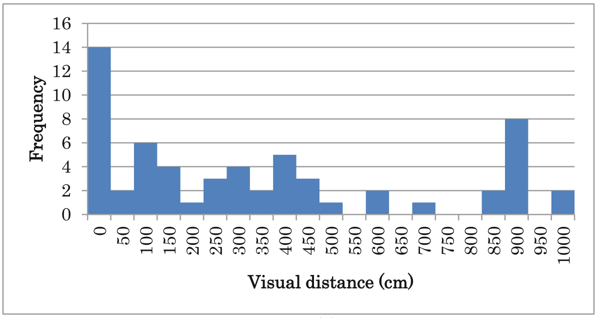
4.2 Low-Vision Capacity Experiment

Figure 9 depicts the visual distance at which the crosswalk entrance was found from the pavement. It is understood that the range of values is large because each subject's symptoms are different. Figure 10 plots the average and standard deviation of the visual distance at which the crosswalk entrance was found from the pavement. There was no significant difference. Figure 11 depicts the visual distance at which subjects were able to discover the opposite side while crossing the crosswalk. Figure 12 plots the average and standard deviation of the visual distance

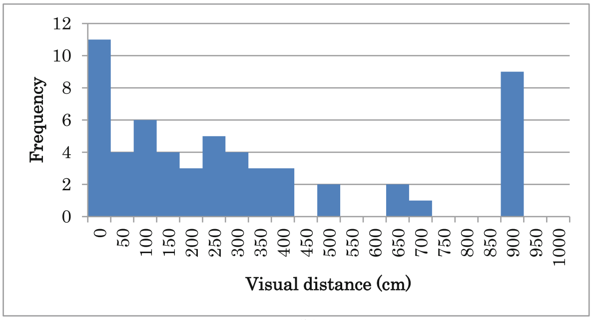


(1) Was it easy to find the LED blocks?
(2) Was it easy to confirm the two bars?
(3) Was it easy to determine the crossing direction?
(4) May I install the LED block?

Fig. 8 Questionnaire result for blind persons



(a)



(b)

Fig. 9 Visual distance from the pavement at which the crosswalk entrance was found. **a** With LED blocks, **b** without LED blocks

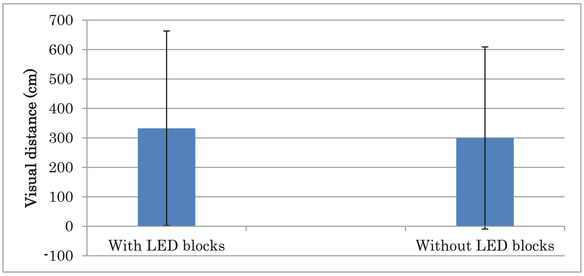


Fig. 10 Average and standard deviation of the visual distance in which the crosswalk entrance was found from the pavement

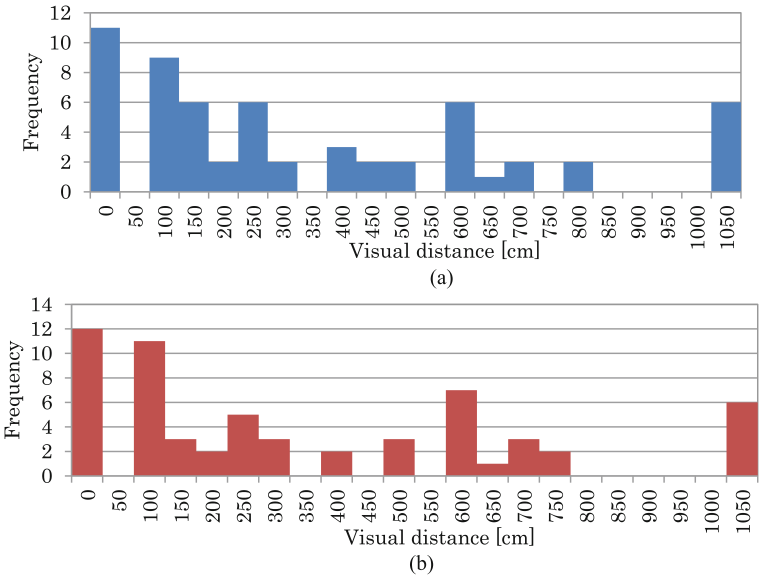


Fig. 11 T Visual distance at which subjects were able to discover the opposite side while crossing the crosswalk. **a** With LED Blocks, **b** Without LED Blocks

at which subjects were able to discover the opposite side while crossing the crosswalk. The reason for the results is that experiment management was made difficult by differences in illumination and by distractions such as the headlights of cars in an actual intersection. Figure 13 summarizes the results of a questionnaire administered after the experiments. The question items were as follows.

1. Was it easy to find the LED blocks?
2. Was it easy to determine the crossing direction?
3. Was it easy to find the LED block on the opposite side?
4. May I install the LED block?

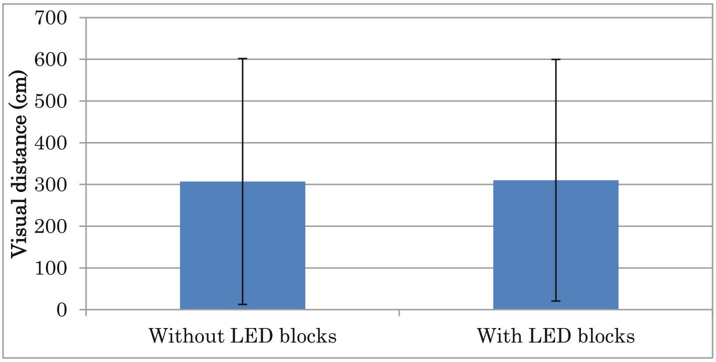


Fig. 12 Average and standard deviation of the visual distance at which subjects were able to acquire the opposite side while crossing the crosswalk

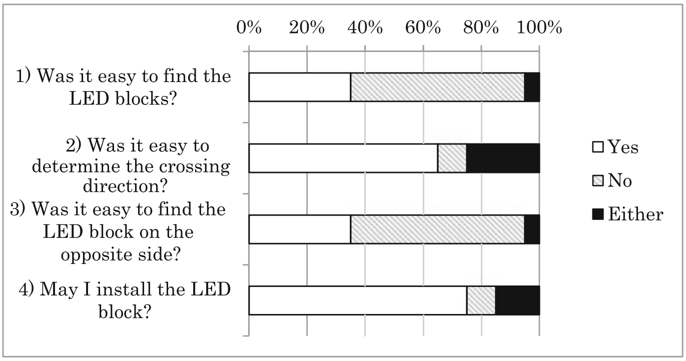


Fig. 13 Questionnaire result for low-vision persons

5 Conclusions

Blind persons can detect this block through the soles of their shoes. The block was installed at the entrance to the crosswalk of an actual intersection, and an experiment was conducted to verify its effectiveness. In experiments with blind persons, the time to acquire the direction of the crossing from the direction location block and the veering distance at the crossing were measured. In experiments with low-vision persons, it was difficult to determine the effectiveness due to the headlights of cars in the actual intersection and differences in the lighting environment. After the experiments, the low-vision and totally blind persons answered a questionnaire concerning the effectiveness of the block. Responses from blind persons confirmed that all persons could use the two bars easily and easily determine the direction of the crossing. Low-vision persons also easily determined the direction of the crossing. There were a lot of answers of wanting blind person and a low vision person to install this block. Many blind persons and low-vision persons wanted this block installed.

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. Vis. 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. Assistive Technol. Res. Ser. **29**, 595–603 (2011)
5. Loomis, J.M., Philbeck, P.Z.: Dissociation between location and shape in visual space. J. Exp. Psychol. **28**(2), 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. Assistive 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. Assistive 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. Assistive 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. Assistive 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. Procedia Manuf. **3**, 3774–3778 (2015)

Effect of Tire Pressure to Driving Forces at a Wheelchair

Masayuki Booka, Hidehisa Oku, Ikuo Yoneda and Shoichiro Fujisawa

Abstract If the same effect is obtained by less driving force in a manually propelled wheelchair, necessary user's physical and mental capability can be decreased. On the other hand, many external factors not relating with wheelchairs and internal factors related with wheelchairs affect the driving force necessary to propel manual wheelchairs appropriately. In these circumstances, the purpose of our research is to clarify the relation between the tire pressure of the driving wheel and the required driving force in manually propelled wheelchairs. For this purpose, a clinical testing to measure required driving forces in different tire pressures at a manually propelled wheelchair has been carried out. The result of the testing indicated that the required driving forces increased according to decrease of the tire pressure. This objectively revealed that appropriate tire pressure in manually propelled wheelchairs is the one of important factors to reduce driving force.

Keywords Manually propelled wheelchair · Tire pressure · Driving force · Quantitative evaluation

M. Booka (✉)
Faculty of Rehabilitation, Hiroshima International University,
555-36 Kurosegakuendai, Higashihiroshima, Hiroshima, Japan
e-mail: m-booka@hw.hirokoku-u.ac.jp

H. Oku
Kobe Gakuin University, Hyogo, Kobe, Japan

I. Yoneda
Nishikyushu University, Kanzaki, Saga, Japan

S. Fujisawa
The University of Tokushima, Tokushima, Tokushima, Japan

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_15

1 Introduction

A factor to influence the driving force of the manually propelled wheelchair (here after, wheelchair) includes the road surface environment such as a step, the incline. Many researches for these external factors have been performed until now. In addition, internal factor of the wheelchair including the mounting location of the large wheels have been also researched. However, relation between tire pressure of the large wheels and driving force has not been analyzed so far.

The purpose of our research is to analyze influence of tire pressure on driving force in wheelchairs. This paper describes the result of clinical testing that driving forces were measured at different tire pressures. The wheelchair used in this research was the standard one that specification was defined in JIS T 9201.

2 Method

2.1 Setting of a Certain Pressure in Wheelchair's Tire

To set a certain pressure to wheelchair's tire, air of high pressure is filled into the tire through a valve. In this process, the tire pressure is measured by a pressure gauge being attached to either a compressor or bicycle pump. On the other hand, a British valve has been generally used as the valve in wheelchair tire. As this valve has rubber to stop the leak of inside air, higher pressure is necessary to push the rubber to measure the inside pressure. So, the authors developed the tire pressure measurement system at first. Figure 1 shows the block diagram of the system. Figure 2 shows the developed system. As shown in these figures, the developed measurement system can monitor the tire pressure statically.

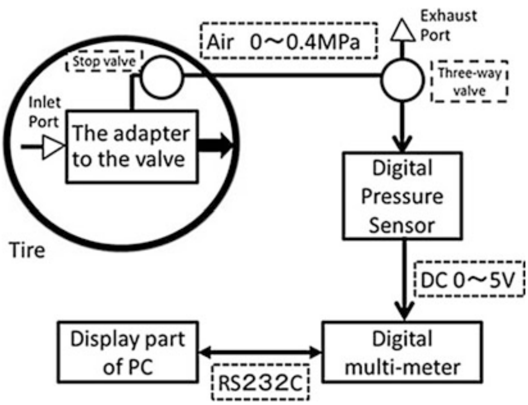


Fig. 1 Block diagram of the tire pressure measurement system

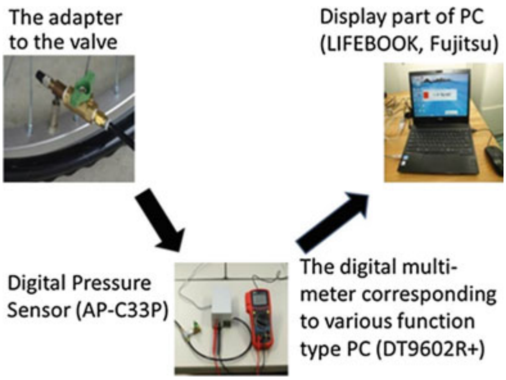


Fig. 2 The developed tire pressure measurement system

Fig. 3 Tire pressure indicator

This system is, however, is not useful to measure the tire pressure dynamically because of necessity of connecting tube between the tire valve and the system. Then, a tire pressure indicator was developed to measure tire pressure dynamically without the connection to other measurement system.

The developed tire pressure indicator is shown in Fig. 3. This tire pressure indicator consists of both a pressure gauge and an additional British valve to set the tire pressure by a compressor or bicycle pump from the outside. The tire pressure is always indicated to the gauge with or without connection to outside high air pressure source. To inspect the accuracy of the developed tire pressure indicator, experiment for calibration was carried out. The experiment for calibration was done by comparing the input pressure to this pressure indicator and indicated pressure. As the tire pressure would be decreased from a standard value to lower value progressively in the target experiment described later, the input pressure in the experiment of calibration was also decreased progressively. In the experiment of the calibration, input tire pressure was inputted from 0.3 to 0.02 Mpa with interval of 0.02 Mpa, and the tire pressure indicated in the gauge at each input was recorded. Figure 4 is the result of experiment of the calibration. In the graph, vertical axis indicates input tire pressure, and the horizontal axis is indicated tire pressure in the gauge. The coefficient of correlation between the input air pressure and the

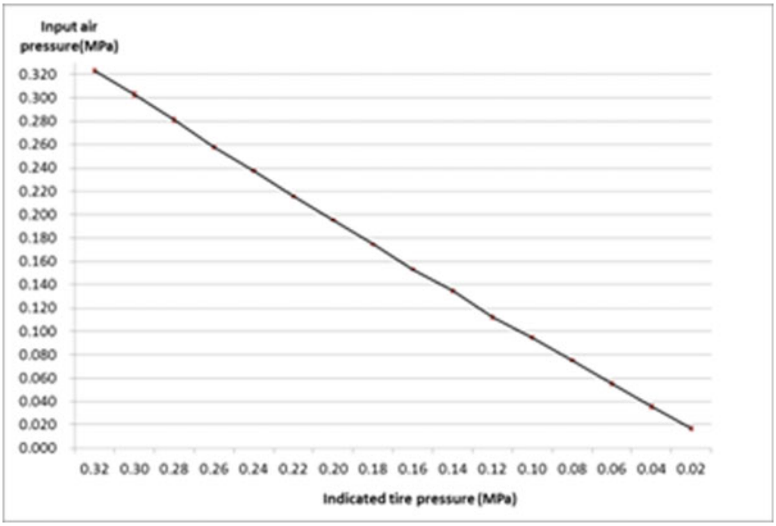


Fig. 4 Relation between the input air pressure and the indicated tire pressure

indicated tire pressure was 1.00, and it was confirmed that the pressure indicator shows the input air pressure correctly.

2.2 Measurement of Both Torque to Hand Rim and Numbers of Rotation in Driving Axis

As described previously, the target experiment was carried out to measure both deft torque to hand rim and moved distance of the wheelchair. The wheelchair used in this experiment was the one in conformity with the Japanese Industrial Standards (JIS). The dimension of the wheelchair was a driving wheel of 600 mmϕ, a hand rim of 554 mmϕ, and weight of 318 N.

To measure torque added to hand rim, a torque converter (Kyowa Electronic Instruments Co. Ltd., TP-10KMSA84F) was attached to axis of each driving wheel to measure. By this, input torque was converted to DC voltage from 0.00 to 2.00 V. This output signal was amplified and converted to digital data by using an A/D converter (Interface Corporation, CSI-320212) to input PC. The sampling rate of this analogue to digital conversion was 50 ms. The converted data was stored in the laptop PC attached to the wheelchair.

The number of rotation in driving axis was measured by encoder count card (Interface Corporation, CSI-631204), the data was stored in the laptop PC by the same procedure as torque measurement.

2.3 Subjects

Subjects were 11 students (9 males and 2 females) without disabilities. Their ages were 21.4 years old on the average. Their profile was shown in Table 1 and all of them had experience of having used wheelchairs so far.

2.4 Procedure of the Experiment

In three kinds of different tire pressures, the experiments of the same content were conducted. The three kinds of tire pressure were 0.3 Mpa that is generally defined as appropriate, 0.2 Mpa (2/3 of the general), and 0.1 Mpa (1/3 of the general). The developed tire pressure indicator was used to set each tire pressure. Two types of courses for test run in the experiment were prepared. Dimension of the each course was 10.00 m of length and 1.50 m of width, and one of large halls in our university was utilized. The road surface of the one course was made of semi-hard composition tile. The road surface of the other one was made of tile carpet.

In each experiment, prior to the test run, the subject recorded the condition of the experiment into the laptop PC. Table 2 shows the content of the condition. Then, the subject was asked to set down on the wheelchair's sheet. The experiment was conducted as follows according to the experimenter's instruction.

1. The experimenter starts the measurement system.
2. The subject is asked both to run the wheelchair forward by turning both hand-rims once and to keep the same posture till the wheelchair will stop.
3. After stopping of the wheelchair, the experimenter stops the measurement system.
4. The subject returns the wheelchair to the start position.

Table 1 Profile of the subjects

Sex (male or female)	Age (years old)	Height (cm)	Weight (N)
M	21	162.0	671.3
M	21	175.0	695.8
M	21	170.0	656.6
M	21	165.2	672.8
M	21	175.0	666.4
M	21	174.0	709.5
M	25	168.0	901.6
M	21	170.0	625.2
M	21	170.0	580.5
F	21	156.0	455.2
F	21	157.0	535.6

Table 2 The content of the recorded condition in each experiment

Date of the experiment	Year, mouth, day
Subject's name	A maximum of 20 word
Subject's sex	Male or female
Subject's age	0–125 years old
Subject's height	50.0–200.0 cm
Subject's body weight	2.0–250.0 kg
Location of the experiment	A maximum of 40 word
Comment	A maximum of 60 word
The maximum measurement time	1–3600 s
Sampling frequency	10, 50, 100, 200 Hz (select)

This is one measurement in a certain tire pressure. This measurement is repeated 3 times. These 3 measurements are the result of the experiment in a certain tire pressure. After 3 times of measurements in a certain tire pressure, the same measurement starts in another tire pressure. Finally, one subject has measurement 9 times in three kinds of different tire pressures.

2.5 Evaluation Methods

In this experiment, both driving force to hand rims and numbers of revolutions were measured. Based on these data, moved distance of the wheelchair, run time, run speed, Deceleration ratio, and total torque were calculated.

Moved Distance of the Wheelchair. The distance that the wheelchair ran is calculated from the number of revolutions of the tire and the diameter of the tire. In this case, it should be considered that the diameter of tire is decreased according to the decrease of tire pressure. Therefore, an additional experiment was carried out to find reduction ratio of diameter that was decided by the tire diameter and the tire pressure. In the additional experiment, a person of 490 N ran 3 m straight road 10 times on the wheelchair which tires had reduced pressure. The reduction ratio the diameter in each reduced tire pressure was calculated by obtained data. Using a revised diameter in each tire pressure, the right moved distance of the wheelchair was re-calculated.

Speed of the Wheelchair. From the wheelchair's moved time and moved distance, speed of the wheelchair was calculated as moved distance in unit time.

Deceleration Ratio. In flat place, the wheelchair is decelerated naturally and stops when no additional power is added to the hand rims. This is caused by negative acceleration generated by the friction between tire and the road surface. In this experiment, each subject is asked to turn hand rims only once. The driving force to the hand rims had disappeared just after this operation. So, both the speed of the wheelchair at this time and interval time before the wheelchair stopped were measured, and negative acceleration was calculated by this speed and interval time. In this research, the deceleration ratio was defined as absolute value of this negative acceleration.

Load/Distance Rate. Cooper reported that necessary power to drive a wheelchair is calculated from driving torque added to the hand rims. According to this, the driving torque is integrated by time, and is used as a parameter to evaluate the power to move wheelchair. Total torque is divided by the moved distance, and the obtained momentum (N m/s) is used as Load/Distance rate in this experiment.

3 Result and Discussion

Figure 5 indicates the one of result of the measurement on the torque, moved distance by the wheelchair, and speed of the wheelchair. The vertical axis indicates the torque (N m), moved distance (m), and speed (m) respectively. The horizontal axis indicates the elapsed time (seconds) from the start of the experiment.

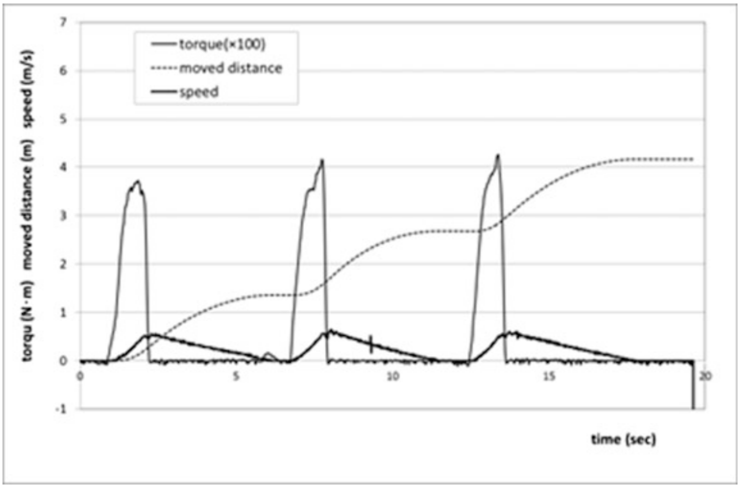


Fig. 5 An example of the measurement result on the torque, moved distance by the wheelchair, and speed of the wheelchair

3.1 Deceleration Ratio

Figure 6 indicates the deceleration ratio at each tire pressure on the hard floor.

The distribution of the deceleration ratio in each subject was indicated in Fig. 7. As shown in Fig. 7, relation between the deceleration ratio and the tire pressure in the hard floor is indicated as the regression line described below.

$$y = -0.1466x + 0.1726 \quad (1)$$

Then, the deceleration ratio on the semi-soft floor is shown in Fig. 8, and the distribution of the deceleration ratio in each subject was also indicated in Fig. 9. As shown in Fig. 9, relation between the deceleration ratio and the tire pressure in semi-hard floor is indicated as the regression line described below.

$$y = -0.2058x + 0.2727 \quad (2)$$

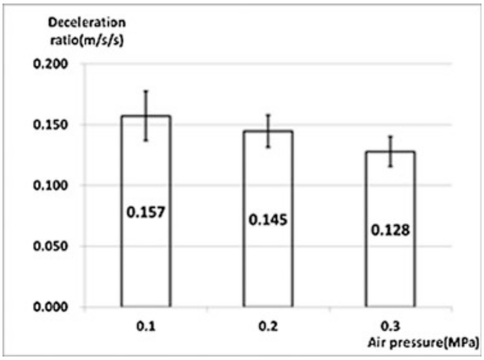


Fig. 6 Deceleration ratio on the hard floor

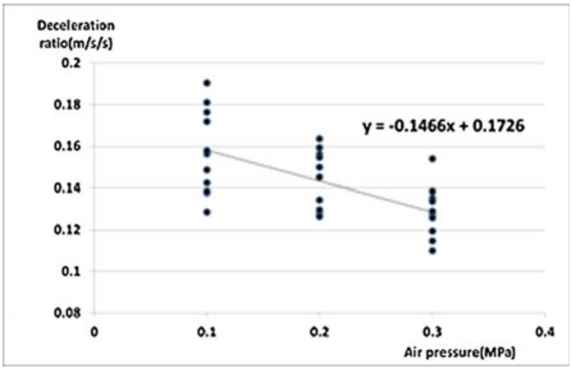


Fig. 7 Distribution of each subject's deceleration ratio on the hard floor

Fig. 8 Deceleration ratio on the semi-hard floor

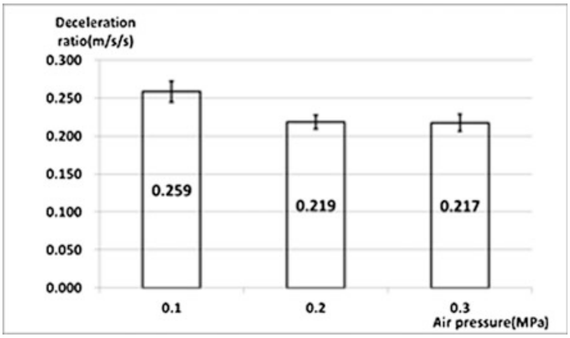
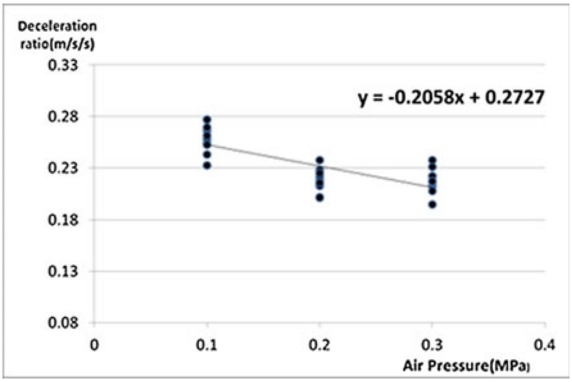


Fig. 9 Distribution of each subject's deceleration ratio on the semi-hard floor



From these results, it was indicated that the decrease of tire pressure increases the deceleration ratio. This tendency appeared more conspicuously in the semi-hard floor than in the hard floor. Although the reason of tendency may be caused by the difference of the materials used for each road, further analysis is necessary in the next step.

3.2 Load/Distance Rate

Figure 10 shows the load/distance rate in the hard floor. The load/distance rate in semi-hard floor is also showed in Fig. 11.

There are not obvious relations between the deceleration ratio and tire pressure when tire pressure was decreased. Although it is subjective, it may be caused by the difference in subjects' postures at driving wheelchairs. For an example, when posture of a subject bends forward, the center of gravity in wheelchair including the subject moves forward. By this phenomenon, the caster of the wheelchair receives

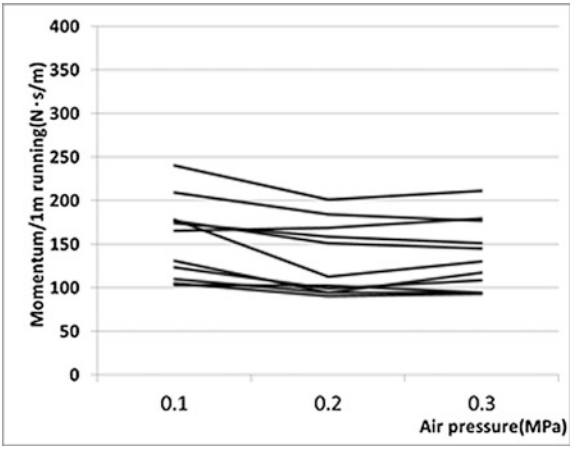


Fig. 10 Load/distance rate on the hard floor

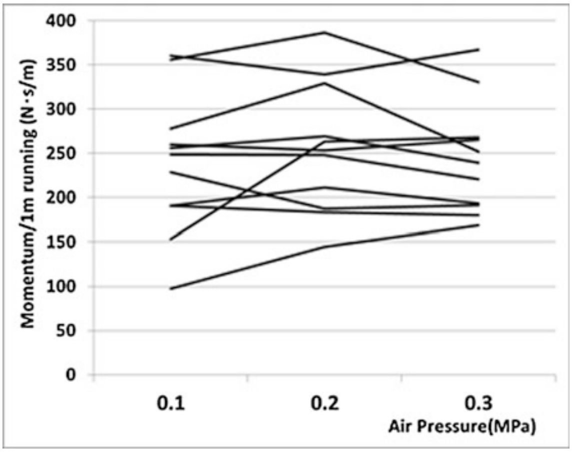


Fig. 11 Load/distance rate on the semi-hard floor

more weight than before. As the rolling resistance of the caster is larger than that of driving wheels, the speed of the wheelchair should be decreased. As the result, it is considered that moved distance of the wheelchair becomes short.

4 Conclusion

In this research, drivability of a wheelchair after reducing tire pressure was experimentally evaluated from the point of the increase and decrease on necessary driving force. As the result, it was indicated that a wheelchair with driving wheels of general tire pressure can run longer distance than a wheelchair with driving wheels of less tire pressure, when given the same driving force. Then, it was indicated that the increase of deceleration ratio caused this.

On the other hand, the decrease of the tire pressure in wheelchairs increased the deceleration ratio whether road had hard surface or soft surface. Although analysis from load/distance rate was conducted to get the answer on this, no objective reason was found. In addition to the tire pressure, it is generally said that the posture of a person on a wheelchair affects to this kind of problem. The further study to this issue will be necessary in the future.

Development of an Effective Pedestrian Simulator for Research

Richard Sween, Shuchisnigdha Deb, Daniel W. Carruth, Daniel Waddell and Masakazu Furuichi

Abstract According to the US Department of Transportation, in 2013 14 % of all traffic fatalities were pedestrians. In Japan, 38.4 % of 2015 traffic fatalities were pedestrians. Studying pedestrian behavior is an important step in preventing pedestrian fatalities on the road. However, to investigate pedestrian behavior, several factors need to be considered. First and foremost, the safety of the participants must be assured. Second, the study environment needs to be controlled to prevent confounding variables and allow for repeated trials. Finally, the costs to develop and perform the study must also be minimized. To address these obstacles, we propose the implementation of a virtual reality (VR)-based simulator for studies of behavior and task performance with full motion. This simulator is composed of a Unity 5 environment, Oculus Rift VR headset, and Kinect or motion capture based position tracking. In this paper, we will discuss the development of the simulator, limitations, and future work.

Keywords Human factors · Simulator · Virtual reality · Motion capture

R. Sween (✉) · D.W. Carruth · D. Waddell
Center for Advanced Vehicular Systems, 200 Research Boulevard,
Starkville, MS 39759, USA
e-mail: rsween@cavs.msstate.edu

D.W. Carruth
e-mail: dwc2@cavs.msstate.edu

D. Waddell
e-mail: daniel@cavs.msstate.edu

S. Deb
Mississippi State University, Starkville, MS 39762, USA
e-mail: sd1278@msstate.edu

M. Furuichi
Nihon University, Tokyo, Japan
e-mail: furuichi.masakazu@nihon-u.ac.jp

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_16

1 Introduction

Crossing a road seems like such a simple task. However, for certain populations, specifically older and younger pedestrians [1–3], this task contains higher level of risk than for the general population. In the United States, 14 % of all traffic fatalities were pedestrians in 2013 [4]. In Japan, 38.4 % of 2015 traffic fatalities were pedestrians [5]. These numbers indicate the importance of research to understand pedestrian behaviors that may contribute to injury or death.

Previous work in studying pedestrian behavior consists of pen-and-paper studies [5], questionnaires [6], simulators derived from driving simulators [1, 3], or 3D environments navigated by joystick, like a video game [2]. Using these simulators, researchers have been able to investigate specific research questions about pedestrian behavior. However, these simulators have problems with validity—participants may not have felt as though they were actually in the environment, and so may have acted without any real feeling of perceived danger.

For example, in Zito et al. [1], participants viewed a crosswalk on three screens that provided a 180° wide by 40° tall field of view in front of the participant where participants viewed traffic flow and waited for an appropriate gap in traffic to begin the crossing. Participants signaled their intent by taking a single step forward, breaking an infrared beam in front of them. However, there were no audio cues from the environment, and participants could plainly see that they were standing in a lab and not on a street corner. This could bias the decisions they made without any real or implied risk.

Charron et al. [2] with their RESPECT simulator developed a 3D virtual world complete with sound and visual cues to study children's behavior when crossing a street. However, participants used a joystick to move through and interact with the environment. As participant's ages ranged from 9 to 12 years old, it would be reasonable to assume at least some of their behavior can be attributed to the video game-like environment and controls, where there is no real-world penalty for risky behavior.

Researchers at Mississippi State University and Nihon University believe that creating an immersive virtual reality simulator capable of supporting free movement over space, including the space required to simulate crossing a street, may address the limitations of previous simulation environments: a lack of immersion and a disconnect between experimental input and real-world action. By incorporating a 3D environment with a virtual reality headset and a room-scale motion tracking system, users are able to explore a virtual environment via a one-to-one correlation to their motions in the real world. While virtual environments combining head-mounted displays and motion tracking for research applications have been described previously (e.g., [7]), the low-cost low-latency high-resolution headsets and motion tracking systems available today offer new possibilities for observing behavior.

One potential pitfall in the use of virtual reality is that of "simulator sickness", defined by Russell et al. as "physiological and perceptual responses to discrepancies among sensory information circuits most commonly involving vision, balance, and movement." [8] This response most commonly manifests in participants as feelings of discomfort, nausea, and/or light-headedness, and most commonly occurs when the user or simulation displays movement to the user without the users' other senses reporting the same feelings of movement.

2 Design

This paper describes the design, implementation, and initial impressions of a virtual reality-based simulator that supports free movement. Based on our review of previous studies, we have identified a few requirements for such a simulator. First, the simulator needs to be immersive. For a simulator to be immersive, it should provide as many accurate sensory inputs as possible, including visual, audible, kinesthetic/proprioceptic, and vestibular sensations. Ideally, the simulator will be immersive enough that the user will forget they are in a simulator at all and react to risk and events in the environment as they would in the real world. Second, the simulator should be easy to use and easy to learn how to use. We want the interactions between the user and the simulator to be intuitive and natural so that the data we record are most similar to how a user would act in the real world. This also reduces the time needed to familiarize the user with the simulator and increases the time available to collect data. The simulator should also be comfortable and safe for the user. This not only means physically comfortable to use, but also means the simulator should not make the user sick or give the user a headache. Finally, the simulator must perform other basic tasks of a research tool. This includes support for experiment designs including the ability to perform repeated, controlled trials and record relevant data points.

The next sections address how we meet the requirements defined above. Our simulation platform is composed of three main parts: a 3D virtual environment, a virtual reality headset and a motion tracking system. The modular design of the simulator allows the headset and motion tracking system to change and adapt to a variety of installations and use cases.

2.1 Virtual Environment

The virtual environment is built in Unity 5. Unity provides built-in virtual reality support and a large developer support community. Unity provides the visual backbone for the entire simulator. As Unity is used typically for game development, it already supports a number of features needed for the simulator, including fast dynamic lighting, three-dimensional sound, basic physics engine, and a robust scripting system that allows for the development of any other features needed not already in Unity.

In order to increase the sense of immersion that participants feel while in the simulator, several audio and visual cues were added to the Unity environments. First, both ambient and specific audio cues were included. Ambient audio cues include cricket and bird noises in the highway scene, or office and building noises in the hallway scene. An example of a specific audio cue is the sound of a car as it passes the user. These specific audio cues make use of Unity's 3D audio system to handle automatically scaling the volume of the sound as the audio source moves closer to or away from the user, including a Doppler effect if the object passes the user.

In the real world, very few environments are motionless. Fans spin and affect objects in an office; cars, bicycles, and trucks drive around outdoor environments; and often there are other people moving around and performing tasks. Our current environments include several examples of this type of motion. For example, in the highway scene, cars drive along the highway and a light at the top of a cell phone tower blinks. A pedestrian standing beside a building randomly cycles through a series of animations, such as talking on a cell phone, looking around, and waving.

While sound and motion increase the feeling of immersion, a world that does not react to a user's presence can still feel sterile. To correct for this, responsive elements were added using Unity's build in trigger colliders. Trigger colliders detect when two collision meshes intersect or overlap. However, no physics calculations performed for trigger colliders, which improves performance. For example, a trigger collider can be used to start a vehicle moving to see how a user reacts to that movement.

2.2 Virtual Reality Headset

In order increase the immersion with the virtual environment, the simulator needs a way to display the environment to the user in an engaging manner. An Oculus Rift DK2 head mounted display provides the user a natural way to look around the environment, meeting the requirements of natural actions and sensory input. Sensors in the Rift allow the user to look around the virtual environment by turning their head in the real world. This improves on projected displays by allowing the user to look at any angle in front of or behind them (and up or down) and still maintain a view of the environment. Moreover, compared to screen-based virtual environments, visual navigation is as simple and natural as turning your head, versus using a controller or mouse to look around. Although the simulator only supports the Rift DK2 at this time, we are planning to add support for the consumer version of the Oculus Rift, as well as the HTC Vive as those devices become available later this year.

2.3 Motion Tracking

Once the user can look around a scene, the next logical step in development is to allow them to move around the environment. As mentioned previously, this step needed to be designed carefully to reduce the chance that the user will succumb to "simulator sickness." To allow the user the freedom to move naturally throughout the environment, we first incorporated a Microsoft Kinect 2 sensor to provide a marker-less motion tracking solution. This solution provided a simple way to obtain quickly a head position that is used to move around the virtual environment. Another benefit to the Kinect 2 sensor is that, while currently the simulator is only using head position, the sensor provides tracking data for 24 other joints that could be used in a future iteration of the simulator.

However, a major limitation of this Kinect 2 sensor is that there is a limited area that the Kinect 2 sensor can track the user's position reliably (a trapezoidal shape, roughly 11 m^2). In order to improve the size of the capture area, a room-scale motion tracking system was built using a Motion Analysis passive infrared motion capture system. Twelve motion capture cameras track a set of markers worn by the participant to obtain their head position and set the camera position in Unity, similarly to the Kinect system. This solution currently provides a rectangular, 24 m^2 capture area, roughly twice the size of the Kinect solution. We are also in the process of procuring a 12-camera Vicon motion tracking system that will also be integrated with the simulator.

One of the first challenges encountered in development was developing a way to calibrate the various coordinate systems that each main sub-system used. Aligning the Unity coordinate system with either the Kinect or Oculus devices was trivial; however ensuring that the Kinect and Oculus coordinate systems were also coordinated was a challenge. This calibration was first performed manually—the Kinect sensor was oriented 180° offset from the forward vector defined by the Oculus Rift. However, this calibration method was inaccurate and inconvenient, as it was not always practical to orient the Kinect sensor in this way. After further research, we identified a Unity method to reset the forward vector of the Oculus, which we use to align the Oculus coordinate system with the coordinate system of either the Kinect or the motion capture system.

2.4 Additional Hardware Components

In addition to the main hardware and software components described in previous sections, two pieces of additional, accessory hardware are used by the simulator. The first piece of hardware is a pair of wireless headphones. The headphones provide a dual purpose: in addition to providing audio information to the user, the headphones also provide a convenient location for mounting the passive infrared markers for the motion capture system.

While the increased capture area provided by the motion capture system certainly increases the amount of the environment that can be explored, it is not a panacea. Many tasks of interest to researchers may involve movement over spaces larger than can be easily tracked using current techniques. Any environment over a few square meters will require some alternative method for navigating great distances around the environment. To this end, the simulator supports using a gamepad to move the entire capture area around the environment. Once the capture area is resituated, the user may resume navigating the environment simply by moving around in the physical world.

2.5 Software Components

The software components of the simulator were designed to be modular, independent subsystems. This design allows experiment designers to include only components needed for a given scenario. The modular design also allows for a simple, drag-and-drop method for adding components to a project, which helps scenario designers quickly get started using the simulator.

The Motion Tracking component comprises both the Kinect and motion capture tracking solutions. The core of the tracker is a socket server that, when the scene is started, begins listening for socket connections from either a Kinect or motion capture client. These clients are responsible for streaming the head point data into the simulator. After connecting to a client, the Motion Tracking component listens for updated head position data, and updates the main camera's position accordingly.

Once users were able to walk around an entire room to explore the virtual environment, it became clear that additional steps would need to be taken to protect the users from injury in the real world. One such measure was the addition of guidelines to the virtual environment that match the confines of the real world. This serves two purposes. First, it protects the user from running into walls or other obstacles outside the designated capture area in the room. It also helps keep the player confined in the region of the room that the motion capture cameras or Kinect sensor are set up to track. It is jarring when a user reaches the edge of the specific systems capture area and attempts to go beyond where the system can reliably track the user, as, once tracking is lost, user's head movements in the real world are no longer mimicked in the virtual world. By showing the user the limits of the system in the virtual environment, it encourages them to stay within the confines of reliably tracked space.

In order to function as a research tool, the simulator needs to be able to record specific data points for later analysis. The Recording and Playback components allow the simulator to keep track of game objects and events in the simulator, record their positions and orientations for later analysis, and playback a trial to review what happened during the trial. The Recording component also tracks any inputs that the user may give from a controller or mouse and keyboard.

The Experiment Control component is responsible for handling tasks related to running an experiment, including managing conditions, participants, and data recording. The Experiment Control module acts as a state manager, which other script components can query to determine how they should behave in a given scenario. This module also keeps track of the current participant number and manages starting and stopping the Recording module as needed.

3 Discussion

As mentioned in the Motion Tracking section above, we are still looking to improve the size of the capture volume and explore new techniques for motion tracking. As mentioned previously, we are procuring a 12-camera Vicon motion tracking system that should improve the quality of our motion capture data. We are also very interested in investigating the use of the HTC Vive with our simulator. The Vive system comes with a pair of motion trackers called Lighthouse sensors. These sensors use infrared lasers in two base stations to allow the headset and included handheld controllers to be tracked across an approximately 12.5 m^2 area. While this area is only marginally bigger than the Kinect's tracked area ($\sim 11\ m^2$), it still provides a marker-less tracking solution and adds additional input mechanisms through the tracked handheld controllers. We are also interested in investigating much larger tracking areas through the use of camera-based tracking systems similar to those used by Zero Latency [9]. This type of system would allow tracking in a gymnasium-sized area or larger and would allow users to explore even larger environments.

As with any type of research simulator, such as a driving or flight simulator, questions of validity will arise. Broadly, these can be described as internal and external validity. Internal validity questions include: does the user behave the same way in the simulator as they would in the real world? External validity is concerned with whether the results from the simulator can be generalized and applied to the real world. We believe that based on the requirements defined and met, our simulator addresses these validity concerns. The simulator is highly immersive and interactions with the environment were designed to be as natural as possible. Further research will be done to quantitatively address these validity questions.

Along with the questions of validity, there is an additional limitation of the simulator. While researchers are able to design a wide variety of environments to observe human behavior and task performance in any number of contexts, these environments must be designed and implemented with the requirements in mind. Put another way, the results and experiences provided by the simulator can only be as good as the environment designer makes them. To counter this, it will be necessary to develop guidelines for achieving the necessary fidelity and validity in virtual environments to ensure that designers are aware of these limitations and requirements and can develop research scenarios that will provide valid and generalizable data.

One observation related to the current simulator is the lack of motion or simulator sickness reported by users. We attribute this to two main causes. First, the latency of the Oculus Rift DK2 display is fast enough that there is minimal lag between moving your head in the real world and seeing the updated environment displayed in the headset. In several early prototype environments that were not fully optimized for the display, as the frames-per-second (FPS) that the simulator could render dropped and latency in the headset increased, there was a jarring disconnect between users' actual and simulated motions, which did result in some reports of discomfort. The other factor we believe aids in reducing simulator sickness is the fact that user's movements in the real world are translated one-to-one into the virtual environment. Put another way, for every movement a user makes in the real world, the user's view in the virtual world updates to match their movement. By reducing the disconnect between users physical and perceived movements, we believe that the simulator provides a more comfortable experience for users in the virtual environment.

3.1 Future Work

Development work continues to improve and develop the simulator. Currently the only way for users to interact with the simulated environment is through reactive elements, elements with specific responses to limited environmental triggers. While reactive elements help to increase immersion by including dynamic elements in the environment, these elements are currently limited to detection of the user's presence, not any intent to act. The inclusion of interactive elements will allow the users to perform actions like pressing buttons at crosswalks or to open doors.

At present, the simulator only allows one user to explore the environment at a time. For coordination studies or other experiences, it would be beneficial to allow multiple users in the Simulator at the same time, whether they are co-located in the same physical space or geographically disparate. Research is currently being done to expand this capability of the Simulator.

While the simulator currently only uses head position for representing the user in the environment, there are instances where tracking and representing more of the user in the environment would be beneficial. For example, if there are multiple users in the environment, it would be helpful to be able to indicate a user's pose to other users in the simulator. Tracking and displaying users' arms would also allow for more natural interactions with the environment. In addition, whole-body tracking is a key capability to support ergonomic research applications.

4 Conclusion

Real world pedestrian research can be difficult to do in a controlled real-world environment. Not only would it be expensive and difficult to secure a large enough area, the risks involved with exposing users to real traffic would be unmanageable. Virtual reality has had the potential to allow a range of controlled and repeatable experiments that are not possible to perform in the real world. Today, emerging low-cost technologies in VR, motion tracking, and graphics software make achieving this potential easier than ever. Our current simulator provides a safe, controlled, and effectively unlimited environment for observing pedestrian behaviors. Our simulator provides more, accurate sensory inputs to increase the user's feeling that they are truly in the environment. All interactions between the user and the simulator were designed to be as natural and realistic as possible. While validation of the current simulator must be done and is likely to be dependent on the behavior(s) of interest, the free movement and enhanced immersion provided by the simulator is expected to address limitations of previous platforms for observation of human behaviors in a wide range of tasks requiring free ranging movement, including pedestrian behaviors.

Acknowledgments The authors wish to thank Elijah Davis, Joshua Church, and Jeffry Herzog for their assistance in development of the Pedestrian Simulator. This research was funded by the Center for Advanced Vehicular Systems.

References

1. Zito, G.A., Cazzoli, D., Scheffler, L., Jäger, M., Müri, R.M., Mosimann, U.P., Nyffeler, T., Mast, F.W., Nef, T.: Street crossing behavior in younger and older pedestrians: an eye- and head-tracking study. BMC Geriatrics **15**, 176 (2015)
2. Charron, C., Festoc, A., Guéguen, N.: Do child pedestrians deliberately take risks when they are in a hurry? An experimental study on a simulator. Transp. Res. Part F Psychol. Behav. **15**(6), 635–643 (2012)
3. Cavallo, V., Lobjois, R., Dommes, A., Vienne, F.: Elderly pedestrians' visual timing strategies in a simulated street-crossing situation. In: 5th International Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, pp. 499–505 (2009)
4. National Highway Traffic Safety Administration.: Traffic safety facts: pedestrians. http://www-nrd.nhtsa.dot.gov/Pubs/812124.pdf
5. Portal site of Official Statistics in Japan. http://www.e-stats.go.jp/
6. Charron, C.: Relations entre les enjeux et les comportements de traversée d'une chaussée. Etudes exploratoires chez l'enfant de 6 à 12 ans (2005)
7. Wu, T., Tian, R., Duffy, V.: Performing ergonomic analyses through virtual interactive design: validity and reliability assessment. Hum. Factors Ergonomics Manufact. Serv. Ind. **22**(3), 256–268 (2012)
8. Russel, M.E.B., Hoffman, B., Stromberg, S., Carlson, C.R.: Use of controlled diaphragmatic breathing for the management of motion sickness in a virtual reality environment. Appl. Psychphysiol. Biofeedback **39**, 269–277 (2014)
9. Zero Latency. https://www.zerolatencyvr.com

Effect of Human Reactions at Signalized Intersections on Intersection Efficiency and Safety

Songsu Son

Abstract An intersection is one of the most critical points in the traffic systems that contain numerous factors considered in the design process. Human factors in intersection designs are complex and their characteristics have a wide range of variability in terms of personality and psychology of road users. However, these human variabilities could be estimated and used as human factors on road designs and traffic operations. The locations of traffic signal at the intersections in influences driver's visibility and reaction relevant to startup loss time on traffic signal timing. When unclear visibility exists, driver's delayed reaction time will results in greater loss time and a sudden startup that might cause safety concern between pedestrian and vehicles. This study investigated the effect of driver's reaction time at varying distance between stop line and signal location on intersection efficiency and safety for a near-side traffic signal and far-side traffic signal.

Keywords Traffic signal · Intersection efficiency · Intersection safety · Near-side traffic signal · Far-side traffic signal

1 Introduction

Geometric design and traffic operation of an intersection are very important and relevant in view of traffic system considering both traffic safety and efficiency. Each country has its guideline to design an intersection and a traffic control device. Traffic signals are generally divided on two types by locations of them on the intersection, that is, near-side signal and far-side signal. Some European countries such as German and the United Kingdom adopt a near side signal for traffic safety but most countries except them adopt a far-side signal.

As safety becomes recently more important on traffic, several countries in Asia plan to change the location of traffic signal from far-side to near-side. In this case, it

S. Son (✉)
American University in Dubai, Dubai, United Arab Emirates
e-mail: sson@aud.edu

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_17

should be checked the traffic operation on the intersection, whether it makes traffic condition worse, level of service down on the intersection and causes wrong phases and other safety problems. Moreover, the application of new equipment on existing traffic system should be seriously considered how it influences on the existing traffic system. This study mainly treats what kinds of factors and how much they are influenced, and the important factors such as control delay, start-up lost time and Stop line Observance Rate (SOR) on a near-side intersection are re-evaluated to measure them as a near-side traffic signal is installed on the existing far-side traffic control system.

2 Literature Review

A proper installation of traffic signal device prevents signalized intersections from inefficient and unsafe operations of intersections on road networks. It is noteworthy that intersection geometric design guideline in each country considers various characteristics of road components such as vehicle characteristics, transportation operating systems, drivers' behavioral pattern, etc. Therefore, direct comparison of intersection design guidelines might mislead the major design concepts embedded in the design guide. However, it is worthy to research various aspects of guidelines in various countries to come up with better guidance on intersection operations. Intersection signal installation guidelines in five countries were considered in this study.

The Korea Installation and Management Manual for Traffic Control Device suggests that drivers should be able to see a traffic signal continuously while driving the intersection segment [1]. According to this manual, the signal should be installed within 10–40 m from the stop line and the minimum distance between traffic signal and stop line must be about 10 m minimum with 20° each left and right side with total angle of 40° to provide a clear visibility of traffic signals. For the parallel signal installation, spacing between two traffic signals should be greater than 2.4 m. When the signalized intersection has greater spacing between the traffic signal and stop line, the secondary signal must be installed prior to the intersection as shown in Fig. 1.

In USA, Manual on Uniform Traffic Control Devices (MUTCD) regulates that a traffic signal must be installed within 12–45 m from the stop line and within 20° each left and right side with total angle of 40° as the same as Korea [2]. The secondary traffic signal should be also placed on the sides of roadway as Fig. 2.

Guidelines for traffic signals [3] in Germany indicate that the signal location is to be installed within 6 m from the stop line. One-way road without left or right turn has the signal device for both vehicles and pedestrians on each side of roadways. When there is no left turn on the street, one or two side signals are installed as shown in Fig. 3b. The intersection with a median strip and left turn has the traffic signal installed on the median strip while a standing traffic signal is placed on the left side of roadways for the intersection without median strip as shown in Fig. 3a.

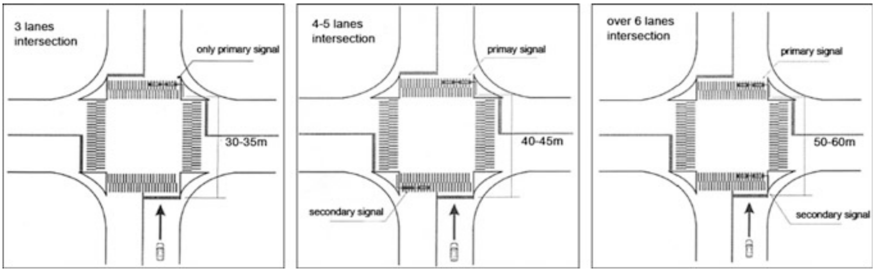


Fig. 1 Signal installation guideline in Korea

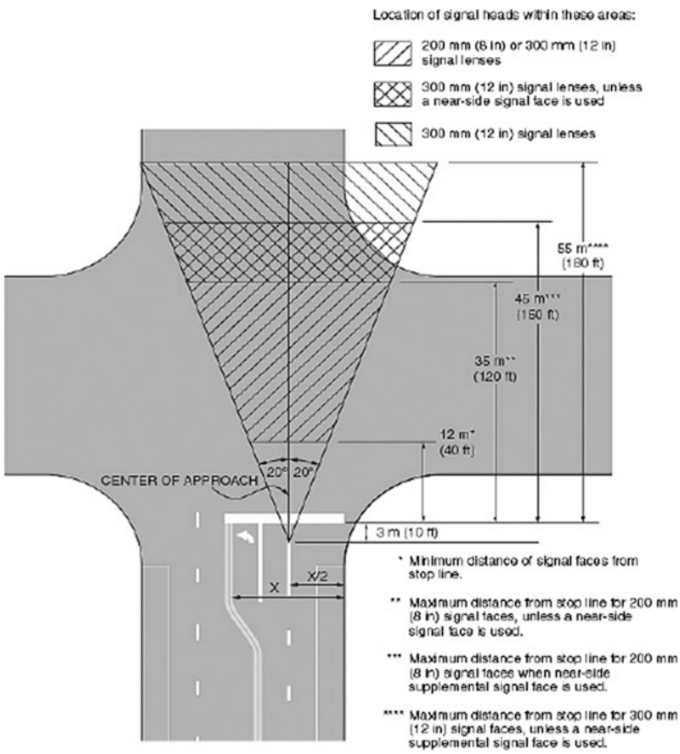


Fig. 2 Signal installation guideline in USA

In Australia, a double primary traffic signal is commonly installed on the street over two lanes with wide median strip [4]. Due to its higher cost, a forehead traffic signal is installed only when the post-mounted signal does not provide enough sight distance and driver cannot perceive a traffic signal on the curve, and the post-mounted signal is not allowed where the next traffic signal is located within 150 m (Fig. 4).

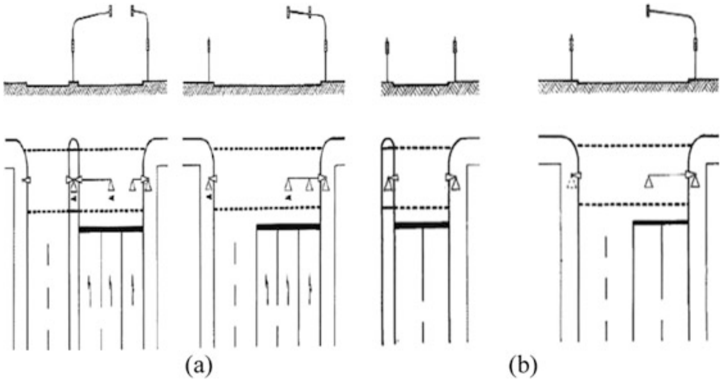


Fig. 3 Signal installation guideline in Germany; **a** with left or right turn, and **b** without turns

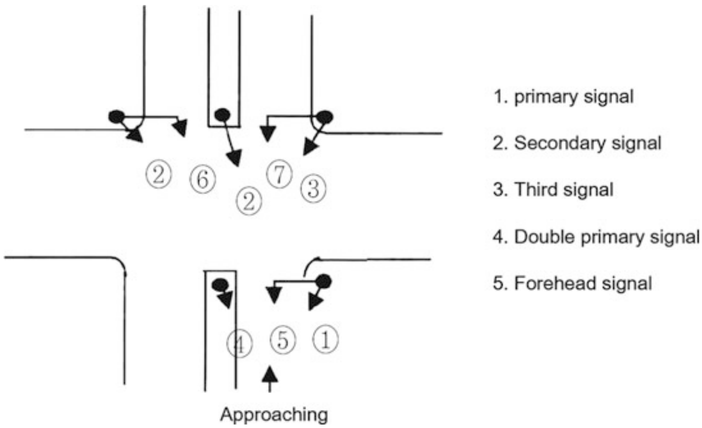


Fig. 4 Signal installation guideline in Australia

The United Kingdom regulates that a signal should be installed for driver to identify two traffic signals on different points. The roadway with a median should have a traffic signal installed on the median strip and is also not allowed to install the right side of the road as shown in Fig. 5.

No significant studies have been made on near-side traffic signal, especially near-side signal on the far-side intersection systems. Jung et al. [5] investigated safety concern on near-side intersection systems and came up with Stop line Observance Rate (SOR). This study showed that the SOR depended on various factors. Ha and Berg [6] suggested safety as level of service (LOS) on intersections instead of control delay that has been typically used as intersection LOS criteria. Lee et al. [7] showed TTC method to evaluate intersection safety. There are various measures for intersection safety as shown in Table 1. Panagiotis [8] concluded that human behavior was the most influencing parameters on intersection safety.

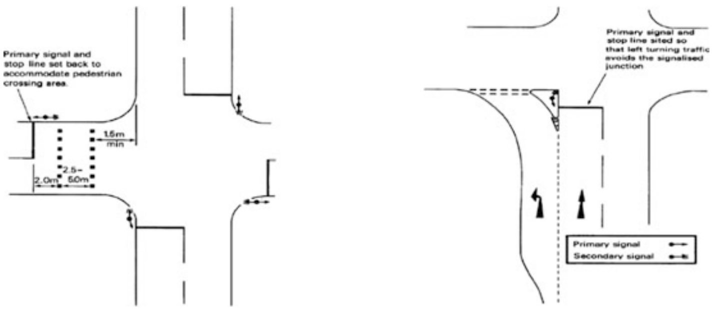


Fig. 5 Signal installation guideline in UK

Table 1 Surrogate safety conflict measures

Surrogate conflict measure	Description
Gap time (GT)	Time lapse between completion of encroachment by turning vehicle and the arrival time of crossing vehicle if they continue with same speed and path
Encroachment time (ET)	Time duration during which the turning vehicle infringes upon the right-of-way of through vehicle
Deceleration rate (DR)	Rate at which crossing vehicle must decelerate to avoid collision
Proportion of stopping distance (PSD)	Ratio of distance available to maneuver to the distance remaining to the projected location of collision
Post-encroachment time (PET)	Time lapse between end of encroachment of turning vehicle and the time that the through vehicle actually arrives at the potential point of collision
Initially attempted post-encroachment time (IAPT)	Time lapse between commencement of encroachment by turning vehicle plus the expected time for the through vehicle to reach the point of collision and the completion time of encroachment by turning vehicle
Time to collision (TTC)	Expected time for two vehicles to collide if they remain at their present speed and on the same path
Stop line observance rate (SOR)	Rate which a driver stops the vehicle before a stop line

3 Methodologies

An intersection LOS is evaluated by control delay according to Highway Capacity Manual (HCM) 2010 [9] and this study indicates the intersection efficiency based on the control delay. This study set limits on the four-leg intersection and assumed several factors properly according to HCM 2010 to evaluate control delay on various conditions including changes on approach speed, startup delay and location of stop line. Regarding intersection safety, the SOR on intersections shown in previous study was selected as indication of intersection safety suggested by Jung et al. [5]. Once intersection safety and efficiency parameters were defined, the effect

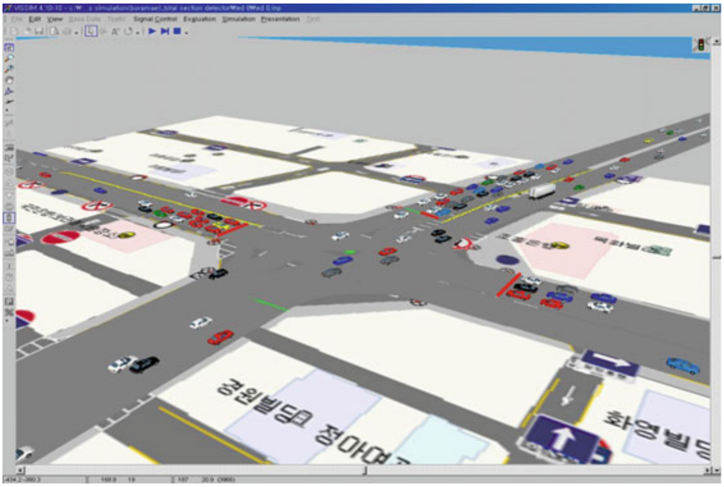


Fig. 6 VISSIM microsimulation

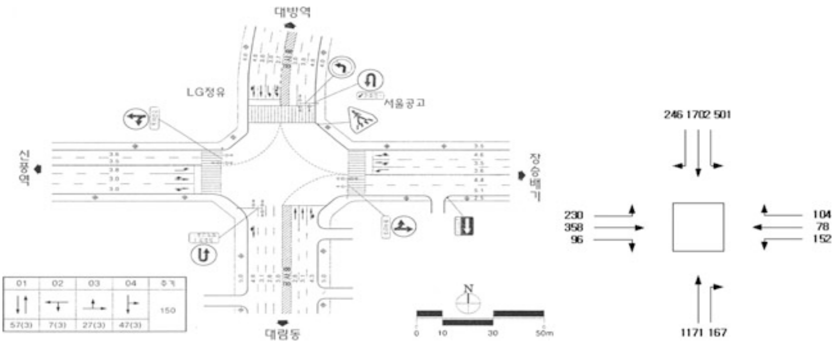


Fig. 7 Geometric and traffic conditions of simulated intersection

of distance between stop line and traffic signal under various conditions were evaluated using microsimulation software (VISSIM) (Fig. 6).

To avoid any uncontrollable situation, the LOS of E and F were excluded on the simulations. The BORAMAE Station in Korea was evaluated at varying distances between stop line and traffic signal as shown in Fig. 7.

4 Analysis and Discussion

Control delays were obtained under varying conditions as indication of intersection efficiency on nearside intersections. Several critical factors influencing control delay have been investigated including dilemma zone, and startup lost time. Yellow

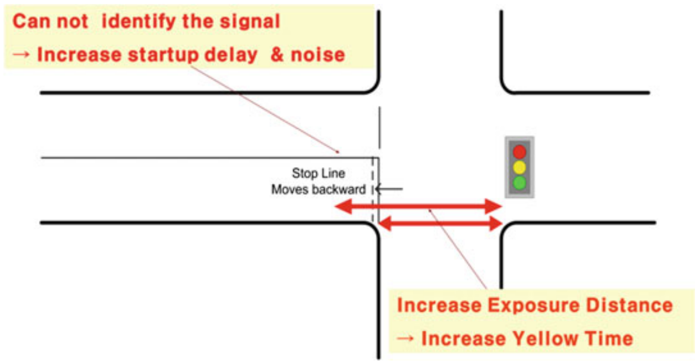


Fig. 8 Startup delay and yellow time on nearside intersection

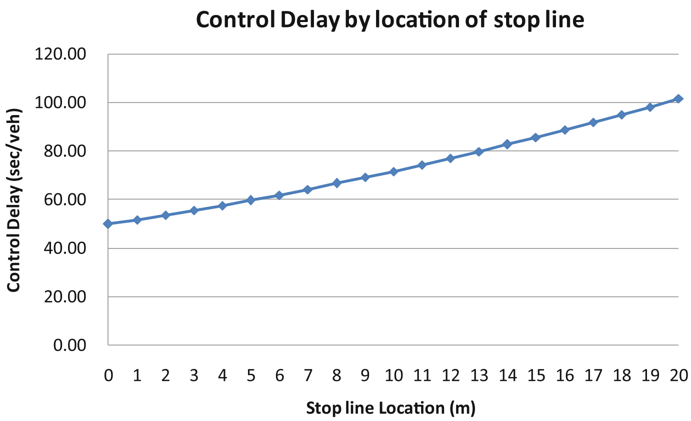
time is a converting time between green and red signal to provide buffering time to driver for safe crossing of intersections. When yellow time is too long, the driver is prone to consider it as a part of green time. Contrarily short yellow time makes a dilemma zone that makes driver neither crossing intersection nor stopping prior to the intersection properly. When near-side signal systems are applied on far-side traffic intersection systems, a stop line is required to move backward to ensure minimum spacing. The yellow time should be adjusted for extended spacing between stop line and traffic signal as shown in Fig. 8.

In this case, the control delay increases and the intersection efficiency become worse due to extended space resulting in longer yellow time or shorter effective green time per hours as shown in Fig. 9a. Startup lost time is also critical because the first several vehicles required longer time to identify the signal change and to pass the stop line than farside signal systems due to unclear visibility with greater peripheral angles to the signal device. The appropriate perceiving space between drivers and traffic signal device is typically 10 m for six-meter signal in height. Drivers within 10 m from nearside traffic signal are not able to identify the traffic signal and decrease effective green time due to extended startup lost time, which also deteriorates the intersection efficiency as shown in Fig. 9b.

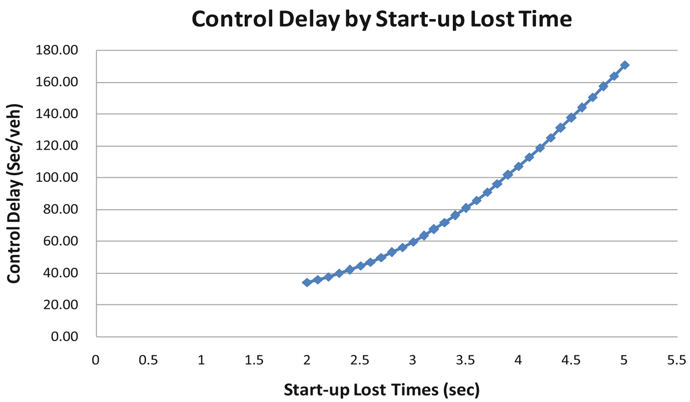
The microsimulation and calculation results for the BORAMAE Station in Korea show almost identical increase of control delay with the increase of spacing between stop line and traffic signal installation location as shown in Fig. 10 [10].

An intersection safety was also evaluated using Stop line Observance Rate (SOR) concept developed by Jung et al. [5]. These studies showed that around 30 % of traffic accidents occur at the intersections and nearside signal systems increased the SOR, which showed that the shorter spacing provided the higher SOR as shown in Table 2.

Based on analysis for both efficiency and safety, the proper spacing between stop line and near-side traffic signal to accommodate both intersection efficiency and safety have been suggested as 10 m of spacing between a stop line and traffic signal location on near-side traffic signal as shown in Fig. 11.



(a)



(b)

Fig. 9 Control delay by **a** Spacing between stop line and signal, and **b** startup lost time

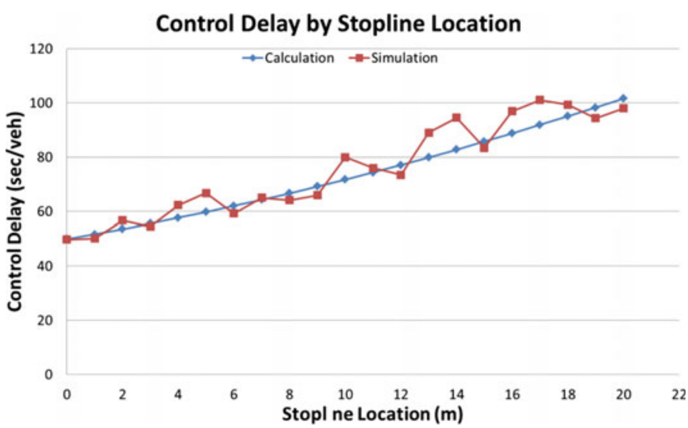


Fig. 10 Control delay by calculation and simulation

Table 2 Efficiency and safety rate at varying spacing

Stop line location (m)	Efficiency (%)	Safety (%)
1	3.60	60.45
2	7.37	57.94
3	11.30	55.47
4	15.40	53.05
5	19.67	50.67
6	24.12	48.33
7	28.73	46.04
8	33.52	43.80
9	38.48	41.60
10	43.61	39.44
11	48.91	37.33
12	54.36	35.26
13	59.98	33.24
14	65.75	31.26
15	71.67	29.33
16	77.73	27.44
17	83.93	25.59
18	90.27	23.79
19	96.74	22.04

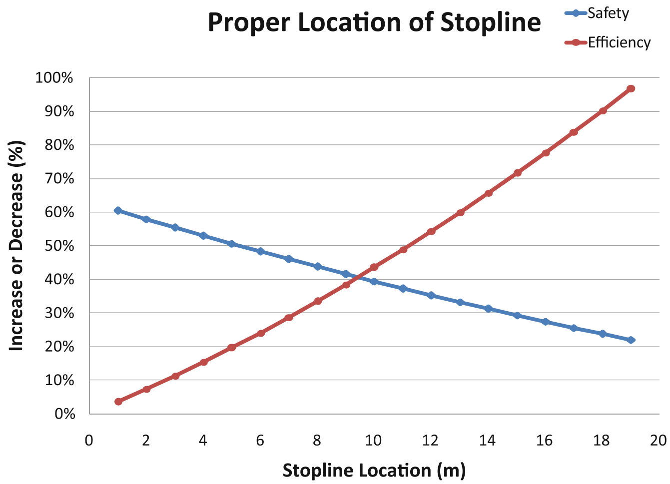


Fig. 11 Proper spacing between stop line and traffic signal location

5 Conclusions

An intersection geometric design is one of the most challenging parts to accommodate efficiency and safety concerns of road networks. As mentioned earlier, many countries put their efforts to improve intersection design guideline to ensure efficient and safe operation of signalized intersection considering all aspects of roadway components. As part of these efforts, near-side and far-side intersection signal systems have been introduced differently in each country. However, improper installation of each traffic signal systems without modification of corresponding intersection component such as stop line location or signal timings could result in worse intersection operations.

This study investigated the effect of near-side traffic signal on intersection control delay and safety, and suggested proper spacing between stop line and traffic signal to optimize the intersection operations for both intersection efficiency and safety. Especially, when the near-side signal is installed to provide better SOR on the far-side signal systems, it dramatically increases control delay because of extended yellow time and startup lost time, which might results in significant congestion after installation of near-side traffic signals. Instead of additional lane at higher cost to increase capacity of intersection to accommodate this congestions, modification of stop line location would reduce startup lost time and increase safety with optimum spacing between stop line and signal location. In addition to the modification of spacing, a secondary traffic signal is also recommended on the sides of the roadway to provide additional traffic signal like European countries. However, it still requires drivers to keep their eyes on the sides until the signal has been changed and take a bit longer time to turn their heads to the traffic direction as part of startup lost time.

This study suggests further studies on assumptions and restrictions made on the study. More realistic startup lost time at varying spacing between stop line and signal location would better estimate optimum spacing on near-side intersection. Moreover, various safety parameters need to be considered as indication of intersection safety in addition to the SOR.

References

1. Korea National Police Agency.: Installation and management manual for traffic control device (2005)
2. Federal Highway Administration.: Manual on uniform traffic control devices (2009)
3. Road and Transportation Research Association.: Guidelines for traffic signal RiLSA (2003)
4. National Association of Australian State Road Authorities.: Guide to Traffic Engineering Practice (1988)
5. Jung, G.B., Kim W.K., Kho, I.S., Lee, T.K.: The study on the correlation between traffic location and stop line observance. J. Korean Soc. Transp. **3** (2005)
6. Ha, T.J., Berg, W.D.: Development of safety-based level-of-service criteria for isolated-signalized intersections. Transp. Res. Board. 98–104 (1995)

7. Lee, J.H., Kang, K.D., Jung, J.H.: Study of intersection safety consolidating by adjustment of traffic light place. J. Korean Soc. Transp. **3** (2005)
8. Panagiotis, P.: Driver behaviour, dilemma zone and safety effects at urban signalised intersections in Greece. Accid. Anal. Prev. **39**, 147–158 (2007)
9. Transportation Research Board.: Highway capacity manual (2010)
10. Son, S.S., Chon, K.S.: Analysis of changes on the intersection efficiency on introducing a near-side traffic signal. In: Proceedings of the Eastern Asia Society for Transportation Studies, vol. 6 (2007)

An Analysis of the Start-up Delay and Safety for Signalized Intersections: Impact of Left-Turn Phasing Sequences

Mohamed Shawky, Abdulla Al-Ghafli and Hussain Al-Harthi

Abstract This paper aims to investigate the start-up delay at signalized intersections in Abu Dhabi (AD) city, UAE. The impact of some external factors that may affect the start-up delay in examined including; left turn phasing sequences (split/lead/lag), movement turning (through/left), intersection location (CBD/non-CBD) and day time (peak/off-peak). The paper also addressed the impact of illustrating lead/lag phasing system on the performance of traffic and safety. A significant number of observations were obtained by using automated license plate recognition cameras at 66 different intersection approaches. The results show that the estimated mean value of the start-up delay is 2.201 s with a standard deviation of 1.823 s. In conclusion, the statistical tests show significant difference in start-up delay the have between observed for through, left, at CDB and non-CDB area and between split and lead/lag phasing. However, no significant difference between peak and off-peak periods and between split and lead phasing are recognized. In addition, the lead/lag phasing improved the traffic performance by reducing the total delay at intersections but has a negative impact on the safety.

Keywords Start-up delay · Saturation flow rate · Lead-lag phasing · Road safety · Abu Dhabi city

1 Introduction

Capacity and delay are two of the commonly used measures of effectiveness (MOEs) in the evaluation of signalized intersections [1]. The startup delay is a part of the total delay time that occurs due to the implementation of the traffic signal

M. Shawky (✉) · A. Al-Ghafli · H. Al-Harthi
Traffic and Patrol Directorate, Abu Dhabi Police, Abu Dhabi, United Arab Emirates
e-mail: m_shawky132@hotmail.com

M. Shawky
Faculty of Engineering, Ain Shams University, Cairo, Egypt

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484, DOI 10.1007/978-3-319-41682-3_18

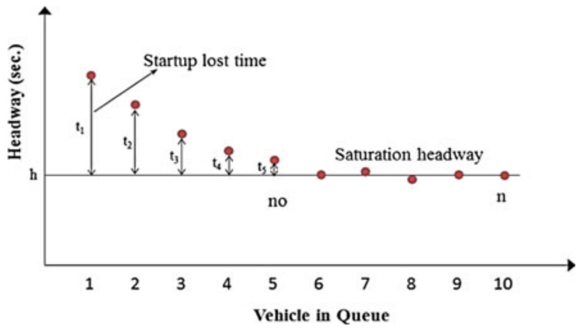


Fig. 1 Headway time for the departure vehicles standing in a queue

control devices. At the beginning of each green time of the traffic signal phase, the first few number of queued vehicles experience start-up time losses that is made up of the response time of the drivers (perception and reaction time) to the change in signal indication along with the vehicle acceleration time to free-flow speed [2]. In this case, the headway time of the departure queued vehicles can be illustrated as shown in Fig. 1. It shows that after a certain number of vehicles (n_o) (usually fall between the third and sixth vehicle) the headway time reaches its minimum value (h) which calls the saturated flow condition [3].

Based on the Highway Capacity Manual (HCM-2000), the start-up delay takes place for the first four vehicles in a standing queue (i.e., $n_o = 4$) and from the fifth queued vehicle the saturation headway can be estimated.

Accordingly, the start-up delay can be calculated as follows:

$$startup\ delay(d) = \sum_{n=1}^{n_o} t_n \tag{1}$$

where: t = headway time—h.

Many factors that affect the value of the start-up delay can be found in the literature as presented in the following section based on data collected from different locations around the world. However, the value of the start-up delay and the contributing factors affecting its value has not been investigated in Abu Dhabi city (AD), the capital of the UAE. On the other hand, lead/lag signal phasing was applied in AD intersection by year 2010. The impact of such phasing system in the operating, safety and start-up delay not intensively discussed before. Thus, the current study aims to find answers for a number of the following questions:

- What is the average value of the start-up delay in AD city signalized intersection?
- Is there any significant difference in the start-up delay in terms of left and through movements, lead/lag/split phasing, off-peak/on-peak hour periods, and CBD or non CBD areas?
- What is the impact of applying the lead/lag phasing systems on the capacity and safety performance of the intersections?

2 Literature Review

The startup delay at signalized intersections has been investigated in a significant number of prior studies. The HCM-2000 [3] mentioned that the typical observed value of the startup delay ranges from 1.0 to 2.0 s. However, in literature the estimated values of the start-up delay have a wide range between 0.75 and 3.04 s. [3, 4]. Table 1 summarizes the findings of the estimated values of the start-up delay from prior studies.

Several factors that affect the value and distribution of the startup delay were also investigated in the prier studies. These factors include the turning movements (trough, left and U-turn), queue length, intersection geometry and location, time of the day, weather condition, visibility of traffic light, phasing timing and sequence, etc.

Regarding the turning movements, Honglongli and Prevedourod [13] found that the startup delay of the through movement is larger than that of the protected left-turn movement. In addition, high standard deviation values were observed for both movements and reflect a big variation of the startup delay among drivers. However, other studies (i.e., [15, 16]) found no significant differences in startup delay between through and left-turn movements. Also, no significant differences between peak and off-peak hours in terms of the start-up delay.

Table 1 Examples of the observed startup delays

Study (source)	Date	Country (location)	Queued vehicle number (n_o)	Average start-up delay (s)
Leong [5]	1964	Australia (Sydney)	4	1.12
Gerlough and Wagner [6]	1967	USA (Los Angeles)	5	2.05
Carstens [7]	1971	USA (Iowa)	4	0.75
Agent and Crabtree [8]	1983	USA (Lexington, Kentucky)	4	1.40
Lee and Chen [9]	1986	USA (Kansas)	5	3.04
Roess et al. [10]	1989	USA (Texas)	4	1.31
Efstathiadis and Machemehl [11]	1995	USA (Texas)	4	1.34
Jacobs [12]	1998	South Africa (Stellenbosch)	5	1.43
Al-Ghamdi [4]	1999	Saudi Arabia (Riyadh)	4	2.99
Honglongli and Prevedourod [13]	2002	USA (Honolulu)	4	1.76
David et al. [14]	2013	USA (various)	5	2.16

The impact of the geometric parameters was also investigated by Bonneson [17]. It was found that the left-turn radii affect the headway of the queued vehicles. The larger radii of the left-turn paths resulted lower headways. In addition, it is indicated that queue length per cycle and lane volume has a negative effect on the headway of the first twelve vehicles. This finding implies that the startup delay of long queue is smaller due to the higher traffic pressure of the long queues. Al-Ghamdi [4] showed that the startup delay of two lane approach is significantly higher than that in the three lane approaches.

Honglongli and Prevedourod [13] showed that a weak negative correlation between the startup delay and queue length and the ANOVA tests indicated that the startup delay is not sensitive to the queue length. Long [16] found that no significant impact of the queue length, number of lanes, intersection location and peak period on the observed average startup delay. In addition, no significant differences were found in average startup delay between queues that contain trucks and queues with passenger vehicles also between different sites with level approach and sites on a 5 % upgrade. Regarding the weather condition, Sun et al. [18] observed that the startup delay increased by 21–31 % in rainy weather compared by clear weather and no significant differences was found between light-medium rainy weather and clear weather.

The studies that addressed the impact of the left-turn phasing sequence on the startup delay are very few compared by the other investigated factors. Most of these studies concentrated on the impact of the permissive-and-protected left turn (PPLT). Noyce et al. [19] and NCHRP report [20] found that no differences in the startup delay were found due to the type of PPLT signal display. However, Brehmer [2] found that the average startup delay was significantly influenced by the PPLT signal phasing. On the other hand, Chris Sheffer et al. [21] compared startup delay between lead and lag protected-only phasing. It was found that that both the mean start-up lost time and fourth vehicle crossing time were significantly lower for lag left turns. In addition, Upchurch and Wright [22] evaluated delays at one intersection for three different lead and lag phasing. It was found that left-turn delay for protected/permitted lead phasing is lower than for protected/permitted lag phasing. However, the study did not consider signal progression adjustments that may have affected the platooning of upstream traffic.

In terms of intersection capacity there prior studies proved the positive impact of installing lead-lag phasing on the capacity. Grover [23] documented a 30–50 % reduction in overall vehicle delay (means 30–50 % increase in capacity). However, other studies presented the negative impact of such left-turn phasing system on the safety performance of the signalized intersections. Randy et al. [24] stated that in the one-year period before installation of lead-lag left-turning 44 accidents occurred, whereas 78 occurred in the year after in Kentucky intersections and about 69 % of these accidents occurred in the first 6 months.

3 Start-up Delay Model Development

Figure 2 illustrate the concept of the model development to find the startup delay (*d*) value of each traffic cycle at a signalized intersection approach. From this figure the startup delay can be calculated form the following equation:

$$d = t_n - \left[(t_n - t_{n_o}) \left(\frac{n_o}{n - n_o} \right) \right] \tag{2}$$

where:

- d startup delay value
- n_o number of vehicle experience with startup delay
- n number of queued vehicle taken into consideration in the analysis
- t_n the elapsed time from the beginning of the green light until the vehicle "n" standing in the queue cross the reference line.

In this study, "n_o" and "n" are taken 4 and 10, respectively. It means that the startup delay will be considered from the first four vehicles and the saturation flow starts from the fifth vehicle to the tenth queued vehicle. Accordingly, Eq. (2) will be as follows:

$$d = t_n - \left[(t_n - t_{n_o}) \left(\frac{4}{6} \right) \right] \tag{3}$$

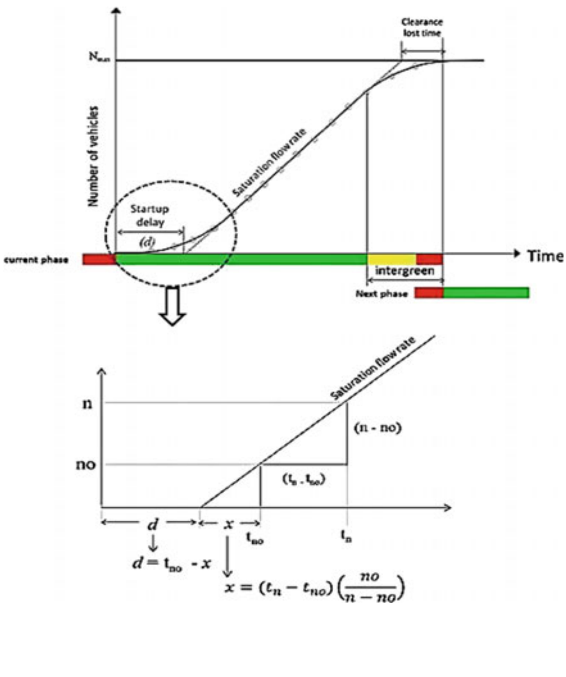


Fig. 2 Concept of developing the startup delay model

4 Case Study Selection and Data Collection

By year 2010, AD DoT applied the lead-lag left-turn phasing at about 38 signalized intersections in order to increase its capacity. At these intersections two approaches are operating as lead/lag left turn phasing and the other approaches are working as split phasing. The leading phase takes place when the left-turn starts at the beginning with though phase. The lag phase takes place when the left-turn at the end of the trough phase. Split phasing takes place when the left-turn and trough movements are start and end with each other.

In this study, about 66 approaches located at 36 different intersections were selected to be taken as case study. These intersections was selected to cover different geometric and operational parameters that may affect the start-up delay value and that will be involved in the analysis process such as; (1) intersection location (CBD/non-CBD area), (2) phasing type (lead/lag/split phasing), and (3) number of through lanes, number of left lanes. About 12,517 traffic signal cycles were involved in the analysis, 6202 traffic cycle during the peak periods and 6310 cycles during off-peak periods. These two periods are defined based on the day time as shown in Table 2.

It is worth mentioning that the selected approaches have been selected to be similar in some parameters such as; 0 % gradient, 0 % heavy vehicles and some lane width. Therefore the impact of these factors in the estimated value of the start-up delay not included in the collected data. Table 3 shows the studied number of intersection approaches under each category and the corresponding number of traffic cycles.

The majority of the collected data in prior studies used the manual technique from videotapes and stopwatch. However, in this study, a new technique of data collection was implied. This technique employed the image processing that taken from the red light violation cameras (TVR cameras) which automatically record the license plan for each vehicle cross the stop line during red light. In our case, the system was adapted to record the time when a vehicle crossed the stop line from the start of green time for each lane separately. So headway time can be accurately calculated. About 125,170 of headway time were obtained by this method.

On the other hand, the traditional video recording and stop watch technique was used at about 16 different approaches was applied to determine headway times. The collected data by this will be used to check and accuracy of the TVR cameras recording times and then to calibrate the time value of the first vehicle in the queue.

Table 2 Day time classification for the headway time data

Day time	Peak period		Off-peak period	
	From	To	From	To
Morning	6:00 a.m.	9:00 a.m.	9:00 a.m.	11:00 a.m.
Evening	2:00 p.m.	4:00 p.m.	4:00 p.m.	16:00 p.m.

Table 3 Number of studies intersection approaches for each category

No.	Intersection approach category		Number of studies approaches	No. of studies traffic cycle
1	Intersection location	CBD	16	2660
		Non-CBD	50	9858
2	No. of through lanes	2	5	538
		3	41	8015
		4	20	3966
3	No. of left lanes	0	6	1027
		1	47	9114
		2	13	2378
4	Traffic signal phasing type	Split	37	6243
		Lead	16	2202
		Lag	13	6243

5 Data Analysis

5.1 *Headway Time Analysis*

The mean value of the observed headway time of the queued vehicle based on the position of the vehicle in the queue is shown in Fig. 3 from the data obtained from TVR camera. Table 4 shows summary of the observed headway time statistics.

Table 5 shows the mean value of the observed headway by the two methods of the data collection. It shows that no significant different between the headway time for the second vehicle to the 10th vehicle in the queue. The observed difference for the first queued vehicle can be justified due to the taken time for the vehicle to fully

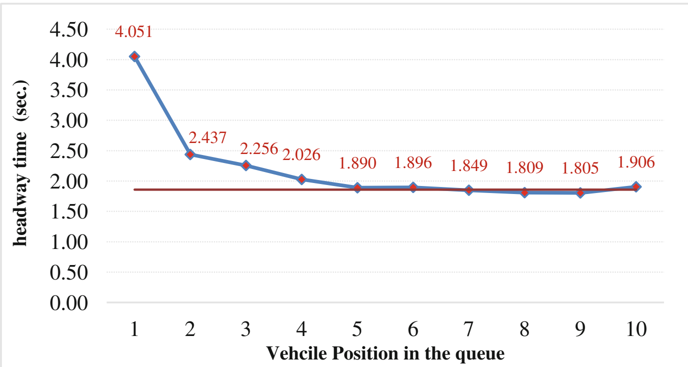


Fig. 3 Mean value of the observed headway time for queued vehicles

Table 4 Statistical parameters of headway time

Vehicle position in the queue	Headway mean (s)	Standard deviation (s)	Maximum	Minimum	Mode	Median	Skewness
1	4.051	0.925	5.367	0.584	3.852	4.232	−0.0010
2	2.437	0.838	5.158	0.758	2.106	2.331	0.0004
3	2.256	0.927	5.367	0.498	1.903	2.043	0.0010
4	2.026	0.741	5.749	0.487	1.607	1.903	0.0006
5	1.890	0.730	4.384	0.524	1.513	1.758	0.0007
6	1.896	0.794	4.539	0.658	1.451	1.717	0.0010
7	1.849	0.759	4.316	0.421	1.591	1.688	0.0009
8	1.809	0.755	4.259	0.398	1.357	1.653	0.0009
9	1.805	0.735	4.025	0.342	1.451	1.654	0.0008
10	1.906	0.897	5.101	0.450	1.513	1.681	0.0011

Table 5 The mean value of the observed headway time based on the two different data collection methods

Data collection method	Vehicle position in the queue (s)									
	1	2	3	4	5	6	7	8	9	10
Video recording	3.038	2.465	2.213	1.945	1.916	1.901	1.858	1.835	1.782	1.889
TVR camera	4.051	2.437	2.256	2.026	1.890	1.896	1.849	1.809	1.805	1.906

cross the stop line and the TVR camera recognize and record the plat number. Thus the difference value of 1.103 s. (i.e. = 4.051–3.038) can be taken as an adjustment value in the calculation on the start-up delay.

5.2 *Estimating the Overall Value of Start-up Delay and Saturation Flow Rate*

Table 6 shows the statistical parameters of the estimated value of the start-up delay. The estimated mean value of the start-up delay was adjusted due to the usage of TVR camera as discussed before. The sample size shown in the table represents the number of traffic signal cycles that taken at each intersection approach category. In general, the estimated start-up delay vale and its standard deviation for all approach types are close except the case of lag phase approach has low mean start-up delay value and higher standard deviation. However, the statistical test shown be applied to identify whether there are a significant difference or not which will be discussed in the next section.

Table 6 Start-up delay statistical parameters for different approach categories of signalized intersections

Approach category		Sample size	Mean (s)	Std. dev. (s)	Max.	Min.	Mode	Median
All intersection approaches		12,916	2.201	1.823	8.705	0.003	3.308	3.199
Movement turns	Through	8907	2.232	1.817	8.705	0.007	4.468	3.235
	left	4004	2.133	1.834	8.689	0.003	4.676	3.122
Intersection location	CBD	2863	2.295	1.802	8.483	0.007	4.681	3.300
	non-CBD	10,048	2.175	1.828	8.705	0.003	3.308	3.172
Phasing sequences	Split	6497	2.265	1.806	8.705	0.007	3.308	3.264
	Lead	4122	2.299	1.780	8.689	0.008	4.108	3.311
	Lag	2292	1.844	1.902	8.417	0.003	1.405	2.774
Day time period	Peak	6479	2.241	1.832	8.705	0.007	6.916	3.250
	Off-peak	6200	2.209	1.812	8.483	0.003	4.628	3.222

For more details in the impact of phasing sequence, the start-up delay values of left turn only was extracted and the statistical parameters for different phasing sequence s are estimated as shown in Table 7. It shows that the start-up delay value of let turn movement in case of lead phasing is significantly higher that both split and lag and lag phasing has the lower value of start-up delay. This result can be interpreted as the drivers in the left turn lane in case of lag phasing are expecting the on-set of the green light because the green of the thought movement has been already turned on. So they are ready to move or sometimes they anticipate the on-set of green.

From the observed headway, the saturation flow rate (SFR) could be also estimated. The SFR is very important parameter in the evaluation of the traffic performance at signalized intersections and can be calculated from the following equation:

$$SFR = \frac{3600}{\bar{h}} \tag{4}$$

Table 7 Start-up delay statistical parameters for left turn movement with different phasing sequences

Left turn phasing	Sample size	Mean (s)	Std. dev. (s)	Max.	Min.	Mode	Median
Left lane with split phasing	1494	2.153	1.843	8.407	0.030	4.208	3.096
Left lane with lead phasing	1524	2.325	1.772	8.689	0.012	5.492	3.363
Left lane with lag phasing	988	1.807	1.870	8.247	0.003	5.259	2.724

Table 8 Statistical parameters of the saturation flow rate

Sample size	Mean (veh/h/lane)	Std. dev. (veh/h/lane)	Max.	Min.	Mode	Median
11,786	1927	276	2494	1210	1981	1923

$\bar{h}$ is the mean value of the observed headway for queued vehicles started by the 5th vehicle in the queue to the 10th vehicle in the queue in our case. Table 8 shows the summary of the estimated parameters of the SFR. More analysis in SFR doesn't take because it is out of current study objective.

5.3 *Investigating the Significant Factors Affecting the Start-up Delay Value*

The statistical t-test was employed to test the significant differences between two pairs of the independent variables; for example between the trough and left movements, CBD and non-CBD locations, Lead and lag phasing, etc. The statistical software program SPSS was used in this analysis. Table 9 shows summary of the output results of the statistical tests. It shows that there are a significant differences between through and left turn movements, CBD and non-CBD, split and lag, lead and lag phasing at significant level of 95 %. However, there are not significant differences between peak and off-peak periods and between lead and split phasing. This result can be interpreted as in the both cases of split and lead phasing the green light starts for the left and through movements at the same time when the start-up delay takes place.

Table 9 Statistical results of the comparison among variables

Compared variables	Levene's test for equality of variances		t-test for equality of means				
	F	Sig.	t	df	Sig. *P*-value	Mean difference	Std. error difference
CBD and non-CBD	2.494	0.114	−3.131	12,909	0.002	−0.120859	0.038601
Split and lead	1.896	0.169	−1.008	10,617	0.313	−0.036062	0.035761
Split and lag	21.942	0.000	9.463	8787	0.000	0.421006	0.044492
Lead and lag	30.259	0.000	9.560	6412	0.000	0.454524	0.047547
Peak and off-peak	0.585	0.444	−0.999	12,677	0.318	−0.032336	0.032377
Left lane with split and left lane with lead	4.007	0.045	−2.606	3015	0.009	−0.171510	0.065818
Left lane with split and left lane with lag	1.542	0.214	4.550	2478	0.000	0.346075	0.076053
Left lane with lead and left lane with lag	9.387	0.002	6.994	2509	0.000	0.517585	0.074005

Table 10 Traffic performance analysis results

Intersection no.	Overall average delay (s/veh.)					
	AM peak			PM peak		
	Before applying lead/lag	After applying lead/lag	% of change	Before applying lead/lag	After applying lead/lag	% of change
1	82.5	52.6	36	103	40.9	60
2	86.5	60.4	30	42.2	39.2	7
3	89.7	74.2	17	59.6	31.3	47
4	105.9	52.4	51	61.3	27.7	55
5	97	22.9	76	41.6	19.2	54

5.4 Impact of Lead/Lag Signal Phasing on the Traffic Performant

Regarding the impact of applying the lead-lag left turn signal phasing on the total delay and traffic performance, a before and after study has been conducted at five intersections. Traffic volume counts before installing lead/lag phasing are available from AD municipality. The traffic performance and average deal value has been estimated using Synchro software. Table 10 shows summary of the output results of the software. It shows a significant positive impact of illustrating lead/lag left turn phasing on the overall average delay whereas the delay reduced by values from 17 to 76 % during AM peak and a reduction from 7 to 60 % during PM peak.

5.5 Impact of Lead/Lag Signal Phasing on the Safety Performant

Drivers at a specific approach used to move most of the time (in our case it is split phasing system), when a lead-lag system was first implemented, some drivers on the lag approach kept doing the same and accelerating once the through movement starts to move thinking that their dedicated signal head is changed to green as well. However, lead lag signal phasing operation can be extremely beneficial but most traffic engineers wary of implementing it because it violates driver's expectations and it may cause accidents [25].

In a previous study the impact of left turn phasing sequences in AD signalized intersections was investigated [26]. It showed that a significant increase in accidents between left turn and through movements after changing the phasing sequence from split phasing to lead lag. Also, the approaches that have changed speed limit from 60 to 80 km/h experienced a significant increase. Referring to the location of the approach, there was no difference observed between CBD intersections and

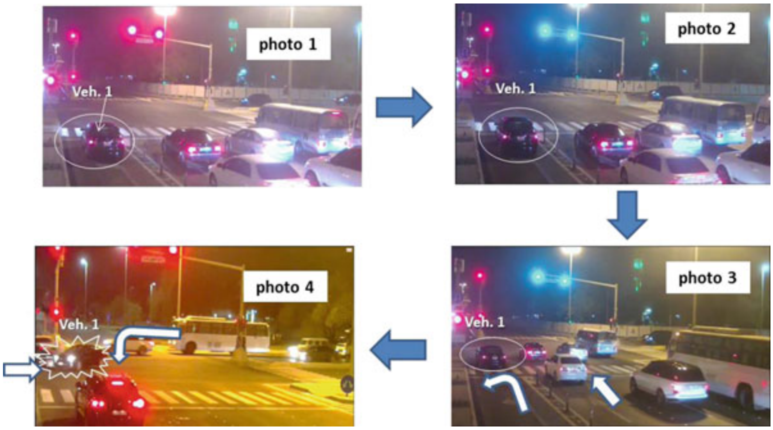


Fig. 4 An Illustration of a real accident occurrence due to lag left turn phase

Non CBD intersections. However, the study showed that the leading left turn approaches were safer than lagging left turns.

Figure 4 shows a real accident photos occurred at lag left turn phasing approach. It shows when the green light started for though the driver of (veh. 1) moved without recognition that he left turn has a separate traffic signal. This video illustrates the high risk of lag-left turn phasing system and why it has negative impact on the safety. So countermeasure should be taken in such cases to improve safety level.

6 Conclusions

This study mainly aims to investigate the value and contributing variables of the start-up delay at AD signalized intersections. The impact of applying the lead/lag left turn movement has been intensively analyzed with respect to delay and safety. The data analysis showed that the overall value of the start-up delay is 2.201 s the start-up delay is significantly differences between though and left movements, CBD and non-CBD locations, split and lag, lead and lag phasing at significant level of 95 %. However, there are not significant differences between peak and off-peak periods and between lead and split phasing. Traffic performance analysis at five intersections showed that the total average delays are reduced by 7–76 % after the illustration on the lead/lag phasing sequence. However, the safety level was reduced due to increasing the accident rates especially o the lag-phasing approaches. In conclusion, the lead/lag phasing systems has a significant impact on the start-up delay, total delay per vehicle at intersections and safety level. It is significantly improve the operational performance of traffic but it has a negative impact on safety. Therefore, countermeasures should be taken at lag-left turn approaches of intersection to improve safety.

References

1. Special Report No. 209: Highway capacity manual, 3rd edn. TRB, National Research Council, Washington D.C. (1994)
2. Brehmer C.L., Kacir, K.C., Noyce, D.A., Manser M.P.: Evaluation of traffic signal displays for protected/permissive left-turn control. National Cooperative Highway Research program NCHRP Washington D.C. Report 493, March 2003
3. Highway Capacity Manual 2000: TRB, National Research Council, Washington D.C. (2000)
4. Al-Ghamdi, A.S.: Entering headway for through movements at urban signalized intersections. Transp. Res. Rec. **1675**(1999), 42–47 (1999)
5. Leong, H.J.W.: Some aspects of urban intersection capacity. In: Proceedings of Australia Road Research Board, vol. 2, Part 1, pp. 305–338 (1964)
6. Gerlough, D.L., Wagner, F.A.: Improved criteria for traffic signals at individual intersection. Highway Research Board, National Research Council, NCHRP Report 32: Washington D.C. pp. 42 (1967)
7. Carstens, R.L.: Some traffic parameters at signalized intersections. Traffic Eng. **41**(11), 33–36 (1971)
8. Agent, K., Crabtree, J.: Analysis of lost times at signalized intersections. Kentucky Transportation Research Program. University of Kentucky, Lexington (1983)
9. Lee, J., Chen, R.L.: Entering headway at signalized intersections in a small metropolitan area. Transportation Research Record 1091, TRB, National Research Council, Washington D.C., pp. 117–126 (1986)
10. Roess, R.P., Papayannoupoulis, V.N., Ulerio, J.M., Levinson, H.S.: Levels of service in shared-permissive left-turn lane groups at signalized intersections. Report DTFH 61-87-C-00012. Transportation Training and Research Center, Polytechnic University, Brooklyn, N.Y. (1989)
11. Efstathiadis S., Machemehl, R.: Congress avenue regional arterial study: variability of departure headways. Center for Transportation Research, University of Texas, Research Report No. 60019-2 (1995)
12. Jacobs, N.: The determination of star-up lost time and saturation flow rates at signalized intersections (in Afrikaans). Final year thesis. Department of civil engineering, University of Stellenbosch (1998)
13. Honglongli and Prevedourod, P. D. Detailed observations for saturated headways and start-up lost time. TRB (2002)
14. David, S.H., Kevin, P.H., Steven, D.S., Joshua, S., Patrick, M., Halston, T.: Implications of distracted driving on start-up lost time for dual left-turn lanes. J. Transp. Eng. **139**(9), 923–930 (2013)
15. George Jr., E.T., Heroy Jr., F.M.: Starting response of traffic at signalized intersections. Traffic Engineering, Institute of Traffic Engineering, Washington D.C., pp. 39-43 (1966)
16. Long, G.: Startup delays of queued vehicles. The 84th Annual Meeting of the Transportation research Board (TRB-2005) Washington D.C., Jan 2005
17. Bonneson J.A.: Study of headway and lost time at single-point urban interchanges. Transportation Research Record 1365, TRB National Research Council, pp. 30–39, Washington D.C. (1992)
18. Sun, H., Yang, J., Wang. L., Li, L., Wu, B.: Saturation flow rate and start-up lost time of dual-left lanes at signalized intersection in rainy weather condition. In: Proceedings of the 13th COTA International Conference of Transportation Professionals (2013)
19. Noyce, D.A., Farmbro D.B., Kacir, K.C.: Traffic characteristics of protected/permitted left-turn signal display. The 79th Annual Meeting of the Transportation research Board TRB, Washington D.C., January 2000
20. NCHRP Project 3-54: Evaluation of traffic signal displays for protected-permitted left turn control. National Cooperative Highway Research program, NCHRP Project 3-54 (2000)

21. Chris Sheffer, P.E., Janson, B.N.: Accident and capacity comparisons of leading and lagging left-turn signal phasing. Presented Transportation Research Board TRB-1999 Meeting Washington D.C. (1999)
22. Upchurch J., Wright, C.H.: Before and after comparison of leading exclusive and permissive/exclusive lagging left-turn phasing. In: Transportation Research Record 1368, TRB, National Research Council, Washington D.C. (1992)
23. Grover, A.: Minimize delay, maximize progression with protected permissive lead/lag phasing. Technical Workshop. Inland Empire Section of ITE, 14 Dec 1995
24. Machemehl, R.B., Mechler, A.M.: Comparative analysis of left-turn phasing sequencing. Federal Highway Research Report 258-2 (1983)
25. Buckholz, J.: The 10 major pitfalls of coordinated signal timing. ITE J. **63**(8), 26–29 (1993)
26. Al-Ghafli, A., Garib, A., Sarhan, M., Al-Harthi, H.: Effect of signal phase sequences on signalized intersection safety performance case study from Abu Dhabi, U.A.E. 17th IRF World Meeting and Exhibition, Riyadh, Saudi Arabia (2013)

Improving a Roundabout Design Through Simulation in a Unity 3D Virtual Environment

Li Wang, David Johnson and Yingzi Lin

Abstract Traffic congestion is one of the most serious problems in big cities. This paper proposed a simplified design of roundabout to simulate and improve the traffic condition of a real-life roundabout in Somerville, MA. A virtual environment was built by using Unity 3D to simulate the Powder House Square roundabout. This study used the driving simulator in Northeastern University's Intelligent Human-Machine System (IHMS) Laboratory to conduct the experiment, using a NASA Task Load Index (TLX) to measure the workload experienced by the participants. The experiment results show that the total time spent on simplified sector was 47.2 % less than the original design, and the number of accidents were reduced by 60.9 %. The average frustration level dropped by 79.2 % on the simplified sector according to the NASA TLX survey. Simulation results showed strong evidence that traffic conditions improved a lot by using the simplified design.

Keywords Roundabout · Virtual environment · Simulation

1 Introduction

Despite being a daily activity, driving in city traffic can be as stressful for drivers as skydiving [1]. What is more serious is that there are lots of traffic accidents every day in the world. The Global status report on road safety 2015 indicates that more than 1.2 million people die each year on the world's roads [2]. Although there are many factors that affect traffic, the most controllable factors are human factors. Solving traffic problems by using human factors methods can make a great contribution to humanity. Intersections are a major constraint to traffic flow, especially when a non-standard style of intersection or a traffic signal is used [3]. The inter-

L. Wang · D. Johnson · Y. Lin (✉)
Intelligent Human-Machine Systems Lab, Department of Mechanical and Industrial Engineering, College of Engineering, Northeastern University, Boston, MA 02115, USA
e-mail: yi.lin@neu.edu

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_19

section that inspired this research is a combination rotary and traffic light system at Powder House Square in Somerville, MA.

There are several challenges presented by this intersection in Powder House Square [4]. There is high level of pedestrian, motor vehicle, and bicycle traffic, which causes congestion and traffic accidents. When drivers approach the rotary, confusion is created by having to be simultaneously aware of traffic lights, traffic flow and pedestrians within the rotary. Pedestrians, having an automatic right-of-way as long as they are utilizing a crosswalk, have an easier time navigating this intersection. The design of the rotary means that their use of crosswalk signals causes all traffic throughout the rotary to come to a complete stop. This study hypothesized that traffic conditions would be safer and more efficient if some of the traffic lights, crosswalks and stop signs were eliminated or moved in order to ensure that drivers only need to make one decision at a time.

A virtual environment [5] was built by using Unity 3D to simulate the conditions experienced in Powder House Square. The driving simulator in Northeastern University's Intelligent Human-Machine System (IHMS) Laboratory seen in Fig. 1 was used in this experiment. To test the hypothesis, a square road map with two roundabouts was designed. Eleven subjects took part in this experiment. Time spent on each sector and the number of mistakes and accidents in each sector were recorded by experimenter. After the experiment, all of the subjects were asked to complete the NASA Task Load Index (TLX) survey in order to effectively measure the workload experienced by the subjects in each sector [6].

The experiment results show that the total time spent in the simplified sector was reduced by 47.2 % and the number of accidents were reduced by 60.9 %. What is more, the average frustration level dropped by 79.2 % on the simplified sector,

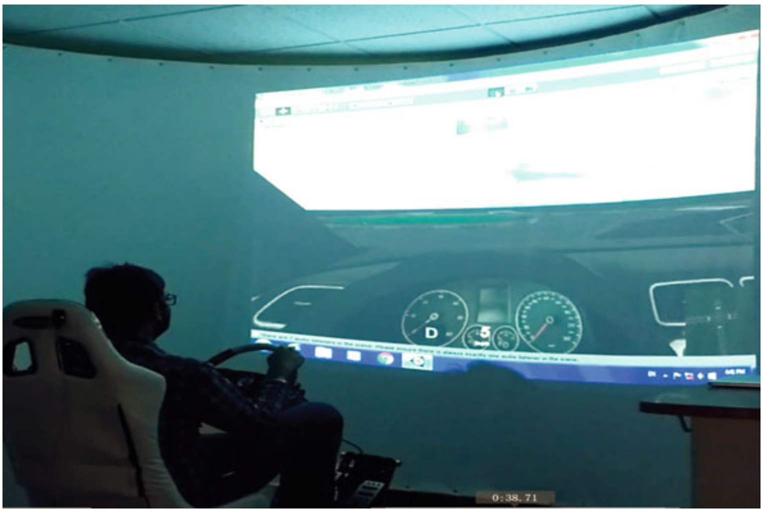


Fig. 1 The driving simulator in Northeastern University's IHMSL

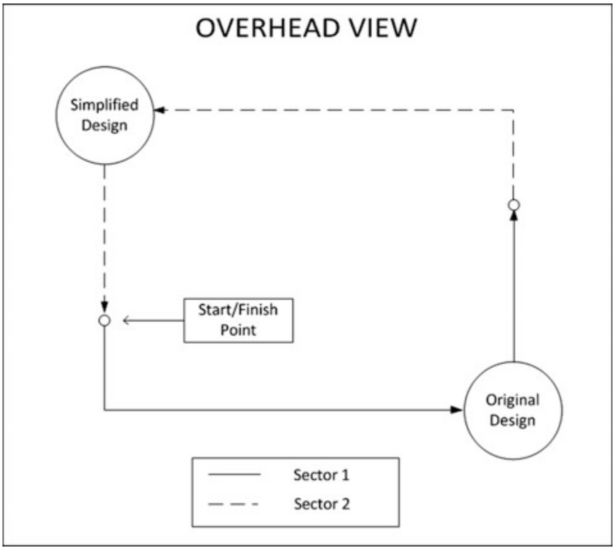


Fig. 2 Overhead view of the square road map with two different sectors

according to NASA TLX survey data. It is clearly shown from the NASA TLX survey data that subjects experienced a much lower mental and physical workload in the simplified design sector than on the original design sector.

2 Methods

Unity 3D was used in this study to build the virtual environment [5]. Unity 3D is a flexible development platform for creating 3D interactive experiences, and is a relatively quick platform to learn while still providing plenty of flexibility. A square road map with two different sectors was the principal design of the study, as can be seen in Fig. 2. One sector had a rotary emulating the original design of Powder House Square, and the other sector had a simplified design roundabout [7–10]. A comparison of the two different sectors were shown in Fig. 3. Part A is the original design and part B is the simplified design.

In the original design, there were 16 traffic lights and 7 crosswalks, all positioned in a way that contributes to confusion and frustration. The original design forced drivers to constantly multitask to maintain awareness of traffic signal status, rotary traffic, and pedestrian traffic. The objective of the simplified design was to ensure that drivers were only completing one task at a time.

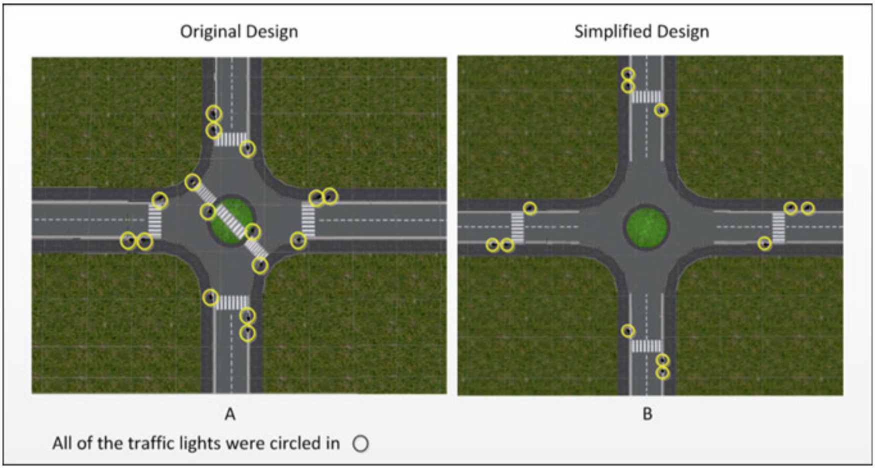


Fig. 3 A comparison of the two different sectors (*A* original design; *B* simplified design)

The first objective in the simplified design was to remove pedestrians from the middle of the rotary by removing the crosswalks cutting through the center as well as the intra-rotary traffic lights. This ensures that drivers going through the rotary only have one task: to focus on other drivers within the rotary. The second objective was to remove pedestrians from the perimeter of the rotary by moving the cross-walks 2 car-lengths away from the rotary's entrance. This setup makes it so that drivers entering the rotary only have to be concerned about waiting for a gap in rotary traffic, and so that drivers exiting the rotary may leave freely. It also limits the pedestrian interaction with drivers to situations where the driver has their full attention. All these factors combine to eliminate multitasking in the simplified design.

2.1 Pilot Test

This experiment conducted a pilot test before the formal portion in order to refine the simulation process [11]. The first change was to the starting point. Originally, the starting point was very close to the first roundabout. It was found that if the start point was very near the first roundabout, the subjects could be strongly affected by having very little reaction time. Another major change was to the frequency of occurrences of computer-driven cars in the simulation. The reason that the change was made was to carefully balance the level of traffic so that it would not be too crowded or too empty [12]. Finally, a trial run for each subject was allowed so that they could acclimate to the sensitivity of the driving simulator.

2.2 Subject Selection

A group of 11 subjects took part in this experiment, all from the graduate student population at Northeastern University. There were 9 male subjects and 2 female subjects, all between the ages of 21 and 23 (μ 21.8 years, σ 0.87 years). All 11 subjects had a driver's license.

2.3 Experiment Instruction

First, each subject completed a trail run to acclimate to the sensitivity of the simulator. Second, each subject completed the whole experiment without interruption. During the experiment, the subjects were instructed to obey all traffic rules. All subjects were expected to recall how they felt in each sector. Finally, all subjects would complete a NASA TLX survey for each sector of the experiment after completing the simulation.

3 Data Collection

All data in this experiment was manually recorded. During the experiment, an observer recorded the time it took for each participant to complete each sector with a stopwatch. In addition, another observer counted the number of accidents and traffic violations in each sector.

After each subject completed the experiment, they were asked to fill out a NASA TLX survey for each sector, original and simplified. The survey asked the subjects to assess their perceived Mental Workload, Physical Workload, Performance Level, Pace of the Task, Effort Required for Normalcy, and Frustration Level. This data is shown in Table 1.

4 Results and Discussion

4.1 Accidents and Violations

Several differences were discovered between the data for the original and simplified designs. The average number of accidents or traffic violations subjects committed in the original design roundabout was 3.8, while for the simplified design roundabout the average was 1.5. This is a decrease in incidents by over 50 %. Another

Table 1 Data from NASA TLX survey

Factor	Type	Subject no.										
		1	2	3	4	5	6	7	8	9	10	11
Mental demand	Original	14	16	10	10	9	16	14	13	11	12	3
	Simplified	10	5	6	10	4	4	20	13	6	6	2
Physical demand	Original	8	6	7	5	2	8	5	10	13	7	3
	Simplified	3	2	6	10	1	4	4	5	2	6	2
Temporal demand	Original	15	15	10	7	7	6	17	15	14	14	12
	Simplified	7	4	8	12	5	1	10	4	4	7	15
Performance	Original	5	5	10	13	8	10	18	5	3	9	1
	Simplified	17	16	11	14	13	17	13	17	16	16	20
Effort	Original	14	17	4	15	8	10	12	14	13	11	10
	Simplified	5	3	9	18	4	1	15	11	3	5	10
Frustration	Original	14	12	1	3	9	1	4	15	3	17	18
	Simplified	4	11	12	2	1	1	18	1	3	4	1

important performance marker is that only 2 subjects were able to complete the original sector without any incidents, while 5 were able to do so in the simplified version. The standard deviation dropped only slightly from 1.38 incidents in the original sector to 1.17 incidents in the simplified sector. It is clear from this data that the simplified intersection is much safer than the original version (Fig. 4).

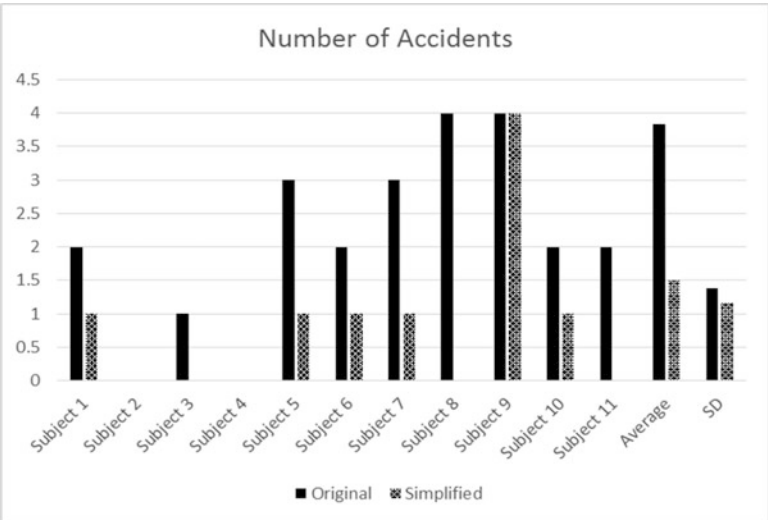


Fig. 4 Number of mistakes that each subject made

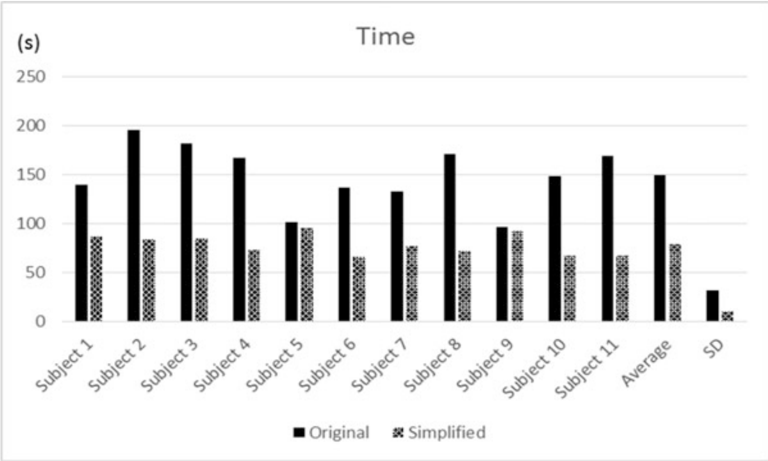


Fig. 5 Time that each subject spent on two different sectors and the average time subjects spent on two sectors

4.2 Time

The time that subjects spent on each sector is also an important factor to evaluate the two differently designed roundabouts. The range of time that subjects spent on original design roundabout was from 96.5 to 195.9 s. For the simplified design roundabout it was from 66.6 to 95.3 s. Even the slowest time in the simplified intersection (95.3 s) was faster than the fastest time in the original (96.5 s) (Fig. 5).

The average times for each sector were 149.1 s in the original design and 78.7 s in the simplified design. A 47.2 % drop is seen in the simplified design, which illustrates that the subjects' efficiency improved greatly in simplified design roundabout. The standard deviations for each sector were also very different. The subjects were much more consistent in the simplified sector, with a standard deviation of only 10.5 s, a 66.9 % reduction from the 31.7 s standard deviation in the original sector.

4.3 Workload

NASA-TLX includes six subscales: Mental Demand, Physical Demand, Temporal Demand, Own Performance, Effort, and Frustration. The maximum value of each subscale is 21 and the minimum value is 1. In this experiment, all of the subjects finished the NASA-TLX survey immediately after the experiment. Figure 6 reflects the average values of the results of the NASA-TLX surveys for each sector.

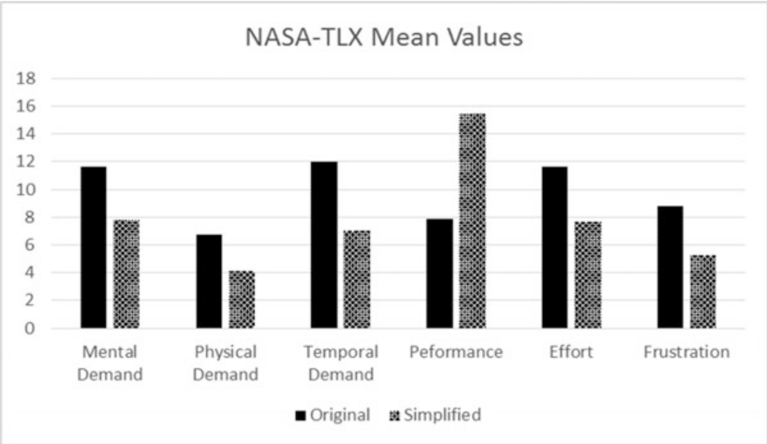


Fig. 6 Average values of NASA-TLX surveys for mental demand, physical demand, temporal demand, performance, effort, and frustration

Mental demand [13] shows how much mental and perceptual activity was required and physical demand shows how much physical activity was required. In this experiment, less mental demand means that the subjects do not need to pay much attention to the traffic lights, pedestrians and traffic flows. At the same time, less physical demand means the subjects do not need to use the brake or turn the wheel frequently. The average value of the mental demand decreased by 32.8 % and the average value of the physical demand dropped by 39.2 % in the simplified roundabout.

Temporal demand was defined as how rushed the subjects felt during each sector. The subjects were under no strict time constraints and were not told that they would be timed. The traffic experienced within each sector likely contributed to the feeling of being rushed. In this study, temporal demand dropped by a factor of 41.7 %, indicating a significant decrease in the pressure subjects felt in completing the simplified sector.

The Performance value indicated how successful the subjects felt they were in completing each sector [14]. The average performance increased by 95.4 % in the simplified sector over the original sector. Effort was defined as how hard each subject had to try to achieve that level of performance. Subjects experienced a 34.4 % drop in effort on average from the original to the simplified sector, while nearly doubling their performance.

Subjects graded their frustration level based on how “insecure, discouraged, irritated, stressed, and annoyed” they felt [6]. A 40.2 % drop in frustration levels from the original to the simplified design, along with all the other factors captured in the NASA-TLX survey, made it very clear that the workload experienced in the simplified sector was greatly reduced while performance nearly doubled.

5 Conclusion

The new location of the signals and crosswalks in the simplified design directly lead to less confusion and therefore a lower probability of accidents within the rotary. Anyone going through the rotary for the first time is less likely to make a mistake or get into an accident. This greatly increased the level of safety experienced in the simplified sector over the original, thus reinforcing the hypothesis that traffic conditions would be safer and more efficient if some of the traffic lights, crosswalks and stop signs were eliminated or moved in order to have each driver only completing one task at a time.

We conclude that the simplified design roundabout works very well in the simulation environment as the simulation results show strong evidence that traffic condition improved a lot by eliminating all of the traffic lights, crosswalks, and stop signs. This simplified design roundabout can also be applied to other roundabouts all over the world, making a great contribution to humanity.

6 Future Work

For further study, more details should be simulated in the virtual environment with a greater number of subjects to get more accurate data sets. One detail that would be important to test is the addition of simulated pedestrian traffic around each rotary. This would help to add a level of realism to the simulation, as well as measure the impact of pedestrian traffic flow through each style of intersection. Another improvement would be to have the simulation software automatically record collisions, near-misses, and traffic violations on its own. This would help to quantify the severity of incidents instead of all incidents being considered equal.

Another consideration for future work would be to add eye tracking, electroencephalographic, and skin galvanic response sensors in order to capture valuable physiological data. This data could be combined with the NASA-TLX survey data to more accurately assess the workload experienced by subjects though each sector.

Acknowledgments The work was supported by NSF grant # 0954579. The authors would like to acknowledge the members of the Intelligent Human-Machine Systems Laboratory and Mr. Sameep Pankaj Khatri for their help with this project. We would also like to extend our gratitude to the participants who took time to help with the experiments.

References

1. Williams, S.: Stress from city driving same as skydiving. Retrieved from: http://www.nydailynews.com/news/national/stress-city-driving-skydiving-mit-study-article-1.1365361 (2013)
2. Toroyan, T.: Global status report on road safety 2015. Supporting a decade of action. World Health Organization, Department of Violence and Injury Prevention and Disability, Geneva (2013)
3. Bergh, C., Retting, R., Myers, E.: Continued reliance on traffic signals: the cost of missed opportunities to improve traffic flow and safety at urban intersections. Insurance Institute for Highway Safety, pp. 6–9 (2005)
4. Google: Google Maps satellite view of Powder House Square. Retrieved Dec 13, 2015, from https://www.google.com/maps/@42.4011387,-71.116808,60m/data=!3m1!1e3 (n.d.)
5. Cai, H., Lin, Y., Cheng, B.: Mimicking human driving behaviors autonomous vehicles in realistic traffic flow simulation. Int. J. Simul. Process Model. **6**(2), 126–136 (2010)
6. 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)
7. Baranowski, B.: Pedestrian crosswalk signals at roundabouts: where are they applicable. Transportation Research E-Circular No. E-C083 (2005)
8. Bergh, C., Retting, R.A., Myers, E.: Continued reliance on traffic signals: the cost of missed opportunities to improve traffic flow and safety at urban intersections (2005)
9. Inman, V.W., Davis, G.W.: Synthesis of literature relevant to roundabout signalization to provide pedestrian access. Access Board, Jan 2007
10. Meyer, M.D.: Transportation Planning Handbook (2009)
11. Van Teijlingen, E., Hundley, V.: The importance of pilot studies. Nurs. Stand. **16**(40), 33–36 (2002)
12. Kimber, R.M.: The traffic capacity of roundabouts. No. TRRL LR942 Monograph (1980)
13. Thorvald, P., Lindblom, J.: Initial development of a cognitive load assessment tool. In: The 5th AHFE International Conference on Applied Human Factors and Ergonomics, Krakow, Poland, AHFE, 19–23 July 2014
14. Cai, H., Lin, Y.: Modeling of operators' emotions and task performance in a virtual driving environment. Int. J. Hum. Comput. Stud. **69**(9), 571–586 (2011)

Part IV
Road and Rail—Driver, Behavior, Distraction and Fatigue

Exploration of the SHRP 2 NDS: Development of a Distracted Driving Prediction Model

Syndney Jenkins, Julius Codjoe, Ciprian Alecsandru and Sherif Ishak

Abstract The objective of this research was to use the SHRP 2 NDS data to predict whether drivers were engaged in any of three specific groups of distracting tasks or no secondary task at all. The tasks that were examined included: talking or listening on a hand-held phone, texting or dialing on a hand-held phone, and driver interaction with an adjacent passenger. Multiple logistic regression was used to determine the odds of driver engagement in one of the secondary tasks given corresponding driving performance data. The results indicated there were differences in the driving performance measures when the drivers were engaged in a secondary task. However, the results of the MLR tests indicated the subset of this data could not be used to develop prediction models with statistically significant predictive power.

Keywords Naturalistic driving studies · Secondary tasks · Distracted driving · Time series data · Driver performance · Multiple logistic regression

S. Jenkins (✉) · J. Codjoe · S. Ishak
Department of Civil and Environmental Engineering, Louisiana State University, Patrick F. Taylor Hall, Room #3416, Baton Rouge, LA 70803, USA
e-mail: sjenkins1315@gmail.com

J. Codjoe
e-mail: jcodjo1@lsu.edu

S. Ishak
e-mail: sishak@lsu.edu

C. Alecsandru
Department of Building, Civil and Environmental Engineering, Concordia University, 1515 St. Catherine W, Montreal, QC H3G 1M8, Canada
e-mail: ciprian@bcee.concordia.ca

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_20

1 Introduction

Distracted driving continues to be a major topic in the area of transportation safety. Some form of anti-distraction legislation has been enacted in forty-four U.S. states to date including Guam and Puerto Rico. The U.S. Department of Transportation has established programs, and even an official government website, all aimed at educating the public on the dangers of distracted driving [1].

Many researchers in past literature have used driving simulators to mimic the driving experience and more specifically, to measure the effect distraction devices have on drivers. However, naturalistic driving studies (NDS) offer the ability to observe drivers in their own vehicles, driving their typical commutes, and exhibiting their normal driving behavior [2]. This aspect, that is unique to NDS, more accurately reflects actual driving behavior when compared to driver simulator studies that use a simulation vehicle and ask the driver to maneuver through a simulated environment. The USDOT's second Strategic Highway Research Program (SHRP 2) organized and funded a massive naturalistic driving study. This study yielded data that includes: 32 million vehicle miles, 12,500 miles driven, 7000 near-crashes, 700 crashes, cell phone records, and a final database that is expected to approach 2 petabytes (or 2000 terabytes) of trip information. SHRP 2 began releasing this database in 2014 to be used by the researching public [3].

The objectives of this research were to identify appropriate performance measures that could be used as surrogate measures of distraction within the SHRP 2 dataset, and to develop a model to predict driver distraction by cell phone using said surrogate measures.

2 Background

Driving simulators are used frequently in traffic research. They are an inexpensive alternative to other experimental methods that can at times be either unethical or unsafe to complete [4]. An exhaustive literature review conducted by Bach found that of the 100 papers reviewed, 52 % of them involved driving simulators while 37 % involved instrumented vehicles (e.g. naturalistic studies) [5]. Although driving simulators with high fidelities can closely mimic an actual driving experience, naturalistic driving studies offer a truly realistic picture of driver behavior because they analyze actual drives on authentic roads.

Naturalistic driving studies (NDS) consist of the observation of drivers in their own vehicles and driving their normal commutes. The vehicles however, are fitted with sensors and other data collection gadgets which are usually add-ons to the in-vehicle systems. While NDS will produce more realistic scenarios, and thereby more valuable data to study driver behavior, they are expensive to use and the collection of data could be problematic. The first large-scale NDS conducted was the 100-Car Naturalistic Driving Study which involved 241 drivers over an

18-month period resulting in about 3 M vehicles miles and yielded 42,300 data hours, 82 crashes, 761 near-crashes, and 8295 critical incidents [6]. Virginia Tech Transportation Institute (VTTI) completed the project with funding from the National Highway Traffic Safety Administration, Virginia Tech, Virginia Department of Transportation and Virginia Transportation Research Council [7].

Due to NDS being a behavioral-based observational experiment method, there are many ways this data can be used to study driver behavior and risk analysis. Some of the studies that have been conducted using the 100-Car NDS include validation of near-crashes as crash surrogates [8], assessing safety critical braking events [9], modeling of driver car-following behavior [10], and examining driver inattention [11]. A study conducted by Feng Guo and Youjia Fang, also used data from the 100-Car Naturalistic Driving Study. They focused on predicting high risk drivers and identifying factors associated with individual driver risk. Driver age, personality and critical incident rate were determined to have major impacts on both crash and near-crash risk and these factors can be used to predict future crashes or near-crashes. The researchers developed logistic models as the prediction method which proved to possess high "predictive powers" [12]. University of Michigan Transportation Research Institute studied driver performance while engaging in secondary tasks using naturalistic data they collected themselves [13].

In the University of Michigan study and other previous research, point estimates such as variance or mean of the performance measures were selected for analysis. However, point estimates represent only one data point over a length of time, so much information is lost when the data is averaged or the variance is computed in order to obtain a point estimate. This study utilizes actual time series data in the analysis. This distinction is important because the plethora of data points accumulated in time series data has the ability to reveal more information than would be possible using only a mean or variance.

3 Methods

The methodology consisted of data retrieval from the InSight website, data editing and aggregation, tests for normality followed by multiple logistic regression.

3.1 Data Retrieval

The data used in this research was the SHRP 2 NDS database which was released through the program's InSight website. Additional data is continuously uploaded to the webpage, therefore it is important to note that all data utilized for this study was released as of March 2015. It is also important to note that all NDS data used in this study was taken from Florida driver samples exclusively. Finally, NDS data specifically from the Events and Trips website categories were used in this study. In addition to these categories additional information was required to link the sample drivers to their trips and event information. This linkage was important because it enabled comparisons of driver performance measures based on driver gender, age and location. Participant ID and additional demographic information including state origin, age and gender of the drivers was retrieved directly from VTTI via a Data Sharing Agreement. The driving performance measures of GPS speed, lateral and longitudinal acceleration, throttle position and yaw rate were selected because literature revealed they were most frequently used in driver behavior research [14].

3.2 Data Aggregation and Editing

Proper data editing before applying said data as input into analyses can aid in the assurance that the results obtained are accurate. In regards to this research, the data editing process included checking data entries to ensure the values were within an acceptable range and logically reasonable as well as identifying outliers or missing data. However, since time series data was used in this research, the first step taken in the data editing process dealt with aggregating the time intervals already in place.

Data on the time series variables was collected over a 20-s time interval for each drive. Within the twenty second time interval the data was broken down into 0.1-s intervals. For example, the data for the GPS speed variable was represented by 200 data points displayed in 0.1-s increments in the database to account for the twenty seconds of data collected. In order to reduce the size of the database, the time series data was aggregated into 1-s increments instead of the original interval of 0.1 s. The "time series" procedure in SAS statistical software was used in order to accomplish this. In order to aggregate each set of 200 data points into 20 points representing the 20 s of data corresponding to each drive, the absolute value of the maximum change for each data point was used for the throttle position, lateral acceleration, longitudinal acceleration and yaw rate performance measures. The throttle position variable contained only positive measurements, therefore, simply the maximum value for each observation per second was kept in the final dataset. Equation 1 displays the function used to aggregate the lateral acceleration, longitudinal acceleration and yaw rate variables into 1-s increments.

$$|\mathrm{Max(x)} - \mathrm{Min(x)}|. \tag{1}$$

Regarding the GPS speed variable, the variance of the driver's speed per second was used to aggregate the data. The Time Series procedure in SAS statistical software was used to perform all of the aggregation. To serve as an example, the code used to aggregate the throttle position variable is displayed below.

```
proc timeseries data=baseline_throttle_position

out=baseline_throttle_position_timeseries;
id time interval=seconds accumulate=maximum;
by event_id;
var value;

run;
```

After the data was aggregated into 1-s intervals, the next step in the data editing process was to ensure the values were within an acceptable range. The upper and lower data ranges of each time series variable were defined in the Trip Data category on the InSight webpage. All values outside of the predefined range limits were removed from the dataset for each of the five time series variables studied. Next, any entry that contained missing information was also removed. Potential outliers were inspected using the distribution analysis task in SAS Enterprise Guide statistical software and removed once identified. The Kolmogorov-Smirnov test was used to test for normality. For an alpha value set equal to 0.05, datasets for all five performance measures resulted in statistically significant outcomes. Due to these findings, it was concluded all of the datasets were non-normally distributed.

3.3 Multiple Logistic Regression Analysis

Logistic regression is frequently used in research to predict the probability that a particular outcome will occur. The outcome can either be a continuous-level variable or a dichotomous (binary) variable [15]. However, the outcomes are usually classified in a binary nature in logistic regression. In this case the dependent variable is dichotomous and is coded as “1” if the event did occur and “0” if the event did not occur. During the analysis, the logistic function estimates the probability that the specified event will occur as a function of unit change in the independent variable(s) [16].

The intent of the analysis was to use all five independent variables (GPS speed, lateral acceleration, longitudinal acceleration, throttle position and yaw rate) to predict whether the driver was or was not engaged in a secondary task. Since five independent variables were considered, multiple logistic regression (MLR) was used instead of simple logistic regression. Three separate MLR tests were run in order to compare the statistical output between the control and the three individual secondary tasks of concern. All instances where drivers were not engaged in a secondary task served as the control group. Table 1 describes the scenarios examined in each of the tests.

In order to accurately interpret the results of the MLR, the binary predictor variables used must be coded in a very specific manner. Typically in MLR, the group that is to be used as the focal or reference group is coded as “0” and the other outcome is coded as “1”. The focal group in each of the comparison tests was the

Table 1 Description of multiple logistic regression tests

MLR test	Description
Control versus Group 1	Engaged in no secondary task versus talking or listening on cell phone (hand-held)
Control versus Group 2	Engaged in no secondary task versus texting or dialing on cell phone (hand-held)
Control versus Group 3	Engaged in no secondary task versus adjacent passenger interaction

individual secondary task in which the driver was engaged. Therefore, for each comparison test the variable that described No Secondary Task was coded as "1" and the specified secondary task was coded "0".

Hosmer and Lemeshow values, Maximum Likelikhood Estimates, Odds Ratios and the Likelihood Ratio Score *p*-values were all used to interpret the MLR test results. Hosmer and Lemeshow values assess whether the predicted probabilities match the observed probabilities using a chi-square statistic. If the *p*-values for this test were not significant, this indicated that the model predictions and actual observations were about the same and the model provided a good fit of the data. The estimates computed in the Maximum Likelihood Estimation (MLE) served as the coefficients used to create a regression line. The odds ratio values were associated with each predictor and described the odds of a case being coded as "0" on the dependent variable. It indicated the amount of change expected in the log odds when there was a 1-unit change in the predictor variable. Finally, the Likelihood Ratio Score *p*-value assisted in identifying the validity of the test. If this measure was significant it could be stated the output resulting from the test were better than chance.

4 Results and Discussion

4.1 Test for Differences in Performance Measures Among Groups

Prior to conducting the MLR, Chi-square tests were conducted in order to examine if there were any statistical differences between the driving performance measures when the driver was not engaged in a secondary task (Control group) and when the driver was engaged in a secondary task (Groups 1–3). Based on the results of the preliminary Chi-square analysis, the performance variables for the secondary tasks that proved to be statistically different from their Control group counterparts were later input into a subsequent Multiple Logistic Regression model to see if any of said variables could be used to predict the driver's behavior.

There proved to be statistical differences between the control performance variables and those when the driver was engaged in a secondary task for almost all of the performance variables. The only instances where this was not the case was for driver's GPS speed when engaged in texting or dialing (Group 2 task). Therefore, in the subsequent MLR tests every variable except the GPS speed in the Control versus Group 2 test was input into the model.

4.2 MLR Results

The MLR tests were all run using the Backward Elimination Method in SAS Enterprise Guide 6.1. Under the Backward Elimination Method, all of the dependent variables that were proven statistically different under the previous Chi-square test were initially input into the model and variables were removed one by one until only variables that produced F statistics significant at the significance level of 0.05 remained. A summary of the variables initially input into the model and the variables that remained at the conclusion of each test are provided in Tables 2, 3 and 4.

The first model compared the Control group to the Group 1 tasks of talking or listening on the phone. The performance variables GPS speed, throttle position and yaw rate proved to have statistically significant F-values and therefore remained in the model at the conclusion of the Backward Elimination Method. However, the *p*-value associated with the throttle position measure was not significant and thus was not included in the final regression equation. The interaction between the GPS speed and longitudinal acceleration as well as the throttle position and yaw rate

Table 2 Control versus Group 1 MLR results

Dependent variables	Effect type in MLR	After backwards elimination	*p*-value at conclusion of test
GPS speed	Main	Remained	0.0166
Lateral acceleration	Main	Eliminated	–
Longitudinal acceleration	Main	Eliminated	–
Throttle position	Main	Remained	0.3615
Yaw rate	Main	Remained	0.0177
GPS speed * Lateral acceleration	Interaction	Eliminated	–
GPS speed * Longitudinal acceleration	Interaction	Eliminated	–
GPS speed * Throttle position	Interaction	Eliminated	–
GPS speed * Yaw rate	Interaction	Remained	0.0445
Lateral acceleration * Longitudinal acceleration	Interaction	Eliminated	–
Lateral acceleration * Throttle position	Interaction	Eliminated	–
Lateral acceleration * Yaw rate	Interaction	Eliminated	–
Longitudinal acceleration * Throttle position	Interaction	Eliminated	–
Longitudinal acceleration * Yaw rate	Interaction	Eliminated	–
Throttle position * Yaw rate	Interaction	Remained	0.0337

Table 3 Control versus Group 2 MLR results

Dependent variables	Effect type in MLR	After backwards elimination	*p*-value at conclusion of test
Lateral acceleration	Main	Eliminated	–
Longitudinal acceleration	Main	Eliminated	–
Throttle position	Main	Remained	0.0032
Yaw rate	Main	Remained	<0.0001
Lateral acceleration * Longitudinal acceleration	Interaction	Eliminated	–
Lateral acceleration * Throttle position	Interaction	Eliminated	–
Lateral acceleration * Yaw rate	Interaction	Eliminated	–
Longitudinal acceleration * Throttle position	Interaction	Eliminated	–
Longitudinal acceleration * Yaw rate	Interaction	Eliminated	–
Throttle position * Yaw rate	Interaction	Remained	<0.0001

Table 4 Control versus Group 3 MLR results

Dependent variables	Effect type in MLR	After backwards elimination	*p*-value at conclusion of test
GPS speed	Main	Remained	0.0077
Lateral acceleration	Main	Remained	0.0363
Longitudinal acceleration	Main	Remained	0.084
Throttle position	Main	Remained	<0.0001
Yaw rate	Main	Remained	0.0233
GPS speed * Lateral acceleration	Interaction	Eliminated	–
GPS speed * Longitudinal acceleration	Interaction	Remained	0.0342
GPS speed * Throttle position	Interaction	Remained	0.0113
GPS speed * Yaw rate	Interaction	Remained	0.0132
Lateral acceleration * Longitudinal acceleration	Interaction	Eliminated	–
Lateral acceleration * Throttle position	Interaction	Eliminated	–
Lateral acceleration * Yaw rate	Interaction	Remained	0.0044
Longitudinal acceleration * Throttle position	Interaction	Eliminated	–
Longitudinal acceleration * Yaw rate	Interaction	Eliminated	–
Throttle position * Yaw rate	Interaction	Remained	0.0027

variables also proved strong enough to remain in the model. Equation 2 displays regression line formed from this test.

$$\begin{aligned} y = & -2.3366 + 0.000014(\text{GPS Speed}) + -0.0456(\text{Yaw Rate}) \\ & + -0.00001(\text{GPS Speed} * \text{Yaw Rate}) + 0.00145(\text{Throt Pos} * \text{Yaw Rate}). \end{aligned} \quad (2)$$

The regression coefficients showed that an increase in driver speed causes an increase in the likelihood that said driver is talking or listening on the phone. The converse was true for the yaw rate variable. When considering the driver's speed and yaw rate concurrently, there seemed to be a negative correlation between the interaction of these variables and the likelihood the driver was talking or listening. The throttle position and yaw rate interaction conversely produced a positive correlation. The coefficients for the GPS speed and GPS speed and yaw rate interaction terms in the equation were relatively small, designating these variables were not tremendously affecting the y-term or the likelihood of the driver talking or listening on the phone. Table 2 displays the Control versus Group 1 MLR Results.

The Odds Ratios for the GPS speed, throttle position and yaw rate were very weak and all very close to one at 1.0, 1.004 and 0.983 respectively. This meant that a one unit change in any of these measures would only increase the odds of the driver being engaged in talking or listening on the phone by one time. Any Odds Ratio this close to one does not indicate predictive power in the model. The Likelihood Ratio Score for this test resulted in a *p*-value of <0.0001, which meant the results output by this test were better than chance. However with R-square value of 0.0023 and the Hosmer and Lemeshow *p*-value at <0.0001 the dependent variables do not account for the majority of the variance in the model and the predicted values cannot be said to statistically match those observed. This model was not a great fit of the data.

The next model compared the Control group to Group 2 tasks of texting or dialing. The performance variables throttle position, yaw rate, and the interaction between throttle position and yaw rate proved to have statistically significant F-values and therefore remained in the model at the conclusion of the Backward Elimination Method. Equation 3 displays regression line formed from this test. Table 3 displays the Control versus Group 2 MLR Results.

$$\begin{aligned} \mathrm{y} = & -3.1796 + 0.00859(\text{Throt Pos}) + -0.1270(\text{Yaw Rate}) \\ & + 0.00396(\text{Throt Pos} * \text{Yaw Rate}). \end{aligned} \tag{3}$$

The regression coefficients resulting from this test showed an increase in the driver's throttle position caused an increase in the likelihood the driver was texting or dialing on the phone. As in the first model, when the driver's yaw rate decreased this indicated an increase in the likelihood the driver was texting or dialing. When considering the driver's throttle position and yaw rate concurrently, there was a positive correlation between the interaction of these variables and the likelihood the driver was texting or dialing. The Odds Ratios for the throttle position and yaw rate were again weak and showed little predictive power due to their close proximity to one at 1.016 and 0.976 respectively.

The Likelihood Ratio Score for this test also resulted in a *p*-value of <0.0001, which meant the results output by this test were better than chance. However with the R-square value for this test was also extremely low at 0.0095. The Hosmer and Lemeshow *p*-value here equaled 0.0151 signifying the predicted values cannot be said to statistically match those observed. Although the Hosmer and Lemeshow *p*-value was significant, it was an improvement from the previous test and is much closer to the alpha of 0.05.

The final model compared the Control group to Group 3 tasks of interacting with an adjacent passenger. The following main effects remained in the model: GPS speed, lateral acceleration, longitudinal acceleration, throttle position, and yaw rate. These interaction effects also remained: GPS speed and longitudinal acceleration, GPS speed and throttle position, GPS speed and yaw rate, lateral acceleration and yaw rate and finally throttle position and yaw rate.

Equation 4 displays regression line formed from this test. The regression coefficients revealed a positive correlation for all of the main effects, an increase in the driver's speed, acceleration in either direction, throttle position and yaw rate corresponded to an increase in the likelihood the driver was interacting with his or her adjacent passenger. The sign convention for the yaw rate variable was positive for this test, however, in the previous tests it was negative. The coefficient values for most of the effects were quite small, meaning these variables were not tremendously affecting the y-term or the likelihood of the driver interacting with an adjacent passenger. The Odds Ratios for the GPS speed, lateral acceleration, longitudinal acceleration, throttle position and yaw rate were all basically equal to one once again (1.0, 1.0, 1.001, 1.008, 1.0). Similar to the previous two tests, this indicated the model had very little predictive power.

$$
\begin{aligned}
\mathrm{y} = & -1.4201 + 0.000047(\text{GPS Speed}) + 0.000923(\text{Lat Accel}) \\
& + 0.000499(\text{Long Accel}) + 0.0106(\text{Thort Pos}) + 0.0198(\text{Yaw Rate}) \\
& + 0.000099(\text{GPS Speed} * \text{Long Accel}) + -2.62\text{E}{-6}(\text{GPS Speed} * \text{Thort Pos}) \\
& + -0.00005(\text{GPS Speed} * \text{Yaw Rate}) + -0.00128(\text{Lat Accel} * \text{Yaw Rate}) \\
& + -0.00177(\text{Thort Pos} * \text{Yaw Rate}).
\end{aligned} \tag{4}
$$

The Likelihood Ratio Score for this test resulted in a *p*-value of <0.0001, which meant the results output by this test were better than chance. However with the incredibly small R-square value equal to 0.0097 and the Hosmer and Lemeshow *p*-value at <0.0001, the dependent variables do not account for the majority of the variance in the model and the predicted values cannot be said to statistically match those observed. This model was also not a great fit of the data. Table 4 displays the Control versus Group 3 MLR Results.

It is common practice for researchers to test or validate the results of a newly developed model. However, since all three models were proven to have such weak predictive power there was no need to validate the equations. It is recommended that further study be conducted to produce stronger models and validation be tested on those.

5 Conclusions

The objective of this research was to use the SHRP 2 NDS dataset to develop distracted driving prediction models. GPS speed, lateral and longitudinal acceleration, throttle position and yaw rate were the driving performance measures tasked with predicting whether drivers were engaged in one of three defined groups of secondary tasks: talking or listening on hand-held phone, texting or dialing on hand-held phone or interacting with the adjacent passenger. The time series nature of the data used provided more robust data than typically used in distracted driving studies. Therefore, the input information used to develop the prediction models was of a very high quality. It combined the beneficial attributes of using time series data, and the more realistic view of driver behavior that is acquired by using the naturalistic method of data collection.

Chi-square tests initially run to compare the Control group to the secondary tasks groups indicated there were differences in the driving performance measures when the driver was engaged in a secondary task. Multiple logistic regression (MLR) was utilized to determine if those differences present could be used to develop models that adequately predict when drivers were engaged in the three secondary tasks of interest. The results indicated this data was unable to accomplish that goal.

Future work should focus on comparing these results to prediction models developed using an alternative to the multiple logistical regression method. If researchers are able to develop models that have improved R-square and Hosmer and Lemeshow Test *p*-values, the potential next step would be the development of a distraction index capable of ranking the impact of each distracting effect.

Acknowledgments This research was funded by the American Association of University Women, National GEM Consortium, Louisiana Transportation Research Center and the Federal Highway Administration through the Dwight D. Eisenhower Transportation GRAD Fellowship. The contents of the paper are the sole responsibility of the authors.

References

1. National Highway Traffic Safety Administration. Traffic Safety Facts. Research Note DOT HS 812012. NHTSA, U.S. Department of Transportation (2014)
2. Second Strategic Highway Research Program. The SHRP 2 Naturalistic Driving Study. Transportation Research Board. http://www.trb.org/StrategicHighwayResearchProgram2SHRP2/Pages/The-SHRP-2-Naturalistic-Driving-Study-472.aspx. Accessed 24 April 2014
3. Second Strategic Highway Research Program. SHRP 2 Publications. Transportation Research Board. http://www.trb.org/StrategicHighwayResearchProgram2SHRP2/Pages/The-SHRP-2-Naturalistic-Driving-Study-472.aspx#publications. Accessed 25 April 2014
4. Kaptein, N.A., Theeuwes, J., Van Der Horst, R.: Driving simulator validity: some considerations. Transp. Res. Rec. J. Transp. Res. Board **1550**, 30–36 (1996)

5. Bach, K.M., Jæger, M.G., Skov, M.B., Thomassen, N.G.: Interacting with in vehicle systems: understanding, measuring, and evaluating attention. Presented at the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology (2009)
6. Virginia Tech Transportation Institute: 100-Car naturalistic driving study, December 12, 2013. Vulnerable Road User Safety. https://www.vtti.vt.edu/research/vrus/projects/100car/100car.html. Accessed 24 Apr 2014
7. Virginia Tech Transportation Institute: 100-car naturalistic driving study fact sheet, 2005. Virginia Tech News. http://www.vtnews.vt.edu/articles/2005/06/2005-834.html. Accessed 23 July 2015
8. Jonasson, J.K., Rootzén, H.: Internal validation of near-crashes in naturalistic driving studies: a continuous and multivariate approach. Accid. Anal. Prev. **62**, 102–109 (2014)
9. Bagdadi, O.: Assessing safety critical braking events in naturalistic driving studies. Transp. Res. Part F Traffic Psychol. Behav. **16**, 117–126 (2013)
10. Sangster, J., Rakha, H., Du, J.: Application of naturalistic driving data to modeling of driver car-following behavior. Transp. Res. Record J. Transp. Res. Board **2390**, 20–33 (2013)
11. Wong, J.T., Huang, S.H.: Attention allocation patterns in naturalistic driving. Accid. Anal. Prev. **62**, 102–109 (2013)
12. Guo, F., Fang, Y.: Individual driver risk assessment using naturalistic driving data. Accid. Anal. Prev. **61**, 3–9 (2013)
13. Sayer, J.R., Devonshire, J.M., Flanagan, C.A.: Naturalistic driving performance during secondary tasks. In: Proceedings of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 224–230 (2007)
14. Codjoe, J., McCarter K., Ishak, I.: Multivariate analysis of distracted driving. Presented at the 94th Annual Meeting of the Transportation Research Board, Washington, D.C. (2015)
15. Meyers, L., Gamst, G., Guarino, A.: Data Analysis Using SAS Enterprise Guide. Cambridge University Press, Cambridge (2009)
16. Karp, A.: Getting started with PROC LOGISTIC

Socializing Under the Influence of Distracted Driving: A Study of the Effects of in-Vehicle and Outside-of-the-Vehicle Communication While Driving

Hanan Alnizami, Ignacio Alvarez and Juan E. Gilbert

Abstract Advancements of in-vehicle technologies and the development of mobile applications that keep a driver connected in a driving environment have caused an increasingly dangerous safety concern. Distracted driving has gained the attention of legislators and governments globally. Countries have constituted bans that partially or fully forbid drivers from using gadgets while driving, especially hindering out-of-the-vehicle communications. This paper introduces Voiceing™, a voice-activated application meant to improve social communications in the car, serving as a safe alternative to distracted driving. Other modalities of interaction such as texting, in-vehicle conversations and outside-of-the-vehicle conversation have been measured and compared with Voiceing™ investigating effects on driver's performance, cognitive load and user acceptance. Results from this study suggest that Voiceing™ is a safer alternative than in-vehicle interactions with humans. Results also show that natural speech interaction of in-vehicle applications and the inclusion of context awareness help improve driving performance while interacting with a vehicle system.

Keywords Measurement · Performance · Design · Experimentation · Human factors · Distracted driving · Social car · In-vehicle distraction · Vehicle communication

H. Alnizami (✉) · I. Alvarez
Intel Corporation, 2111 NE 25th Avenue, Hillsboro, OR 97124, USA
e-mail: hanan.alnizami@intel.com

I. Alvarez
e-mail: ignacio.j.alvarez@intel.com

J.E. Gilbert
University of Florida, P.O. Box 116120, Gainesville, FL 32611, USA
e-mail: juan@ufl.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_21

1 Introduction

The use of mobile phones in cars has risen significantly in the last decade. This rise is attributed to the need for drivers to feel connected to their outside world [1]. Vehicle information exchange takes place inside of the vehicle, typically in the form of conversations, and outside the vehicle enabled by vehicle telematics. In both case, speech is utilized as the main channel of these interactions. When driving with passengers, a driver engages in conversations that establish a social communication in the vehicle. When driving alone, research has also found that drivers have the urge to use their mobile phones while driving [2], performing socializing activities such as calling, texting or checking email and social networks. Vehicle manufacturers nowadays are continuously integrating features that help drivers stay socially connected to the outside world. Studies show that use of these features increases the driver's cognitive load impairing driving performance [3] causing slow reaction time, poor judgment, and poor lane maintenance, due to reduced lateral and longitudinal vehicle controls and reduced critical event detection [4–7]. Driver's distraction is defined as the diversion of a driver's attention away from activities that are critical for safe driving. Such distraction is caused by a competing activity [8], object, or event within the internal or the external vehicle environment [9]. More than 34 US states, and over 52 countries have passed distracted driving legislations banning the use of mobile devices while driving. But banning the direct use of mobile devices while driving has only caused further distraction and increased dangers. Drivers that use their mobile phones while driving take extra handheld operation measures for fear of being caught, increasing the chances of an accident. The Federal Communications Commission (FCC) is working to recognize and help develop technologies that could potentially reduce driver's distraction [10].

This paper introduces Voiceing™, a voice activated application that utilizes speech to send/receive voice messages and send emails while driving to fill the void in vehicular social communications safely. This research investigates four different conditions of socializing inside the vehicle vs. outside of the vehicle while driving a simulator: (1) using Voiceing™ to send and receive voice messages, (2) interacting with a passenger, which represents an in-vehicle human interaction condition, (3) interacting with a familiar person over the phone, which represents an—outside of the vehicle—human interaction condition, (4) interacting with an unfamiliar person at a call center, which represents an interaction with a trained professional. (5) Texting while driving. The rest of this paper presents previous work on distracted driving studies, introduces the Voiceing™ technology, explains the methods in which driving distractions were measured and the tools used to measure them, and finally describes the results comparing the above-mentioned modalities.

2 Distracted Driving

Driver distraction occurs when a driver is delayed from the recognition of information that is necessary to safely accomplish the driving task [11]. The "100 car naturalistic study" found that secondary tasks were performed during 40 % of all trips [11]. Many drivers find it difficult to resist the temptation of using mobile devices to stay connected while driving. Text messaging, for example, has been found to affect lateral control and reaction time, lowering it up to 35 % [12]. In parallel, the amount of attention given to the use of In-vehicle Infotainment Systems (IVISs) has peaked considerably and raised safety concerns [13] and studies have demonstrated that the use of IVISs contributed to 25–30 % of crash risk [14].

The cause of such distractions relates to the fact that humans have a limited amount of available cognitive resources, according to the Dual-Task Paradigm [15]. In driving conditions a great deal of those resources are allocated for the primary driving task, leaving little capacity to secondary actions. However, the multiple resource theory of attention has shown that different pools of resources could be used in parallel, allowing to perform multitasking if actions performed are being allocated in different modalities [16]. This principle supports the design of hands-free, eyes-free vehicular interfaces. Many present auditory interfaces are preferred media for in-vehicle warnings [17] or menu navigation [18]. Driver distraction effects of utilizing them have been proven lower, comparing voice interfaces to manual interfaces [19]. This paper analyzes the effect of using Voiceing™ on driving performance as opposed to the effects of talking to a person inside of the vehicle(passenger), talking to a person outside the vehicle (call center), and texting while driving.

3 Voiceing™

Since one of the speech conditions investigated in this research is voice messages, this section explain Voiceing™, the voice user interface benchmarked in this study. Voiceing™ is a hands-free, eyes-free voice activated system developed by the Human-Centered Computing lab at Clemson University. It allows for sending and receiving short messages using speech. The uniqueness of this tool lies in its independence of being exclusive to any mobile platform. It could be used with any telephone number, mobile or landline. Users of Voiceing™ could send voice messages and receive replies from non Voiceing™ users. The system allows the user to operate diverse features such as using voice to send emails, short auditory voice messages, setting reminders, and making calls. In order to interact with it, a user would perform a call to their unique Voiceing™ number and would interact with the speech interface accordingly. Once a user creates an account, s/he is able to set the delivery of voice replies one or multiple delivery modalities according to the user's preference. Participants used a dummy account created for the purpose of this study. The interaction with the system begins when the Voiceing™ systems prompts the participant with "how may I help you?". The participant then says "send a text to John Doe". The system prompts them to record their message, verifies it with the participant and then sends it to out to the contact(s). When the user receives voice messages, the system calls the participant. When the call is answered, the system prompts "you have a message from John Doe" then will go forth with playing the voice recording of the message.

4 Experiment Design

This study examines the effects of various social communication methods on driving performance and distraction in a driving scenario on a simulator. Different conditions were investigated and were compared to a single driving condition as a baseline, and a texting while driving condition, as it has already been proven to be a highly distracting action. The conditions were: (1) using Voiceing™ to send and receive voice messages, (2) interacting with a vehicle passenger, (3) interacting with a familiar person over the phone, and (4) interacting with an unfamiliar person at a call center.

4.1 Apparatus and Study Logistics

The driving simulator operated on a desktop PC, running Windows 7. The driving controls used in the study consisted of (1) a Logitech G27 dual-motor force feedback racing wheel mounted on a table; (2) an HD LCD 40 in. screen positioned in-line with the racing wheel and directly in front of the participant, and (3) steel gas, brake, and clutch foot pedals placed on the ground under the table. Simulator was set up for automatic gear. An adjustable desk chair was used to provide different height variations according to participants' liking. A second desktop PC running Windows 7 was used in connection with a Sony WCS-999 wireless microphone system that the participant wore during the drive. Skype version 4.2, a VoIP software application, was used to place telephonic calls. Figure 1 depicts the experiment setup. The software used to collect performance data was the Lane Change Task simulation and analysis software (LCT) [20].

4.2 Demographics

Participants were recruited from university campus and were required to be licensed drivers. A total of 127 participants, 53 females and 74 males, took part in this study. Participants ranged in age from 20 to 59 and the mean age of the sample was

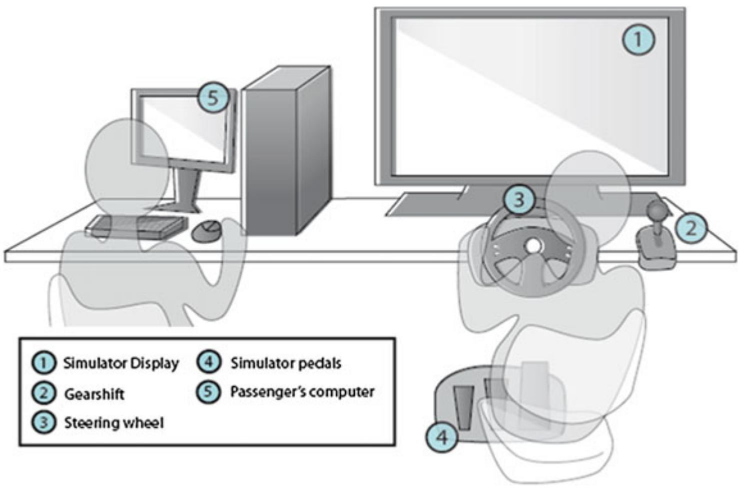


Fig. 1 Experiment setup

25.1 years old, *SD* = 10.19. Only one percent of the participants suffered from disabilities, but this did not affect the participant's driving skills. All participants owned a driver's license with a mean driving experience of 7.62 years. Participation in the study was voluntary. Subjects were rewarded $10 for their participation.

4.3 Design

The experiment used a single factor, repeated measures within-subjects design. The independent variable examined was the method of communication as a secondary task while driving the simulator. Driver distraction was measured as a combination of driving performance measures and a subjective cognitive load assessment. User experience measures were also collected in the post-experiment survey. All participants were fluent in English; 70 % of them were native speakers and 30 % spoke English as a second language.

4.4 Secondary Task Conditions

This study analyzed the interaction of a driver in four different conditions, each handling a different modality of interaction while driving. All conditions were hands free. Participants were asked to press a button on the steering-wheel in order to engage with a person or a system. The four conditions were:

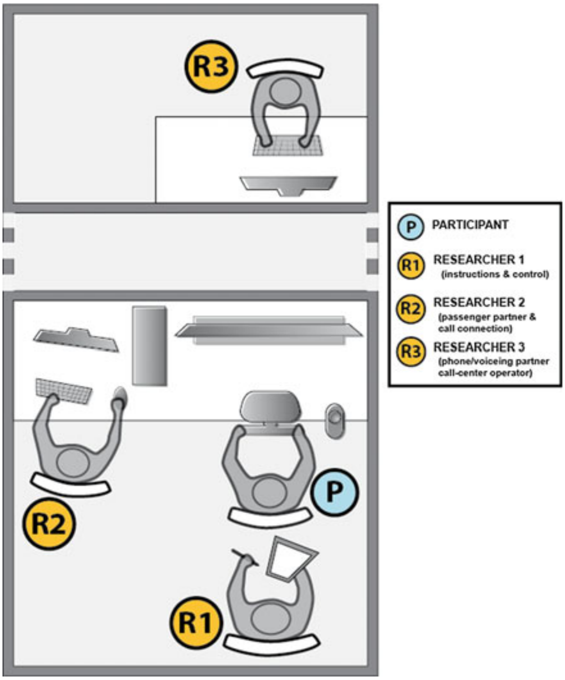


Fig. 2 Researchers and participant placement

1. Interacting with Voiceing™: Participants were asked to interact with a pre-determined contact, R3 in Fig. 2, an experimenter not present in the same room, via sending and receiving voice messages. Participants called the Voiceing™ system by pressing a button on the steering wheel and interact with the system, sending and receiving voice messages until they reached the end of the track.
2. Interacting with a familiar person on the phone: Participants were instructed to call, and have a friendly conversation with, Michael Jackson for the length of a track. We used the name Michael Jackson for ease of remembering. Researcher R3 in Fig. 2 played Michael Jackson. Participants were instructed to discuss any topic that came to mind. R3 stimulated the conversation by asking further questions such as “what did you eat this morning” to keep the participant engaged in the conversation.
3. Interacting with a human passenger: Participant was presented with a problem scenario and were asked to collaborate with their passenger to find a solution. Researcher R2 in Fig. 2 played the passenger and sat next to the driver. R2 used a web-based system to look up answers for participant’s specific questions. When presented with the problem, it was up to the participant to explain the problem to R2 and seek help solving the problem.

4. Interacting with a call center operator: Participant was presented with a problem scenario and was asked to consult a call center operator for information on how to solve the problem at hand. The participant initiated the call by pressing the button on the steering wheel. As shown in Fig. 2, the call center operator, R3, was able to provide the driver with road assistance information according to the questions asked.
5. Texting while driving: Participant was asked to text a continuous series of text messages while performing the LCT task using their own mobile phone. The messages were sent to R3 who engaged in a text message conversation with the driver.

5 Procedure

First, participants were asked to consent to participating in the study. Once obtained, participants were provided with information about the study and asked to complete a pre-survey that collected demographic data, as well as previous experience operating the simulator, using voice technologies, and texting habits. Next, participants practiced driving the simulator. A practice session involved driving three consecutive tracks on the LCT simulator to get familiar with how to operate it. When completed, an interaction task was randomly assigned to each participant to counterbalance the conditions. All conditions required participants to drive maintaining the lane changing and maneuvering task, while they underwent the requirements of each condition.

An example of a problem scenario presented in conditions passenger and call center is “This is your first time operating the vehicle and you don’t know how to place a call. Consult the passenger”. While driving the simulator, the participant asked the partners (passenger or call center agent) questions like “how do I make a phone call”. The interlocutor then queried web-based search engine for the correct answer and read it back to the participant. If the question had no answer, then the participant would be asked to rephrase the question. This process was repeated with various scenarios until the end of the track was reached.

After all tasks were completed, the participant was asked to complete a post survey questionnaire. NASA–TLX is a multi-dimensional 20-point Likert scale that measures six workload dimensions (mental demand, physical demand, time demand, performance, effort, and frustration level). The participant also completed a usability questionnaire using a modified version of the technology acceptance model, TAM [21].

6 Results

6.1 Driver Performance Results

A within-subject baseline was computed in the training phase and each secondary task was compared against the individual baseline to obtain driver performance metrics. All performance metrics were extracted running the raw LCT data files trough a data mining script in NI DIAdem [22]. Metrics collected following this procedure included mean lateral deviation, standard deviation (SD) of the mean lateral deviation, reaction time, and wrong lane changes.

Results on mean lateral deviation, shown in Fig. 3, present lower mean values for the baseline and driving alone than the secondary tasks.

One-Way analysis of variance (ANOVA) was performed to compare the differences. The results, ($p = 0.000$), proved significant differences across secondary tasks. Bonferroni test for pairwise comparison found significant differences between baseline and texting tasks ($p = 0.007$) as well as between baseline and Call Center tasks ($p = 0.001$). No significant difference was found between the baseline and Voiceing.

Figure 4 shows that the mean values of lateral standard deviation (SD) were lowest during the use of Voiceing™ and surprisingly during the texting task. Participants had a tighter control of the steering wheel than during the baseline. This results suggest that during these two tasks participants where focusing on the interaction and they performed a tighter control on the wheel. The phone and

Fig. 3 Mean lateral deviation on the LCT (error bars represent 95 % confidence interval)

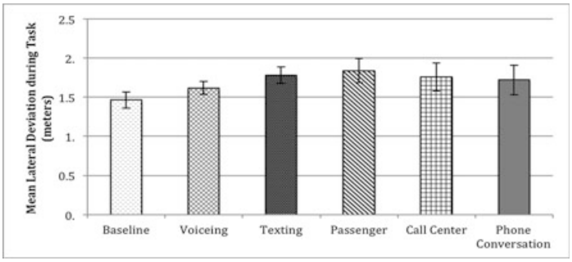
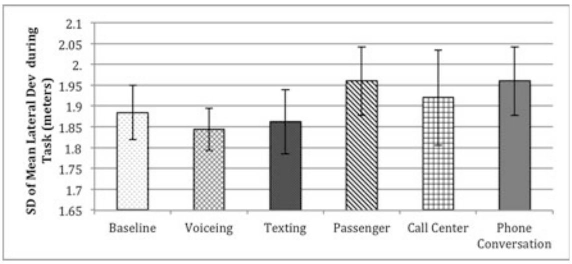


Fig. 4 SD of lateral deviation on the LCT (error bars represent 95 % confidence interval)



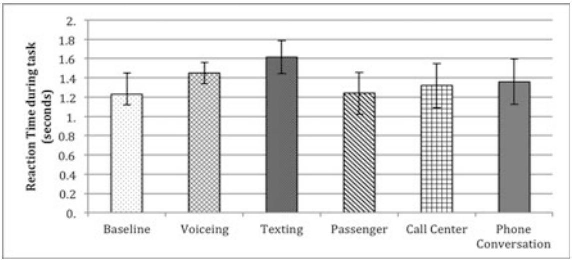


Fig. 5 Reaction times on the LCT (error bars represent 95 % confidence interval)

passenger conversations resulted in the highest SD values suggesting looser control of the vehicle during these tasks. ANOVA analysis reported however no statistical significance between tasks.

Reaction time in the LCT is calculated from the time the sign is legible until the time the user started the lane change maneuver. Given that the point of recognition for the LCT is known, the reaction time was calculated using steering angle measures. Figure 5 displays the mean reaction time values for each task. Fastest reaction times were achieved during the single drive baseline and increased as participants performed secondary actions. ANOVA test comparing mean reaction times showed significance between tasks. Following post hoc test showed only significance on texting while driving ($p = 0.038$). The reaction time during conversation with the passenger was shorter than during phone conversations or Voiceing™ usage.

Extremely illustrative behaviors were found in the wrong lane changes count for each task, as can be seen in Fig. 6. A wrong lane change means the participant missed the sign. Results of the mean values show that there was, on average, close to zero mistakes during the single task drive. The number of wrong lane changes increased an average of 40 % performing secondary tasks like conversations with a passenger or the operator of a call center. However, while performing secondary actions that required hand and eye attention, such as texting, the number of wrong lane changes increased dramatically. Kruskal-Wallis ANOVA showed very high significance, $p = 0.000$, across tasks. The Tukey-Kramer test showed significant differences between the baseline and texting ($p = 0.000$). Texting also presented significant differences compared to Voiceing™ and the Call Center conditions.

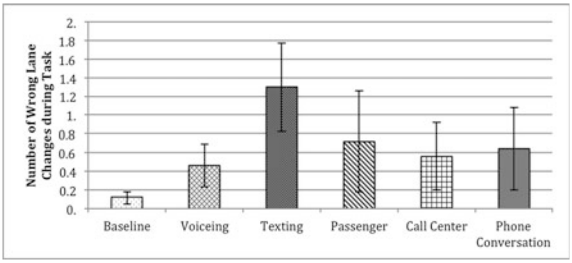


Fig. 6 Wrong lane changes on the LCT (error bars represent 95 % confidence interval)

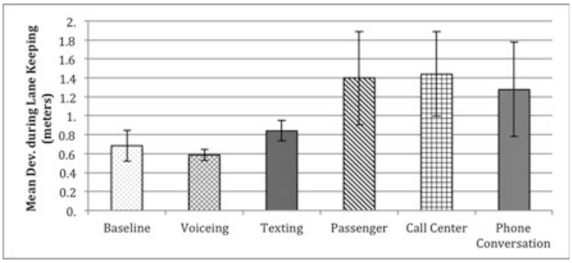


Fig. 7 Mean deviation during lane keeping (error bars represent 95 % confidence interval)

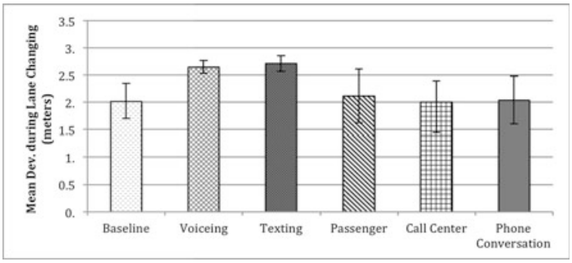


Fig. 8 Mean deviation during lane changing (error bars represent 95 % confidence interval)

To further investigate the effects across modalities, the two driving actions of which the LCT is composed, lane keeping and lane changing, were analyzed separately. The effects of the different secondary actions were studied during these phases. Results for the mean lateral deviation are presented in Figs. 7 and 8. During the lane keeping phases, the driver had to concentrate on maintaining the vehicle in the center of the lane. Results, Fig. 7, showed that mean deviations when the driver was interacting with people were the highest. The values for lane deviation were lowest for the Voiceing™ task, suggesting that participants kept better control than even the single drive baseline. The results were surprisingly similar for the texting condition. Interacting with human partners revealed to be more distracting during the lane-keeping task, based on the lateral deviation values. Statistical analysis reported very high significance in group comparison, $p = 0.000$, and post hoc tests showed significant differences comparing the baseline to all but for the Voiceing™ task. Also, the conversational tasks involving human partners showed high statistical significance compared to Voiceing. The results of analyzing SD during lane keeping showed the same results as the mean deviation values. On the lane changing phases, however, the results were opposite to the lane keeping, see Fig. 8. The secondary tasks that involved interaction with people showed better performance than those that implied interacting with technology. Furthermore, results for phone conversation and the call center showed a slight increment in vehicle control during lane changes compared to the baseline. Post hoc pairwise comparisons reported significant differences between the baseline and Voiceing, $p < 0.05$, and very high significance, $p = 0.001$, between baseline and texting. Likewise, significance was found between the human-to-human modalities, texting and Voiceing.

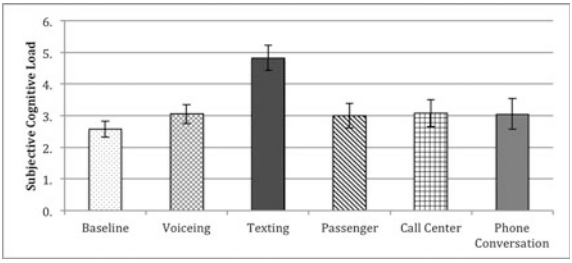


Fig. 9 Average TLX ratings for subjective cognitive load (error bars represent 95 % confidence interval)

6.2 Cognitive Load Results

After performing the secondary driving tasks, participants were asked to rate their cognitive workload using the NASA-TLX questionnaire [23]. The values of the overall cognitive workload were calculated giving equal weight to the six workload dimensions present in the TLX-questionnaire. Figure 9 shows the mean values per task on a normalized Likert scale from 0, lowest, to 7, highest.

Subjective cognitive level reported by the participants during the baseline drive was lower than when asked to perform a secondary task. The most demanding task was clearly texting and the rest of the conditions presented similar mean values. Statistical analysis followed to determine the significance of the mean value differences. The Tukey Kramer procedure revealed pairwise very significant differences, $p < 0.001$, between texting while driving and the baseline, but also between texting and all other tasks. While effort, temporal, physical and mental demand showed similar results to the overall measure, performance and frustration levels clearly benefitted the Voiceing™ task above the rest. The highest performance score was reported for the Voiceing™ task, where participants believed to drive better than during the baseline task, Fig. 10.

In an inverse correlation, the frustration levels reported during the different conditions were lowest for the baseline and Voiceing™ tasks, Fig. 11. The highest frustration was experienced while participants were chatting on the phone which suggests that participants were aware of the difficulty to perform accurately on the LCT task while they were on the phone.

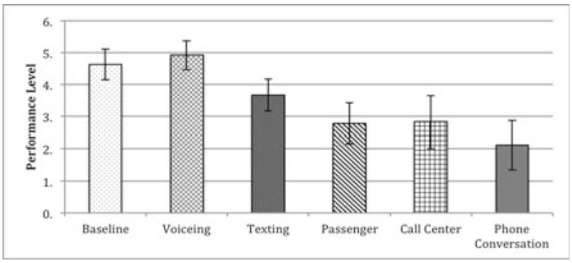


Fig. 10 Subjective performance level reported on the TLX questionnaire

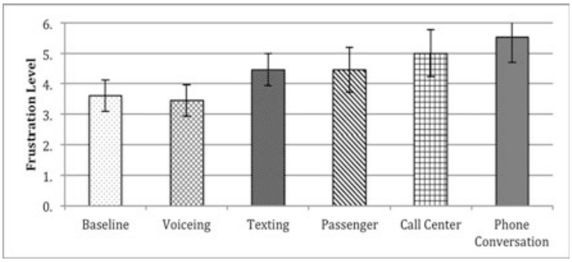


Fig. 11 Subjective frustration level reported on the TLX

7 Conclusion

The results presented in this paper suggest that in order to fulfill the socializing needs of drivers, voice-activated applications have the least effect on the driver in terms of driving performance. Even compared with human conversations in the vehicle and through telephony lines, we see results that are favorable towards Voiceing™. The driving performance metrics were also supported by the behavioral intention of use of the participants, the majority of which rated Voiceing™ as a preferred communication media, especially when compared to texting while driving. These results suggest that the social activity in the vehicle is inherently spoken even though many of the participants were very skilled at texting. The self-reported performance measures indicate as well that participants felt comfortable using the Voiceing™ application.

All in all, these results suggest that speech technology is the most promising interaction method in the vehicle to enable the social act. A driver-centric design for in-vehicle communication systems has the potential to lower driver distraction. The implementation of such a system like Voiceing™ improves communication without significantly impacting driver cognitive load or driver distraction. We hope to use these results when designing other socially-inspired communication devices and other in-vehicle social services. Future improvements on natural speech interaction of the voice user interface and inclusion of context awareness in the application could potentially help further palliate the effects of interacting with a vehicle system while satisfying the communication needs of drivers.

References

1. Dimmick, J., Feaster, J.C., Ramirez, A.: The niches of interpersonal media: relationships in time and space. New Media Soc. **13**(8), 1265–1282 (2011)
2. Sohn, T., et al.: A diary study of mobile information needs in computer-human interactions CHI'08 (2008)
3. Bruyas, M., et al.: Does making a conversation asynchronous reduce the negative impact of phone call on driving? French National Institute for Transport and Safety Research (2008)
4. Alm, H., Nilsson, L.: The effects of a mobile telephone task on driver behaviour in a car following situation. Accid. Anal. Prev. **27**(5), 707–715 (1995)

5. Horrey, W.J., Lesch, M.F., Garabet, A.: Assessing the awareness of performance decrements in distracted drivers. Accid. Anal. Prev. **40**(2), 675–682 (2008)
6. Ranney, T.A., Harbluk, J.L., Noy, Y.: Effects of voice technology on test track driving performance: implications for driver distractions. Hum. Factors J. Hum. Factors Ergon. Soc. **47**(2), 439–454 (2005)
7. McKnight, A.J., McKnight, A.S.: The effect of cellular phone use upon driver attention. Accid. Anal. Prev. **25**(3), 259–265 (1993)
8. Lee, J.D., Regan, M.A., Young, K.L.: Defining driver distraction. CRC Press, Boca Raton FL (2008)
9. Stutts, J.C.: The role of driver distraction in traffic crashes. AAA Foundation for Traffic Safety, Washington DC (2001)
10. FCC: Texting while driving guidelines. http://www.fcc.gov/guides/texting-while-driving
11. Scherer, K.R.: Vocal markers of emotion: comparing induction and acting elicitation. Comput. Speech Lang. (2011)
12. Roesch, E., Bänziger, T., Scherer, K.: Preliminary plans for exemplars: theory. Hum. Netw. Excelence (2004)
13. Pitterman, J., Pitterman, A., Minker, W.: Handling emotions in human-computer dialogues. Springer, Ulm, Germany (2010)
14. Merriam Webster Online Dictionary. http://www.merriam-webster.com/dictionary/emotion
15. Goselin, P.A., Gagné, J.: Use of a dual-task paradigm for measure listening effort. Can. J. Speech Lang. Pathol. Audiol. **3**(2), 159–177 (2010)
16. Eyben, F., et al.: Emotion on the road; necessity, acceptance, and feasibility of affective computing in the car. Adv. Hum Comput. Interac. (2010)
17. Scherer, K.R.: Psychological models of emotion. In: Borod, J.C. (ed.) The Neuropsychology of Emotion, pp. 137–162. Oxford University Press, New York (2000)
18. Jeon, M., et al.: Enhanced auditory menu cues improve dual task performance and are preferred with in-vehicle technologies. In: Proceedings of the First International Conference on Automotive User Interfaces and Interactive Vehicular Applications, ACM, Essen, Germany. pp. 91–98 (2009)
19. Scherer, K.R.: Vocal markers of emotion: comparing induction and acting elicitation. Comput. Speech Lang. (2011)
20. Plutchik, R.: The nature of emotions. Am. Sci. **89**, 344–350 (2001)
21. Venkatesh, V., et al.: User acceptance of information technology: towards a unified view. MIS Quaterly **27**(3), 425–478 (2003)
22. McDonald, R., et al.: Structured models for fine-to-coarse sentiment analysis. In: 45th Annual Meeting of the Association of Computational Linguistics. Association of Computational Linguistics, Prague, Czech Republic (2007)
23. Picard, R.W.: Affective Computing. MIT Press, Cambridge, MA (1997)

Risk Factors for Driver Distraction and Inattention in Tram Drivers

Anjum Naweed, Janette Rose, Sangeeta Singh and Damien Kook

Abstract Tram driving is a complex task requiring high levels of workload, route knowledge, and attention. Metropolitan tram networks typically contain many routes and share roads with other road users. Collision potential is highest at road intersections and areas where the track runs along the road with no segregation and when collisions occur they can cause serious injury and disruption. A study was conducted on an Australian tram network to identify collision risk factors. The approach included focus groups and discussions with 22 drivers, and observations at two high-risk locations. Data were coded thematically using a recently published taxonomy for driver distraction and inattention. The majority of factors fell into the Driver Cursory Attention category, with a large representation also in the Misprioritised and Neglected Attention categories and instances of Diverted Attention were mainly driving-related. Findings are discussed in terms of potential mitigation strategies and their implications for further refinements to driver distraction and inattention taxonomies.

Keywords Driver distraction · Inattention · Tram driving · Risk perception · Road safety · Rail safety

A. Naweed (✉) · J. Rose
Appleton Institute for Behavioural Science, Central Queensland University,
44 Greenhill Rd, Wayville, SA 5034, Australia
e-mail: Anjum.Naweed@cqu.edu.au

J. Rose
e-mail: Jan.Rose@cqu.edu.au

S. Singh · D. Kook
Keolis Downer Rail (Yarra Trams), GPO Box 5231, Melbourne,
VIC 3001, Australia
e-mail: Sangeeta.Singh@yarratrams.com.au

D. Kook
e-mail: Damien.Kook@yarratrams.com.au

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_22

1 Introduction

A tram driver needs to drive safely and efficiently while also providing a high level of customer service. Good tram driving not only requires an awareness of the various elements in the environment but also the ability to understand its complexity and be prepared for unexpected events. In most tram networks, rail tracks run over roads (i.e. streetcars) as well as dedicated tramways, creating a mixed-traffic environment with complex signalling that is traversed using both train dynamics and road navigation tactics [1]. For this reason, the demands on tram drivers may be considered similar to those of motor vehicle drivers and train drivers combined. Not only do drivers need to be skilled in handling a variety of trams, all of which have different handling characteristics, they also need to be highly alert and vigilant at all times.

The environment is highly dynamic and heavily populated with passengers, pedestrians, motor vehicles and other road users demanding constant attention, and often requires actions necessary to avoid accidents. Thus a high level of situation awareness is vital at all times. As defined by Endsley [2], this includes perceiving and comprehending appropriately, and projecting ahead. For tram driving, perception includes being aware of the location, destination, required positioning of points along the route, train handling characteristics specific to the tram type, current speed, speed limits, and so on. A tram driver's comprehension of the situation must be constantly updated with the evolving environment, and being able to project ahead plays a major role in relation to predicting the behaviour of other trams drivers, pedestrians and other road users. Experience may improve this skill, especially in tasks where there is an element of repetition such as tram driving (e.g. travelling the same route, approaching stations, following other trams) on a daily basis.

Because of the nature of the work and job design, the tram-driving role is also impacted by specific rail human factors concerns such as fatigue, shift-work, and reaction time. According to Williamson [3], the relationship between fatigue and distraction in motorists may be two-way, with both positive and negative effects. There is some evidence to suggest that fatigued drivers may be less prone to distraction than non-fatigued drivers, and that certain forms of distraction may assist in maintaining alertness [3]. However, both fatigue and distraction can be detrimental to safety and performance, thus a combination of the two in a task such as tram driving may increase the risk of accidents.

With such a high demand on attention, it comes as no surprise that accidents and near misses occur due to distraction and inattention. With distraction estimated to be the cause of between 10 and 23 % of motor vehicle crashes [4], similar percentages are likely to apply to tram collisions. Although accidents are not a frequent occurrence on metropolitan tram networks, when they do occur they may cause major disruption to service and serious injury [5]. Despite this, there is a dearth of research into causes of tram collisions, and in particular driver distraction and inattention, with most published literature on Australian trams having an engineering focus on the history, growth, design of trams and tram networks rather than the elements of driving [1].

In this study, potential human factors causes of tram collisions on an Australian metropolitan tram network were investigated as part of a broader examination of the types of incidents that occur on the network, and analysis was undertaken using a recently published taxonomy for driver distraction and inattention developed by Regan et al. [6]. The taxonomy divides driver inattention into five categories based on the different mechanisms by which the inattention may arise, and while the taxonomy is for drivers of motor vehicles, it may well be suitable for other similar tasks. Here the taxonomy is applied to tram driving to determine its suitability for that task, and the paper investigates risk factors for distraction and inattention in tram drivers. This includes not only accidents that actually occurred but also situations in which there is a high risk of accidents occurring.

1.1 Aims and Objectives

The aims of the paper are to determine the extent to which distraction plays a part in tram-on-tram collisions, and to contribute more on this topic to an area of driving that is under-researched. A secondary objective is to use a recent driver distraction taxonomy [6], and determine how well it applies in the tram-driving context, and whether it would be a useful tool for determining strategies to mitigate risks.

2 Methodology

Figure 1 illustrates the stages of research for a study.

A mixed-methods design was used which included:

- A review of industry documentation and accident reports.
- On-site observations at two complex locations with conflicting moves, including observations from the passenger area of the tram and driver's cabin.

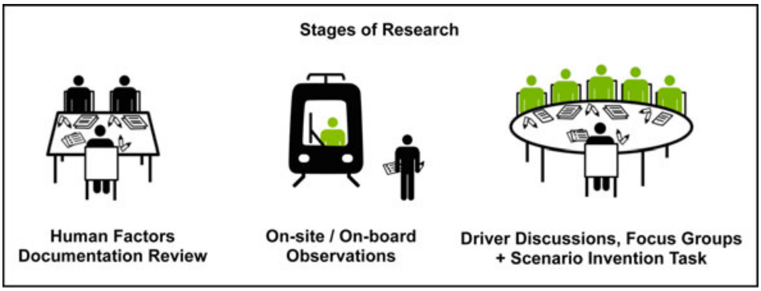


Fig. 1 Stages of research used to process and collect data in the study

- Discussions with individual drivers at three depots, and focus group discussions at two depots, including drivers, trainers, team managers, and other management staff.

The methodology comprised an early stage of accident report review, both to gain familiarity with the domain and to scope out the requirements for the rest of the methodology. A combination of observations, focus groups and direct discussions were adopted to gather representative data. Two locations were chosen for the main focus of this investigation, based on history of previous accidents; one major intersection where road meets rail and where there are multiple road users; and one junction that is separated from the road but positioned next to a major intersection that affects tram activity within and approaching the junction. These two locations were selected based on their representation of the complexities found throughout the network.

Observations were conducted from various vantage points around each location as well as aboard trams travelling through each intersection, from every point of entry. In order to ensure a comprehensive picture of the intricacies of negotiating these intersections, observations included: tram activity/frequency during peak and non-peak times; frequency of potential conflicting situations; pedestrian activity in and around the intersections; activity and behaviour of other road users including motor vehicle drivers, cyclists, and horse-drawn carriages; and passenger behaviour, including those travelling in trams, waiting at stations, and boarding or alighting the tram.

In the two focus groups, a scenario invention task [7] was conducted where each driver was instructed to draw a diagram of the first location. This task has been designed as an effective means of externalising knowledge (drawing out knowledge that is held subconsciously) that can be very difficult to recall and talk about. Using various coloured pens, each driver then created a scenario that would create a high risk of collision. These scenarios could be situations the drivers had experienced themselves, situations they had heard about from other drivers, or completely fictional scenarios that could potentially occur. When all drivers had drawn their diagrams and scenarios for the first intersection, they described their scenarios to the group in turn and discussions were held between all participants and the researchers. Drivers were then asked to write down any potential technical and non-technical solutions that they believed could be implemented to reduce the risk of collisions. The same procedure was followed for the second location. Figure 2 shows an example diagram from the task.

In the individual driver discussions, participants discussed concerns they had in relation to the two locations, and described the strategies they used to negotiate those locations safely. These discussions also included other more general concerns such as strategies used throughout the network to ensure safety and on time running.

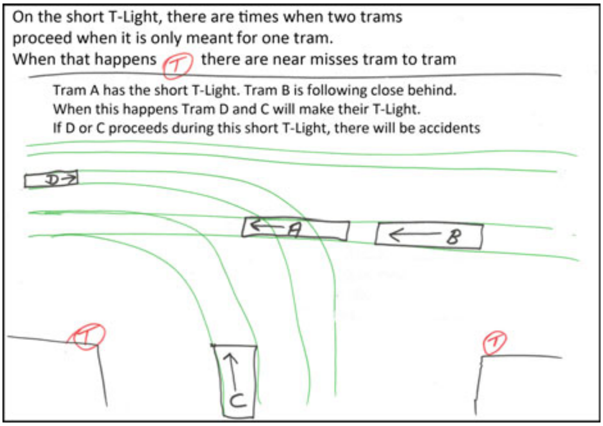


Fig. 2 Stages of research used to process and collect data in the study

2.1 Participants

A total of 23 participants were involved in the study (22 Male, 1 Female). Nine of these participated in the focus groups, and 14 took part in individual driver discussions. The age of the participants ranged from 25 to 67 years, with an average age of 50 (standard deviation = 8.89 years). Tram driving experience ranged from 1.5 to 21 years, with an average of 10 years (standard deviation = 6.47 years). Thirteen of the participants possessed more than 10 years driving experience.

2.2 Data Capture and Analysis

Data was captured via voice recordings for the focus group and individual interviews. Written notes and researcher voice recording were used for observations on-site and aboard trams. Photographs were taken from numerous perspectives around the study sites. A review of the accident reports was conducted, from which key themes and categories were derived. Voice recordings were transcribed. Thematic analysis was used to classify and codify transcript data according to distraction and inattention types [8]. The codes and were applied deductively during the process, and based on the five categories of inattention and corresponding definitions from Regan et al.'s taxonomy [6], with the fifth category having two sub-categories as described below:

- *Driver Restricted Attention (DRA)*—Insufficient or no attention to activities critical for safe driving brought about by something that physically prevents (due to biological factors) the driver from detecting (and hence from attending to) information critical for safe driving.

- *Driver Misprioritised Attention (DMPA)*—Insufficient or no attention to activities critical for safe driving brought about by the driver focusing attention on one aspect of driving to the exclusion of another, which is more critical for safe driving.
- *Driver Neglected Attention (DNA)*—Insufficient or no attention to activities critical for safe driving brought about by the driver neglecting to attend to activities critical for safe driving.
- *Driver Cursory Attention (DCA)*—Insufficient or no attention to activities critical for safe driving brought about by the driver giving cursory or hurried attention to activities critical for safe driving.
- *Driver Diverted Attention (DDA)*—The diversion of attention away from activities critical for safe driving toward a competing activity, which may result in insufficient or no attention to activities critical for safe driving. This category has two sub-categories:
  - *DDA non-driving-related (DDA-NDR)*—The diversion of attention away from activities critical for safe driving toward a competing non-driving related activity.
  - *DDA driving-related (DDA-DR)*—The diversion of attention away from activities critical for safe driving toward a competing driving-related activity.

2.3 Ethical Considerations

Ethical approval for all aspects of the research design was sought, reviewed and granted by the CQUniversity Australia Human Research Ethics Committee Number: RSH/3337 Proposal for Human Factors Assessment of Tram-on-tram Collisions. Participants were provided with an information sheet about the study and informed consent was obtained before beginning each stage of data collection.

3 Results

The analysis revealed a range of strategies that the drivers used in their day-to-day driving, though it also highlighted many situations in which driver behaviour was suboptimal from a safety perspective. Consistent with thematic analysis methods, not every incident or issue was included and the points were representative of the many incidents and issues arising from the investigation. A brief preliminary account of the identified risk mitigation strategies is also presented.

3.1 Restricted Attention

In several scenarios, a cause of potential accidents was associated with the driver passing out from several physiological issues including overheating in the cab, lack of food and water, and insufficient sleep. This was also featured in an actual accident report and fell into the DRA category. In general, fatigue and sleepiness are common problems in rail, as with other shift-work, and create a risk for maintaining safety in the system [e.g. 9]. It is often difficult to determine the effects of fatigue and sleepiness on performance and especially accident causation [10], however it is reasonable to assume that it would have an impact. Many drivers reported feeling fatigued and again this was associated with restricted attention.

3.2 Misprioritised Attention

Conflicting goals in relation to safety and on time running were a recurrent theme, and correspond closely with rail driving in general [11]. All drivers reported safety as their main concern, yet there was also a lot of time pressure associated with maintaining on-time running (i.e. timetable). Despite the obvious concern for safety displayed by all drivers, they indicated that they would make risky decisions to make up lost time, with the aim to keep their time at zero. This factor could fall into the DMPA category as drivers found themselves focussing more on time management to the exclusion of safety. For example, one participant stated that although drivers would not blatantly run a stop signal, he believed there had been occasions where drivers pushed through on amber despite having time to stop.

An issue strongly connected to risk-taking behaviour and time management was found to be inter-depot competition and rivalry. The main cause of this was the desire to maintain on-time running, which was considered to be jeopardised by being caught behind a tram from another depot, causing delays. For example, a few routes went through the city and continued on through the other side to other suburban areas but most trams turned around at a shunting yard on the edge of the city centre. The trams that travelled through usually had large numbers of passengers boarding along the route shared by those trams that turned around at the edge of the city. Therefore the trams that did not travel through were likely to be held up by the large number of passengers boarding the tram in front at every stop along the route. As with the pressure of on time running, this led to risk-taking behaviour, caused by misprioritisation of goals and therefore corresponded most with misprioritised attention.

3.3 Neglected Attention

At one of the study sites, a small building impeded the view of tram drivers travelling down a gradient towards an intersection and track split. Here, priority was given to the tram entering from the adjoining track, so tram drivers travelling down the gradient had to be prepared to stop in order to give way. A signal at this junction provided only a few seconds notice to approaching tram drivers, meaning that the driver should look for approaching trams before beginning their descent as a safety check. Many participants indicated that this check was not made, and fell into the DNA category as the drivers failed to attend to safety-related activities, i.e. head check to the right.

During onboard observations, it was noted that many drivers did not turn their heads to the left to check for oncoming traffic when they entered a road/rail intersection. Although they were taught that danger typically came from the right, it was considered good practice to check their left side to ensure it was clear rather than assuming the case. Thus a failure to make a full head check corresponded with neglected attention.

3.4 Cursory Attention

There was some expectation bias associated with the manoeuvring of other vehicles and over-familiarity with stopping patterns, which led drivers to make assumptions about (and therefore plan around) the behaviours of other road users and other tram drivers. For example, participants reported that when they approached a stationary tram at a stop with no passengers waiting to board it, they would assume that the tram was about to move away and therefore aim to stop where the tram was currently located. This expectation was correct the vast majority of the time, however were are occasions when the tram in front did not move away when expected, thereby requiring the approaching tram driver to heavily apply brakes, to result in a collision or near miss. Drivers in the focus groups mentioned three different recent occasions when a tram-on-tram collision had actually occurred as a result of this expectation bias. This factor could fall into the DCA category as drivers gave only cursory attention to the tram in front.

Another situation was created by lack of clear rules. For example, there was a shelter and a flag situated a few metres apart at the off-road junction. Some tram drivers stopped their tram at the flag while others stopped at the shelter if people were waiting there. Drivers in the focus group pointed out that this was a problem because they believed the flag was the correct place to stop, thus most drivers would assume the tram in front would also stop at the flag and subsequently aim to stop their tram just behind them. If the tram in front happened to stop several metres short of the expected location, the following tram would not have time to stop. This fell into the Cursory attention category as tram drivers gave insufficient attention to what was actually happening because their expectations over-rode their attention efforts.

An element of safety related specifically to passengers was the operation of doors. During onboard observations, tram door closes were observed being activated while passengers were still disembarking. Some drivers were careful about ensuring the doors were clear before allowing them to close (they would hold the 'open' button down until they were sure passengers were clear). However other drivers allowed the doors to close when passengers were still trying to alight the tram. It is often difficult to see the doors properly when the tram is full including standing passengers, however some drivers were only gave a cursory glance in their mirrors to check the doors. Thus, this risk corresponded with cursory attention.

3.5 Diverted Attention

Non-driving-Related. In one of the reviewed Incident Investigation Reports, the driver reported that he had been thinking about his upcoming wedding and did not notice that the tram in front was stationary, which corresponded to non-driving related diversion of attention.

Drivers in the focus groups commented that they often arrived at their destination not remembering how they got there. Although they could not say why this happened, there are two possibilities for this common phenomenon; they may have operated with automaticity (i.e. outside of conscious awareness) and "switched off" from all thoughts, or their thoughts may have been non-driver related.

Driving-Related. At all times tram drivers must be constantly alert for pedestrians in the environment. This was especially so at the junction separated from the road where pedestrians commonly violated rules by walking alongside the track with no barrier between them and the moving tram. Drivers reported being very vigilant of pedestrians in this situation, often to the detriment of other important task-related attention demands, thus the driving-related category of diverted attention was most suitable for this situation.

Due to the pressure of on time running, drivers constantly checked their display or timetable to see how close they were to schedule. During observations both on-site and on the trams, several drivers were seen to be referring to the display or timetable while the tram was moving, creating a diversion of attention (away from activities critical for safe driving) that corresponded with a driving-related task.

3.6 Risk Mitigation

The findings were also used to conduct a preliminary analysis of possible risk mitigation strategies. A more extensive and detailed analysis is required to fully test the usefulness of the taxonomy for this purpose, however several potential strategies were determined from the preliminary analysis, including: improved fatigue management training; a review of rostering; a review of timetabling; placement of appropriate signage; clearer definition of certain rules and regulations; environmental modifications; and training for tram drivers, including use of specific simulated scenarios [12].

4 Discussion

As part of a broader investigation into the causes of tram collisions on a metropolitan tram network, Regan et al.'s taxonomy [6] was used to categorise causes that related to distraction or inattention. The first objective from this exercise was to determine what sorts of risk factors were applicable to distraction and inattention in tram driving. These was presented in the results and for the most part, correspond well with psychological factors driver distraction and inattention in heavy rail (e.g. trains) [11]. The second objective was to determine if the taxonomy was appropriate for use with the tram-driving task and rail in general. The most notable issue with the coding process associated with the taxonomy was that several issues could arguably have fallen into more than a single category. In this instance, the most appropriate category was chosen based on consensus during the analysis. In particular, the distinction between the neglected and cursory category was somewhat underspecified in the taxonomy and difficult to code for tram driving. For example, when a tram driver did not slow down quickly enough when approaching another stationary tram, it could be that they had paid only cursory attention or had neglected to pay any attention to the tram altogether. Similarly, the issue with drivers looking at their timetable was coded as 'Driver Diverted Attention Driving-Related' but this could also fall into the 'Misprioritised Attention' category because the drivers were prioritising time over safety.

There were also some events that may have been due to inattention but did not fall into any of the categories of the taxonomy. For example, habituation was a common problem raised by the drivers in the focus groups and individual discussions. An example of habituation was when a driver took out the same tram class for several weeks then changed to a different tram, for example a longer one with different handling characteristics. They became habituated to the first type of tram and then became inattentive about the appropriate way of handling the new tram, effectively driving in an automatic mode suitable to a different, in this example, a smaller type of tram. A number of drivers stated that this was a common occurrence and it also feature prevalently in the scenarios.

Additional categories and modifications to existing categories may assist with making the taxonomy more suitable for tram driving. For example, Driver Restricted Attention currently relates to biological factors such as micro-sleeps [6]. This could be modified to include restricted attention caused by elements external to the driver as was the case described under Driver Neglected Attention where drivers could not see approaching trams due to the location of a building on the approach to the track split. This modification could also potentially be added to the existing taxonomy for the purposes of categorising inattention in motorists.

A comparison of this method with other methods used in the investigation showed that the thematic coding often did not reveal the underlying causes of accidents and potential incidents. For example, simply stating that a tram driver was inattentive because they neglected to perform checks to the right for oncoming trams does not explain why they failed to check. Much more detail is needed to determine the root cause and possible mitigating strategies, and this an area for future work.

With regard to determining mitigating strategies, the coding was of some assistance but thinking of the causes in terms of the type of distraction or inattention did not lead to mitigation strategies over and above the broader analysis and in fact resulted in less strategies than other methods used in the investigation. However, as noted, at this stage it was a preliminary analysis only and a more detailed analysis may highlight more potential mitigation strategies. Additionally, only a relatively small quantity of data was suitable for coding which suggests that the taxonomy, in its current form at least, may not be as suitable for the tram-driving task as it is for the motor vehicle driving task.

5 Conclusion

The categories detailed in Regan et al.'s taxonomy [6], in their current form, may not be appropriate for the tram-driving task. However modifications such as clarification of codes and addition of new codes may make it more appropriate. A more detailed analysis is required with respect to the usefulness of the taxonomy in assisting with mitigating strategies for tram driving. While an example of distraction was detected for every category in the taxonomy, the types of distraction that are associated with managing competing activities not critical to safety, such as on time running, are arguably the most insidious. For the most part, others types of distraction and inattention, such as neglected and cursory attention, correspond with the nature of habituated behaviours and automaticity in the task itself. Future work in this area is needed.

Acknowledgments The authors are extremely grateful to the staff at Yarra Trams who participated in this research and facilitated the planning so that it could be possible.

References

1. Naweed, A., Moody, H.: A streetcar undesired: investigating ergonomics and human factors issues in the driver-cab interface of Australian trams. Urban Rail Transit. (2015). doi:10.1007/s40864-015-0021-x
2. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Fac. **37**(1), 32–64 (1995)

3. Williamson, A.: The relationship between driver fatigue and driver distraction. In: Regan, M. A., Lee, J.D., Young, K.L. (eds.) Driver Distraction: Theory, Effects, and Mitigation, pp. 383–392. CRC Press, Boca Raton, FL (2009)
4. Salmon, P.M., Young, K.L., Regan, M.A.: Distraction and public transport: case study of bus driver distraction. In: Regan, M.A., Lee, J.D., Young, K.L. (eds.) Driver Distraction: Theory, Effects, and Mitigation, pp. 229–245. CRC Press, Boca Raton, FL (2009)
5. Landy, S.: Couple injured after car, tram collide on Royal Pde, Parkville. Herald Sun. The Herald and Weekly Times, News Corp Melbourne, 14 April 2015
6. Regan, M.A., Hallett, C., Gordon, C.P.: Driver distraction and driver inattention: definition, relationship and taxonomy. Acc. Anal. Prev. **43**, 1771–1781. doi:10.1016/j.aap.2011.04.008 (2011)
7. Naweed, A.: Psychological factors for driver distraction and inattention in the Australian and New Zealand rail industry. Acc. Anal. Prev. **60**, 193–204 (2013)
8. Saldaña, J.: The coding manual for qualitative researchers. Sage, NY (2012)
9. Dorrian, J., Hussey, F., Dawson, D.: Train driving efficiency and safety: examining the cost of fatigue. J. Sleep Res. **16**(1), 1–11 (2007)
10. Naweed, A., Filtness, A.: Fatigue and sleepiness: complex bedfellows. Track Sig. **19**(1), 60–61 (2015)
11. Naweed, A.: Investigations into the skills of modern and traditional train driving. App. Erg. **45** (3), 462–470 (2014)
12. Naweed, A.: Simulator integration in the rail industry: the Robocop problem. Proc. Inst. Mech. Eng. Part F: J. Rail Rapid Transit **227**(5), 571–582 (2013)

7 Years of Experience with Demerit Point System in the Czech Republic: Is It Effective?

Petr Zámečník, Vít Gabrhel, Veronika Kurečková and Pavel Řezáč

Abstract This study maps the effectivity of the demerit point system in order to find possible complementing measures to the demerit point system and improve the existing ones. Our analysis was done on data from point system since 7/2006 till 12/2013 and national road safety and traffic accident statistic since 2006 till 2015. Results show, that decrease in road fatalities and serious injuries is partly caused by socioeconomic development and that Czech demerit point system has some severe weaknesses and inappropriate focus. That is why it fails in preventing young drivers and repeated offenders from (re)offending.

Keywords Demerit point system · Traffic offenders · Driver remediation · Driver recidivism · Traffic safety

1 Introduction

Regarding public health, road traffic safety is the most crucial issue of the 21st century [1]. World Health Organization informed on road traffic safety situation for the first time in a broader scale in its report in 2004 [2], when, based on the data of 99 % world population, WHO calculated that 1.3 million people are killed and 20–50 million people are injured in traffic every year. The results have not changed very much since that time, because the last two reports [3, 4] concluded with very

P. Zámečník (✉) · V. Gabrhel · V. Kurečková · P. Řezáč
Centrum Dopravního Výzkumu, v.v.i. (Transport Research Centre), Líšeňská 33a, 63600 Brno, Czech Republic
e-mail: petr.zamecnik@cdv.cz

V. Gabrhel
e-mail: vit.gabrhel@cdv.cz

V. Kurečková
e-mail: veronika.kureckova@cdv.cz

P. Řezáč
e-mail: pavel.rezac@cdv.cz

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_23

similar figures. In addition, the reports claim that in case the trend continues, in 2020 road traffic accidents will become the third most frequent causes of deaths in the world.

However, the causes of this risk do not stem from the imperfection of the system, but rather from human factor error. Nevertheless, the issue of human factor is too complex to formulate simple conclusions and recommendations. Driving is such a complex and diverse activity that it is impossible to identify several individual qualities which could help us determine safe and responsible drivers and dangerous and irresponsible drivers. In addition, drivers are in no way a homogeneous group, which prevents generalization and interpretation of collected findings. What the studies generally agree with is the fact that the degree of human factor impact on road traffic accidents is very high. For example [5] claims that human factor is a single cause of 57 % road accidents and a co-factor of more than 90 % road accidents. Later research shows a slightly increasing trend and consider human factor as a co-factor of 90–95 % road accidents [6–8].

However, in most cases traffic rules violations are more common than for example insufficient intake of information [9]. One of the main reasons is the absence of a negative feedback which increases the false belief in a low probability of accidents [10] and a subjectively felt higher chance of avoiding a road accident than others [11]. The absence of social feedback was changed with demerit point systems introduction. European countries were gradually introducing the so-called demerit point systems in order to provide negative feedback; more effectively prevent driving for risky individuals, and generally improve road safety. The Czech Republic is among the 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 demerit points have been recorded since 1st of July 2006.

Although the long-term road safety situation in the Czech Republic shows a positive trend, the country lags behind in reducing the consequences of road accidents in comparison with countries with similar infrastructure and traffic figures (e.g. regarding mortality in relation to mileage, which amounts to 15.7 fatalities per 1 billion driven kilometers, the Czech Republic ranks among the last in the IRTAD world database). The reasons for long-term reduction in road accident and fatality rate in the Czech Republic are often assigned to the introduction of the demerit point system, socio-economic development and technological progress. Relatively high long-term road accident rate and mortality rate on Czech roads in comparison with foreign countries is based on two reasons. (1) Insufficient system of gradual education of drivers and (2) minimum number of measures aimed at a small group the most risky drivers. However, these two factors are dealt with by the demerit point system with difficulties. The decline in road accident and mortality rate on Czech roads over the last five years (2009–2013) may thus have a different explanation than well-working demerit point system. Therefore, it is suitable to look at road traffic safety in the context of the socio-economic development.

It is quite clear that the introduction of the demerit point system leads to the reduction of road accidents and fatalities [12, 13], but the effect fades over time as there is, for example, lower effect on fatalities on motorways [13, 14]. Regarding some other aspects, we can see a stable long-term effect, e.g. on the use of safety belts and child restraint systems [15] on drivers as well as on passengers [16]. From this point of view we can consider demerit point systems effective and successful measures, since the use of safety belts and child seats are the most effective methods to prevent serious injuries and fatalities at road accidents [17].

The demerit point system works on the principle of improving road safety by deterring road users through the threat of punishment (general level) and through sanctions for those who are caught and convicted (specific level). The specific deterring effects can be evaluated by monitoring of the recidivism rate and general effects by prevalence of the negative phenomenon in driving community [18]. Nevertheless, the term "recidivist", which is often used to define drivers with problematic driving history, is frequently mistaken for many other definitions, such as "repeated offender" or "hard core offender" [19–23], therefore it is operationalized with difficulties. In addition, recidivism can be understood by many different ways and each of them has the potential to provide different result.

Firstly, the category of recidivists include those who commit crime or offences deliberately and unintentionally, secondly, the seriousness of given offences or criminal acts (just points for given behaviour are deducted regardless of its context and real hazard) is not taken into account. What is missing is scaling of recidivism seriousness, since the term in legal, not research, sense of word is concerned [18]. Furthermore, if a driver is caught drinking and driving for the first time and driving under the influence of drugs for another time, such driver is not considered as recidivist by the point system, since those are different law violations. In general, the effectiveness of sanctions in the form of recorded points as a feedback tool for the change in behaviour is doubtful, since the relationship is still unclear [24].

2 Demerit Point System in the Czech Republic

2.1 Form and Development of Demerit Point System in the Czech Republic

The existing demerit point system in the Czech Republic was introduced on 1 July 2006, based on the Acts Nos. 411/2005 Coll. and 226/2006 Coll., which superseded the Act No. 361/2000 Coll., on Road Traffic. "Demerit" points are recorded in the register of drivers by authorities of municipalities with extended powers within 5 days from the day the driver receives particular decision or is informed of effective sanction for violating selected obligations in traffic. The reason for introducing the demerit point system in the Czech Republic was particularly based on sanctioning recidivistic delicts against road traffic safety. In addition, regarding certain groups of drivers, e.g. drivers from higher socio-economic groups, the financial sanctions, even repeated ones, proved to be insufficient and failed to prevent them from committing further offences, since there was no threat in the form of driving license suspension. The demerit point system introduced an administrative procedure which may lead to driving license suspension. Since its establishment, the system has changed several times and has been further modified by amendments of other acts. The changes particularly follow the directive of the European Parliament and of the European Council 2006/126/ES from 20 December 2006 on Driving Licenses, as amended. Some of the most significant changes are specified below:

The first change occurred on 1 September 2008. The Act No. 374/2007 Coll. and Decree No. 156/2008 Coll. have allowed to perform "safe driving training" course, after which 3 demerit points can be eliminated under conditions specified by law.

Probably the most significant change occurred on 1 August 2011, when the Act No. 133/2011 Coll., which amended the act on road traffic, became effective. This amendment revoked the point deduction for offences against road traffic safety which were considered less serious (in total from 44 to 27 penalized items). They particularly concerned unauthorized use of special warning lights, failure to use assigned traffic lane, driving without medical assessment on fitness to drive. Regarding aggressive behaviour (drinking and driving, speeding), the number of deducted points increased; based on this amendment, no points are deduced for behaviour related to relatively low hazard (e.g. no lights or unauthorized use of assigned traffic lane). Furthermore, the amendment brought about the obligation to participate in traffic psychological examination, when applying for driving license to be returned.

A prohibition has been effective since 1 January 2012 to issue a driving license to a person whose driving license is suspended or revoked in another EU or EEA member state or who has been disqualified from exercising professional activities based on driving motor vehicles, unless the time period for returning driving license has elapsed (e.g. time of disqualification from exercising professional activities or the "out of points" time).

So-called objective responsibility of vehicle owner was introduced on 19 January 2013, in the manner of other EU countries, e.g. France, the Netherlands, Slovakia, which is to prevent the abuse of the status of a "close person". In case the identity of driver who committed a traffic offence cannot be determined, the sanction is paid by the given vehicle owner.

2.2 Demerit Point System Statistics

The data for the mentioned statistics were collected form publicly available statistical data of the Ministry of Transport, particularly regarding the National road safety strategy [25], and information on the status of the demerit point system in the Czech Republic [26] a partly from a non-public database of Ministry of Transport of the Czech Republic called "Eliška" [27]. Nevertheless, even within these non-public accesses for research purposes, it is possible to monitor just basic indicators, such as the number of individual offences, age and sex of offenders, and the region where offence was committed.

Regarding the changes which has occurred since the beginning of the introduction of the demerit point system (particularly the number of offences with point deduction and the number of points), the outputs for individual years are difficult to compare. However, based on [27], on average 0.1–0.2 % of drivers reach the limit of 12 demerit points each year, while two types of drivers are overrepresented—young drivers and male drivers. Furthermore, the rate of recidivism ranges around 50 % for failure to use seat belts, 25 % for using mobile phones, 30 % for driving without driving license, 15 % for drinking and driving and 15–50 % for speeding (depending whether any speeding is considered or how much the speed limit was exceeded is taken into account as well). Finally, the offence profile of people "out of points" for the first time and for the second time is very similar.

References [25] and [26] statistics show that on 31 December 2015 there were 39,507 drivers with 12 points deducted, i.e. 7.33 % drivers with deducted points and 0.59 % of all registered drivers. Out of the total number of drivers with 12 points deducted, there were 36,851 (93.28 %) men and 2656 (6.72 %) women. The gender ratio of all registered drivers is relatively stable—approx. 55.73 % men and 44.27 % women. The number of men and women with deducted points differs greatly from the ration of all drivers—it is currently formed by 81.72 % men and 18.28 % women. The differences between men and women with deducted points in individual regions are negligible (the ration of men ranges between 79.45 and 84.06 %). On 31 December 2015, 539,277 drivers were registered in the central register of drivers for committing an offence (criminal act) with at least one point deduction. This is 8.05 % of all registered drivers (6,700,000), i.e. every 12th driver has a point deduction. It is necessary to take into account the fact that a part of drivers is inactive and therefore the proportion of point deducted and "out of point" drivers among active drivers is actually higher. It is interesting that a combination of all types of speeding (55 %), or slight speeding (36 %) in combination with failure to use seat belts (20 %) form more than half of the offences.

Regarding the road accident rate statistics [25], the national statistics show that the number of fatalities in 2010 was reduced by 42.7 % in comparison to 2002 and in 2012 by 48.2 % (in comparison to 2002). However, the decreasing trend reversed in 2014 and in the last two years (2014 and 2015) the increasing trend appeared again. In 2014 the set target of NSBSP 2020 (National road safety strategy 2011–2020) was not met with regard to the number of fatalities in both monitored categories—within 24 h and within 30 days and their number even significantly increased in comparison with 2013—by 7.9 % (within 24 h), or by 5.2 % respectively (within 30 days). The statistics of 629 fatalities in 2014 (target 594) and 660 fatalities in 2015 (target 505) indicates that in order to reach good results from the previous years, it will be necessary to introduce new measures, and that very optimistic estimate of the number of 360 fatalities in 2020 will most likely fail to be met. The statistics of calculation of losses from road accident rate stay stable in the long-term—they range between CZK 40 and 50 billion per year and roughly correspond with 2 % GDP. Another stable trend is the group of young drivers as the most hazardous group of road users [28].

2.3 Demerit Point System Evaluation

Effectiveness of Sanction Part of the System. The effectiveness of the demerit point system to prevent recidivism depends on the perception of sanctions by drivers. The most important appears to be the certainty of punishment [29], while the increase in the severity of punishment only slightly reduces the recidivism in risky behaviour [30] and a slight change in the behaviour of population regarding the general level of deterrence [18, 31]. The severity of sanctions is an effective factor in case the certainty of be caught and punished is high [18, 32–34]. However, this is particularly difficult with hard-core recidivists, when we can find several differences in comparison with "common" recidivists. In comparison with common recidivists they consider the risk of being caught as lower, benefits from traffic offences higher, and losses related to offences lower [35]. Consequently, a large part of the imprisoned drivers returns to the prison within three years [29]. Regarding drinking and driving, we can see the effect of addicted life style. Those who drink more drink and drive more frequently, while those with low or medium alcohol consumption are more able to separate drinking and driving [36].

Regarding the effect of the demerit point system on recidivism, the system is the least effective on speeding (more than 50 % of registered offences and the recidivism rate of 50 %) and on seat belt use (recidivism of 50 % likewise). The system is seemingly the most effective on drinking and driving, although the low rate of recidivism (15 %) is caused by the longer period of driving license revocation (the demerit point system has not been introduced for sufficiently long time to fully manifest the recidivism rate) and by the fact that a part of those drivers has not been registered in Eliška database since 2011, due to being allocated in a separate specialized database [27].

Effect on Young Drivers. NSBSP 2020 also focuses on young drivers and introduction of stricter conditions for point deduction for drivers beginners with up to 2 years of driving practice [35]. No stricter conditions for point deduction for drivers beginners were introduced by the amendments in 2011 and 2013, since the constitutional principle of equality and prohibition of discrimination would have been breached. This in turn leads to the fact that drivers under 25 years of age in the Czech Republic are the most point-deducted and out-of-points group, die in road accidents more frequently than other groups, and cause road accidents more frequently than other groups. This is particularly misfortunate in combination with potential elimination of points for training to deal with critical situations. The introduction of this measure in the system in 2008 has neither been evaluated, nor sufficiently justified. However, based on the available literature, this measure seems rather counterproductive, particularly for novice drivers.

It is well known that some criticism has been reported against this type of training since the late 1980s. For example [37] reported an increase in slippery road accidents after skid training became a mandatory part of the training. Keskinen et al. [38] pointed in the same direction: after the introduction of skid training, a larger proportion of young driver's accidents occurred in slippery road conditions. The suggested explanation was that the skid training courses were mostly based on exercising maneuvering skills and this resulted in overconfidence of driving skills. Instead of using their better skills in critical situations the students used their improved skills in ordinary driving situations, e.g. for higher speed [39]. The skill-training group of novice drivers rates their skill as greater than other groups, although the actual skill level in skid-test conditions did not differ [40]. We can generally said, that any training programmes that teach drivers to respond in emergency situations lead to aggravated road traffic safety [40–42].

3 Discussion

Although we can see a reduction in the accident rate and mortality rate on Czech roads in the last decade, the question is how much is this reduction influenced by the demerit point system. In the context of development trends in the Czech Republic the situation is similar to most other countries in Central and Eastern Europe, despite the fact that demerit point systems were introduced at different times (or not implemented at all). The statistics since 2009 were probably influenced by improvements in technology and health care and also by the economic crisis, which is also indicated by a negative trend in road traffic safety since 2013.

European statistics show that the largest decrease in accident rates can be seen in the regions with the highest unemployment rates and that the biggest decrease was obvious in large cities [43]. Therefore, the economic recession is associated with reduced traffic volume, especially in the field of freight transport and commuting. Furthermore, we can see a change in travel behaviour, where part of the drivers changed trips for leisure activities for local activities and focused more on "fuel save" driving, which, among other things leads to reduced speed [44]. Regarding mortality, a significant help comes from technological advancement in passive safety and improved health services, which is mainly beneficial for elderly drivers and road users [45]. However, the progress in technology does not necessarily contribute to safety, as some drivers feel safer thanks to better technology and tend to more risky behaviour [43].

The stable decrease in road fatalities (with the exception in 2007) and serious injuries between 2006 and 2013 can be therefore explained as follows. In the first years the demerit point system worked very well—the effectiveness of the demerit point system is generally the highest in one or two years after its introduction [46]—and performed its function. But it is the starting economic crisis which has had a significant contribution to the reduction of accident rate since 2009. The subsequent negative trend since 2013 shows that there is a need for a modification of the system. The main shortcoming can be considered the low certainty of punishment based on the system (law enforcement in the Czech Republic is generally low, e.g. [47]) and also that changes in the demerit points system went mainly towards higher severity of sanctions, not reflecting the need to focus on young drivers and interconnecting the system with a counterproductive measure. Furthermore, during the lifetime of the demerit point system, there were no support campaigns or actions for better compliance (only general road safety actions). In addition, the administration related to the demerit point system has excessive impact on the justice system, and the system of state and public administration, while the costs related to the administration could be spent on preventive and rehabilitation programmes and campaigns.

4 Conclusions

Although the evaluation of effectiveness of the demerit point system is very difficult, we consider the system in the Czech Republic insufficient, based on the described data, particularly due to the absence of focus on the most risky populations—young drivers and recidivists. The sanction part of the demerit point system fails to reach its optimum function for its focus on sanction severity and lower enforcement. The preventive part of the demerit point system includes a counterproductive measure of a critical situations training course and lacks a rehabilitation part. The positive trend in fatalities and serious injuries in this context can be rather attributed to the effects of the economic crisis, which is confirmed by the aggravated situation since 2013.

Acknowledgments 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. Moller, J.H.: Driver health and traffic safety: an overview, pp 1–22. In: Verster, J.C., Pandi-Perumal, S.R., Ramaekers, J.G., de Gier, J.J. (ed.) Drugs, driving and traffic safety (2009)
2. Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A.A., Jarawan, E., Mathers, C.: World report on road traffic injury prevention. World Health Organization, Geneva (2004)
3. WHO: Global status report on road safety 2013: supporting a decade of action. ISBN 9789241564564
4. WHO: Global status report on road safety 2015. ISBN 9789241565066
5. Treat, J.R., Tumbas, N.S., McDonald, S.T., Shinar, D., Hume, R.D., Mayer, R.E., Stansifer, R. L., Castellan, N.J.: Three-level study on the causes of traffic accidents. Final report. 1979 (US DOT HS-805-086)
6. Najeeb, M.A.: A study of the psychological factors that influence the rule violation behavior of drivers. In: Conference Road safety in a globalised and more sustainable world—current issues and future challenges (2012)
7. Risser, R.: Validity of driver s diagnostics and selection. In Risser, R. (ed.), Assessing the driver, pp. 89–101. Braunschweig (1997)

8. Sommer, M., et al.: Cognitive and presonality determinants of fit to drive. Transp. Res. Part F Traffic Psychol. Behav. **11**(5), 362–375 (2008)
9. Reason, J.T., Manstead, A.S.R., Stradling, S.G., Baxter, J.S., Campbell, K.: Errors and violations on the road: a real distinction? Ergonomics **33**(10–11), 1315–1332 (1990)
10. Fuller, R.: Behavior analysis and unsafe driving: warning learning trap ahead! J. Appl. Behav. Anal. **24**(1), 73–75 (1991)
11. 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)
12. 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. Community Health **64**, 274–276 (2010)
13. Butler, J.S., Burke, J.P., Healy, D.G., Stephens, M.M., McManus, F., McCormack, D.: Trends in RTA related spinal injuries: the post penalty points era. Ir. J. Med. Sci. **175**, 20–23 (2006)
14. 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. Inj. Prev. **13**, 60–64 (2007)
15. Zambon, F., et al.: Sustainablility of the effects of the demerit points system on seat belt use: a region-wide before-and-after observational study in Italy. Accid. Anal. Prev. **40**, 231–237 (2008)
16. 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. Community Health **61**, 877–881 (2007)
17. Waller, P.F.: Challenges in motor vehicle safety. Annu. Rev. Public Health **23**, 93–113 (2002)
18. 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**(2015), 27–37 (2015)
19. Freeman, J., Liossis, P.: Drink driving rehabilitation programs and alcohol ignition interlocks: is there a need for more research? Road Transp. Res. (2002)
20. Hedlund, J., Fell, J.C.: Repeat offenders and persistent drinking drivers in the U.S. Paper presented at the association for the advancement of automotive medicine 39th annual conference, USA (1995)
21. Hubicka, B., Laurell, H., Bergman, H.: Criminal and alcohol problems among Swedish drunk drivers—Predictors of DUI relapse. Int. J. Law Psychiatry **31**, 471–478 (2008)
22. Styles, T., Imberger, K., Cairney, P.: Development of a best practice intervention model for recidivist speeding offenders. Austroads Project No. SS1389, Sydney (2009)
23. Yu, J.: Punishment and alcohol problems recidivism among drink-driving offenders. J. Crim. Justice **28**, 261–270 (2000)
24. Weatherburn, D., Moffatt, S.: The specific deterrent effect of higher fines on drink-driving offenders. Br. J. Criminol. **51**, 789–803 (2011)
25. Report on national strategy of road safety (NÁRODNÍ STRATEGIE bezpečnosti silničního provozu). Ministry of transport of the Czech Republic, http://www.ibesip.cz/cz/besip/strategicke-dokumenty/narodni-strategie-bezpecnosti-silnicniho-provozu/nsbsp-2011-2020
26. National report on demerit point system. Ministry of transport of the Czech Republic. http://www.mdcr.cz/cs/Silnicni_doprava/bodovy_system/
27. Statistic information about "out-of-points drivers" between 1st of July 2006 and 31st of December 2013. Non-public data of demerit point system of Ministry of transport of the Czech Republic. Accessed in June 2014
28. Report on fullfiling national strategy of road safety (Národní strategie bezpečnosti silničního provozu: Informace o plnění v roce 2014), http://www.ibesip.cz/cz/strategie/narodni-strategie-bezpecnosti-silnicniho-provozu/nsbsp-2011-2020/plneni-nsbsp-2011-2020
29. Vito, G.F., Maahs, J.R., Holmes R.M.: Criminology: theory research and (criminal justice illuminated). Theory Res. Policy (2006)

30. 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. **6**, 483–491 (1991)
31. Elvik, R., Christensen, P.: Deterrent effect of increasing fixed penalties for traffic offences: Norwegian experience. Paper presented at the transportation research board 85th annual meeting (2006)
32. Briscoe, S.: Raising the bar: can increased statutory penalties deter drink-drivers? Accid. Anal. Prev. **36**, 919–929 (2004)
33. Homel, R.: Policing the drinking driver: random breath testing and the process of deterrence. Federal Office of Road Safety, Canberra (1986)
34. von Hirsch, A., Bottoms, A.E., Burney, E., Wikstrom, P.O.: Criminal deterrence and sentence severity. Hart Publishing, Oxford (2000)
35. Loughran, T.A., Piquero, A.R., Fagan, J., Mulvey, E.P.: Differential deterrence: studying heterogeneity and changes in perceptual deterrence among serious youthful offenders. Crime Delinq. **58**(3), 3–27 (2012)
36. Mann, R.E., Smart, R.G., Stoduto, G., Adlaf, E.M., Vingilis, E., Beirness, D., Lamble, R., Asbridge, M.: The effects of drinking-driving laws: a test of the differential deterrence hypothesis. Addiction **98**(11), 1531–1536 (2003)
37. Glad, A.: Phase II driver education, effect on accident risk. Oslo, Norway: Transport Institute, TOI rapport 15 (1988)
38. Keskinen, E., Hatakka, M., Katila, A., Laapotti, S.: Was the renewal of the driver training successful? The Final Report of the Follow-Up Group. Psychological Report No. 94. University of Turku (1992)
39. Engström, I., Gregersen, N.P., Hernetkoski, K., Keskinen, E., Nyberg, A.: Young novice drivers, driver education and training. Literature Rev. VTI rapport 491A (2003)
40. Gregersen, N.P.: Young drivers' overestimation of their own skill-an experiment on the relation between training strategy and skill. Accid. Anal. Prev. **28**(2), 243–250 (1996)
41. Elvik, R., et al.: The Handbook of Road Safety Measures, 2nd edn. Emerald Group Publishing Limited, Bingley, UK (2009)
42. Mayhew, D.R., Simpson, H.M.: The safety value of driver education and training. Inj. Prev. **8** (Suppl 2), ii3–ii8 (2002)
43. Lloyd, L., Wallbank, C., Broughton, J.: A collection of evidence for the impact of the economic recession on road fatalities in Great Britain (2015)
44. Sivak, M., Schoettle, B.: Toward understanding the recent large reductions in US road fatalities. Traffic Inj. Prev. **11** 561–566 (2010)
45. Cheung, I., McCartt, A.T.: Declines in fatal crashes of older drivers: changes in crash risk and survivability. Accid. Anal. Prev. **43**(3), 666–674 (2010)
46. Castillo-Manzano, J.I., Castro-Nuno, M., Pedregal-Tercero, D.J.: An econometric analysis of the effects of the penalty points system driver's license in Spain. Accid. Anal. Prev. **42**, 987–1420 (2010)
47. National report on competitivness of the Czech Republic. Ministry of industry and trade of the Czech Republic. http://download.mpo.cz/get/43266/48536/573453/priloha001.pdf (2010)

Driving Related Fear—A Complex Problem with a Complex Treatment

Veronika Kurečková, Aleš Zaoral, Pavel Řezáč and Petr Zámečník

Abstract It is estimated that up to 15 % of the population have symptoms of driving phobia. Strong driving related fear affects 1 % of the population, mostly women. From 2012 until 2015 we had 40 clients with driving phobia. For most of them driving related fear manifested as a part of a more complex self-esteem problem, not as an isolated issue. We used various methods concerning the whole contexts of client's life and the system of client's relationships. 12 out of the 17 participants of group program and 20 out of the 23 participants of the individual programs experienced a significant decrease of the driving related fear symptoms and started to drive regularly. Based on our experience—driving related fear is rarely an isolated symptom, so it has to be treated as a part of a more complex problem. The complex approach significantly contributes to the efficacy of the therapy.

Keywords Driving related fear · Driving phobia · Therapy · Treatment · Trauma

1 Introduction

Driving is an integral part of lives of the majority of population and the lives in turn are closely connected with driving and transport. It is well documented, that emotions play important role in driving at the site of human factor. They directly

V. Kurečková (✉) · A. Zaoral · P. Řezáč · P. Zámečník
Centrum Dopravního Výzkumu, v.v.i. (Transport Research Centre),
Líšeňská 33a, 63600 Brno, Czech Republic
e-mail: veronika.kureckova@cdv.cz

A. Zaoral
e-mail: ales.zaoral@cdv.cz

P. Řezáč
e-mail: pavel.rezac@cdv.cz

P. Zámečník
e-mail: petr.zamecnik@cdv.cz

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_24

influence willingness and ability to drive. As the scientific interest is most commonly focused on anger and hostility, they are the best-documented emotion and there are the clearest effects for these emotions. Most studies explored the relationship of these emotions with hazardous and aggressive driving [1]. But there is also other part of the spectrum. Not only drivers with high recidivism rate and aggressive drivers deserve attention. Besides anger and hostility, fear and depression are rated among the most important emotions affecting driving. Feeling of anxiety, fear or stress is related to the higher number of errors and reduced ability or inability to drive what leads among others to reduced wellbeing and lower position on the labor market.

Clinical studies about fear of driving are usually focused on driving related fear (DRF) after crash and posttraumatic stress disorder (PTSD) or on psychiatric part of this phenomenon as the driving phobia is often connected with agoraphobia or panic disorder [2]. A significant proportion of the existing research on DRF has come from an interest in psychological consequences of motor vehicle accidents (MVAs). PTSD and DRF appear relatively often in the package of symptoms that can manifest as a result of MVAs. There are (mostly clinical) studies which explore DRF and PSTD arising from MVAs and MVA-related driving fears [3], but exploring DRF which is not connected with MVAs is performed very rarely. Typically, there is one way direction in research MVA—DRF/PSTD symptoms—clinical treatment. However, the other sources of DRF, differences among various types (and consequently differences in causal treatment), description of symptoms, triggers and development of DRF still remain unclear.

Due to the fact that fear is the most potent negative force, anxiety, fear and phobia of driving are one of the psychotherapeutic and psychiatric themes of growing interest. There are many studies as well as therapeutic approaches and procedures concerning this disorder. The driving phobia seems quite common. It is estimated that up to 15 % of the drivers' population have symptoms of driving phobia [4]. Strong phobia from driving (amaxophobia) affects 1 % of the population, mostly young women. The problem of driving phobia is not about young or novice drivers only, one quarter of people aged 55+ feel driving anxiety, despite the fact that many of them have been driving for many years [5]. Most of the people that report about their suffering from the driving phobia are women. This is proved by the research [6–8] as well as by the experience from the Transport Research Centre, Czech Republic [9]. However, in the Czech Republic we have neither exact numbers nor estimations, as the problem of driving phobia has so far not been considered as an issue. Despite serious anxiety symptoms many people are not aware that their problems are an indicator of a potentially serious phobia that deserves treatment.

Despite of the growing interest in the driving phobia, there are still some controversies about categorizing the anxiety, fear or phobia of driving. The essential question is—what really is a driving phobia? However, it seems that there is no simple answer. A mild fear from driving is a very common issue, especially with novice drivers. A mild anxiety with no avoidance tendencies and no impact on social, occupational and personal functioning might be perceived as physiological. After all, driving a vehicle is a complex psychomotor process requiring sensorimotor coordination, psychophysical judgment, attention, emotion and reaction time [10]. It can never be absolutely safe so some concerns are reasonable. However, in some cases the concerns can exceed a certain limit and exclude drivers from the active car using or cause serious psychical states during driving, resulting in serious phobic symptoms with huge impact on many aspects of patient's life, that can be considered as a psychological disorder.

The causes and pathogenesis of the driving phobia seem not be very clear. It has been stated many times that the MVA and non-MVA studies suffer from inconsistency in defining fear of driving [11]. Early research and conceptions [7, 8, 12, 13] considered driving phobia as a disorder following motor or non-motor vehicle accidents. Even some recent research shows that 18–77 % of people after a serious accident suffer by a phobia [4], but this does not implicate that driving phobia is always a result of a traumatic experience on the road. It can have various causes as well as levels of severity and symptom patterns. In some cases driving phobia is manifested by fear of driving or fear of the sudden panic attack or anxiety while driving [14]. Some of the drivers can be worried about potential criticism or negative evaluation of their performance [13]. Many of the patients experience a fear of losing control of the vehicle and the situation or a fear of specific driving situations (high speeds, night driving, open roads driving, tunnel driving, driving over bridges, steep roads etc.).

Some diagnosis as panic disorder, posttraumatic stress disorder and social phobia are considered to be part of the driving phobia [15]. Driving fear frequently appears as an accompanying symptom of agoraphobia. The image typically combines agoraphobic symptoms, panic and specific phobia [6]. Other authors state that driving phobia can develop after the individual experiences of unexpected panic attack in the feared situation [11]. Some authors [11] perceive the difference between specific phobias and agoraphobia in the focus of apprehension. Patients with agoraphobia feel a fear of panic and its consequences—anxiety expectancy; in contrast patients with specific phobia feel a fear of specific danger—danger expectancy [11, 16, 17].

There is a connection between driving phobia and actual driving skills; however, there is not much research of this issue [10]. In some cases people are not feared of driving only, but overall by travelling in any motor vehicle. In many cases there are other phobias or psychopathologic symptoms continuously developing and changing which sometimes complicates the diagnosis, as well as treatment. In short, driving anxiety, fear and phobia include many levels of intensity, as well as wide and various scales of symptoms. The nature and development of the disorder cannot be easily described, as it involves complex pathogenesis patterns that have to be concerned. So as the symptoms and their intensity the methods and proceedings of treatment vary too. Experts usually emphasize the need of the individual therapeutic schedule, tailor-made for each of the clients.

Cognitive behavioral psychotherapy has proved a good efficacy in many of the cases [18, 19]. de Jongh et al. [20] reported using of trauma-focused Cognitive Behavioral Therapy (TF-CBT), and Eye Movement Desensitization and Reprocessing for travel phobia. They noticed that the stronger the symptoms were, the more significant was the decrease of its intensity after the therapy. da Costa et al. [17] as well as Wald and Taylor [14] reported the effects of virtual reality exposure therapy (VRET)—especially as a factor facilitating the in vivo exposure. Walshe et al. [21] suggest some basic measures to ensure an appropriate level of immersion during the VRET exposures. Other authors [22] describe a case study of creating a virtual reality in hypnosis. Hypnosis and imagination proved to be an effective tool for driving phobia treatment—separately or combined with some other methods. E.g. Hill and Bannon-Ryder [23] combined the hypnosis with cognitive behavioral programs. Williamson [24] successfully combined self-hypnosis training, instituting a calmness anchor and dissociative imagery with positive mental rehearsal. However, the driving phobia does not have to be treated by worlds and images only. E.g. transcranial magnetic stimulation as a potentially effective method for anxiety treatment was mentioned by some authors [25]. The maladaptive behavior people with driving phobia often engage themselves in [26] is a factor that contributes on the road safety decrease. The authors state it is probably because a high frequency of cognitive distortions facilitates the tendency to protect themselves from unpredicted dangers when driving. Also sudden panic attacks while driving increase the risk of traffic accidents and therefore constitute a risk for all road users.

In some cases driving phobia manifests itself in the driving situations only. Therefore it seems to be an isolated problem. However it influences a lot the patient's life. The restriction of freedom, career handicaps and social exclusion [26, 27] are only some of the disadvantages people with driving phobia face with. In some cases it can be the cause of serious existential problems. Research of da Costa et al. [26] comparing the clinical characteristics and quality of life of the women suffering by driving phobia in comparison with the women without such symptoms. The research showed lower scores in the quality of life scales, especially in these items: "functional capacity", "social aspects", and "mental health". They obviously scored higher in anxiety.

2 Experience from the Czech Republic

Transport Research Centre has begun the driving phobia research and therapy programs in 2012. During the years 2012–2015 we had 40 clients; 17 of the clients attended group programs, 23 of them attended the individual therapy. With all of our clients following symptoms appeared: (a) fear of driving (self-reported); (b) driving avoidance; (c) anxiety symptoms while driving. The reports from the therapies were analyzed.

2.1 Methods of Treatment

In the group programs (2 groups, 17 participants) we worked with expectations and goals and motivations for driving. During the individual therapy many methods and approaches were applied. For most of our clients driving phobia manifests as a part of a more complex self-esteem problem, not an isolated issue. We used various methods concerning the whole contexts of client's life and the system of client's relationships. Important supporting fellows were often invited to the therapy. We also used practical training with the driving school instructor and psychologist—it had a significant effect as an exposition in vivo but also as a method of the driving skills development that influenced both actual skills and self-efficacy of our clients. With some of the clients we used a driving simulator for routine training and anxiety reduction. 12 out of the 17 participants of group program and 20 out of the 23 participants of the individual programs experienced a significant decrease of the driving related fear symptoms and started to drive regularly.

2.2 Driving Related Fear—Isolated Symptom or a Part of the System Problem?

For most of our clients driving related fear was not an isolated symptom. In many cases the clients had symptoms of social phobia and general anxiety. More than 75 % of the clients showed some psychasthenic symptoms. Serious problems with self-esteem were noticeable, mostly in connection with the relationship to mother. Our client's mothers very often showed anxiety and we could also notice attachment issues in our client's childhood. Partner issues were also one of the determinants of driving related fear. In some cases there was a conflict with partner that resulted in the discouraging behavior as a part of the fight for the hegemony. Usually there were also supporting persons that encouraged our clients to drive. We invited those persons to the therapy as supporting elements.

Driving itself is also a complex problem that extents to many aspects of the client's life. There are four levels of driver competencies that can be considered with the DRF issues, based on the GDE matrix [28].

- *Vehicle control—driving skills issue*: Actual driving skills were quite low with women with DRF, mainly due to the lack of routine. Sometimes the driving performance can be affected by the temporary or permanent cognitive deficit. In three cases our clients had significant cognitive deficits. With two of them the deficits were probably determined by the psychical overload and faded away after the therapeutic treatment. On the other hand many of the clients showed concerns about their driving potential that proved to be inadequate. Lack of driving skills determines significantly the client's self-esteem. For such clients driving training can be a reasonable part of the treatment.

- *Mastery of traffic situation*: Other traffic participants are usually a stressful element. Clients are worried about the potential risks the others can cause, as well about the potential risks they can cause when interacting other road users.
- *Strategy and planning*: Lack of self-confidence and self-esteem and strong concerns about the others influence a lot the planning and strategy. Respects to the needs of follows also should be considered.
- *Lifestyle*: The context of client's life and significant relationships should be considered on this level.

2.3 Men Versus Women

There were only 4 men out of the group of 40 clients. We noticed significant differences between men and women in our sample. With men we noticed significantly more serious symptoms of anxiety. It can be caused by the fact that men usually hesitate to ask for a professional help and they join the therapy only if their symptoms are beyond endurance. With our male clients we revealed a strong father person as a problematic element—the fathers were usually dominant and despotic, with no support of their son's self-confidence. We have also noticed a significantly higher level of the driving skills that contrasted with the high level of anxiety. Men very often reported sudden outbreak of the DRF symptoms after many years of unchallenged driving experience. Women usually experienced DRF symptoms right from the beginning of the driving practice.

2.4 DRF as a Consequence of Traumatic Experience?

With our clients serious post-accidents traumas were extremely rare. On the other hand there were traumatic experiences from the driving school. Many of the clients reported sexual attacks or humiliation from the driving instructors. During the process of novice driving traumatic conflicts with fellows (parents or partners) were also very often. In some cases other serious traumas (not connected to driving) acted as a determinant of anxiety and influenced the driving related fear.

3 Conclusions

Driving related fear is a serious problem that influences social inclusion, individual well-being as well as the traffic safety itself. There are many methods that can be used with appropriate results. However we noticed that the DRF is rarely an isolated symptom. Therefore it must be treated as a part of a complex context of client's life, concerning individual specifics as well as the significant follows of the client but also the more general social and cultural context. In some cases all levels of driving competencies must be developed, in some cases the therapy is focused on the psychopathological symptoms. However the complex approach is what significantly contributes to the efficacy of the therapy.

Our experience differed from the experience of many foreign experts. However our experience is based on limited number of cases, therefore it cannot be generalized. It probably indicates a high variability of symptoms, determinants as well as the methods of treatment of the DRF issues.

We have to emphasize that individual therapeutic approach reflecting individual specifics is always necessary. There is also a need of the systematic social measures that can help to prevent the DRF (especially driver education system) or facilitate the treatment (system of early diagnostic of the DRF, system of effective treatment, accessible for everyone, advantages and risks of alternative therapeutic methods—e.g. medical treatment etc.). Further research of the DRF issues, including prevalence studies in specific populations, can also be helpful.

Acknowledgments 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. Mesken, J., et al.: Frequency, determinants, and consequences of different drivers' emotions: an on-the-road study using self-reports, (observed) behaviour, and physiology. Transp. Res. Part F: Traffic Psychol. Behav. **10**, 458–475 (2007)
2. Taylor, J.E., Deane, F.P., Podd, J.V.: The relationship between driving anxiety and driving skill: a review of human factors and anxiety-performance theories to clarify future research needs. NZ J. Psychol. **37**(1), 28–37 (2008)
3. Taylor, S., Koch, W.J.: Anxiety disorders due to motor vehicle accidents: nature and treatment. Clin. Psychol. Rev. **15**, 721–738 (1995)
4. Taylor, J.E., Paki, D.P.: Wanna drive? Driving anxiety and fear in a New Zealand community sample: a pilot study. NZ J. Psychol. **37**(2), 31–37 (2008)
5. Taylor, J.E., Alpass, F., Stephens, C., Towers, A.: Driving anxiety and fear in 55–70 year-old New Zealanders. Age Ageing **40**(1), 62–66 (2011)
6. Ehlers, A., Hofmann, S.G., Herda, C.A., Roth, W.T.: Clinical characteristics of driving phobia. J. Anxiety Disord. **8**, 323–339 (1994)
7. Mathew, R.J., Weinman, M.L., Semchuk, K.M., Levin, B.L.: Driving phobia in the city of Houston: a pilot study. Am. J. Psychiatry **139**, 1049–1051 (1982)
8. Munjack, D.J.: The onset of driving phobias. J. Behav. Ther. Exp. Psychiatry **15**, 305–308 (1984)
9. Kurečková, V. Zámečník, P., Daňková, L.: Driving phobia treatment—experience from the Czech Republic. In: Impact Conference Proceedings (2014)
10. Taylor, J.E., Deane, F.P., Podd, J.: The relationship between driving anxiety and driving skill: a review of human factors and anxiety-performance theories to clarify future research needs. NZ J. Psychol. **37**(1), 28–37 (2008)

11. Taylor, J.E., Deane, F.P., Podd, J.V.: Determining the focus of driving fears. J. Anxiety Disord. **14**, 453–470 (2000)
12. Blanchard, E.B., Hickling, E.J., Taylor, A.E., Loos, W.: Psychiatric morbidity associated with motor vehicle accidents. J. Nerv. Mental Dis. **183**, 495–504 (1995)
13. Taylor, J.E., Deane, F.P.: Comparison and characteristics of motor vehicle accident (MVA) and non-MVA driving fears. J. Anxiety Disord. **3**, 287–298 (2000)
14. Wald, J., Taylor, S.: Preliminary research on the efficacy of virtual reality exposure therapy to treat driving phobia. Cyberpsychology Behav. **6**, 459–465 (2003)
15. Lewis, E.J., Walshe, D.G.: Is video homework of benefit when patients don't respond to virtual reality therapy for driving phobia? Cyberpsychology Behav. **8**, 342 (2005)
16. Antony, M.M., Brown, T.A., Barlow, D.H.: Heterogeneity among specific phobia types in DSM-IV. Behav. Res. Ther. **35**, 1089–1100 (1997)
17. da Costa, R.T., de Carvalho, M.R., Nardi, A.E.: Virtual reality exposure therapy in the treatment of driving phobia. Psicologia: Teoria e Pesquisa **26**(1), 131–137 (2010)
18. Townend, M.: Cognitive therapy for driving phobia: two single case studies. Behav. Cognitive Psychother. **31**(3), 369–375 (2003)
19. Townend, M., Grant, A.: Integrating science, practice and reflexivity—cognitive therapy with driving phobia. J. Psychiatr. Ment. Health Nurs. **13**(5), 554–561 (2006). doi:10.1111/j.1365-2850.2006.00978.x
20. de Jongh, A., Holmshaw, M., Carswell, W., van Wijk, A.: Usefulness of a trauma-focused treatment approach for travel phobia. Clin. Psychol. Psychother. **18**(2), 124–137 (2011)
21. Walshe, D., Lewis, E., O'Sullivan, K., Kim, S.I.: Virtually driving: are the driving environments "real enough" for exposure therapy with accident victims? An explorative study. Cyberpsychology Behav. **8**(6), 532–537 (2005)
22. Kraft, T., Kraft, D.: Creating a virtual reality in hypnosis: a case of driving phobia. Contemp. Hypn. **21**(2), 79–85 (2006)
23. Hill, R., Bannon-Ryder, G.: The use of hypnosis in the treatment of driving phobia. Contemp. Hypn. **22**(2), 99–103 (2006)
24. Williamson, A.: A case of driving phobia treated with dissociative imagery. Contemp. Hypn. **21**(2), 86–92 (2006). doi:10.1002/ch.294
25. Pallanti, S., Bernardi, S.: Neurobiology of repeated transcranial magnetic stimulation in the treatment of anxiety: a critical review. International clinical psychopharmacology **24**(4), 163–173 (2009)
26. da Costa, R.T., de Carvalho, M.R., Cantini, J., Freire, R.C., Nardi, A.E.: Demographics, clinical characteristics and quality of life of Brazilian women with driving phobia. Compr. Psychiatry **55**(2), 374–379 (2014)
27. Ku, J.H., Jang, D.P., Lee, B.S., Lee, J.H., Kim, I.Y., Kim, S.I.: Development and validation of virtual driving simulator for the spinal injury patient. Cyberpsychology Behav. **5**, 151–156 (2002)
28. Hatakka, M., Keskinen, E., Gregersen, N.P., Glad, A., Hemetkoski, K.: From control of the vehicle to personal self-control; broadening the perspectives to driver education. Transp. Res. Part F **5**, 201–215 (2002)

Using a Prospect Theory Approach to Studying the Car-Following Model

Mohammadadel Khodakarami, Yunlong Zhang, Bruce X. Wang, Mohammadali Shirazi and Maryam Shirinzadeh Dastgiri

Abstract Car following is a fundamental traffic feature that has been widely studied in literature using vehicles' speed and acceleration. This study investigates car following from an entirely different perspective, a psychological approach based on the Prospect Theory (PT). PT is a behavioral economic theory that explains human reaction under risk situations. Since car following can be regarded as a risk containing task that addresses the need for balancing safety with travel time reduction, PT is an ideal approach to model car following. Employing PT can provide a spectrum of probabilistic locations of following vehicles in contradiction with traditional methods that define the exact position. This study presents a sensitivity analysis in order to validate the results and calibrate PT's parameters. The results reveal that PT generates similar probability distributions to the simulation scenarios that proxy the space headway in real situations.

Keywords Car following model · Prospect theory · Risky Decision-Making

1 Introduction

A car-following model determines the properties of the following vehicle such as spacing, speed, and acceleration. The car following phenomenon has been researched for a long time; one of the first models dates back to the early fifties when Pipes (1953) offered his model by defining a safe distance between the leading and following vehicles. Car following models are now fundamental components of every microscopic simulation program since the base of any microscopic traffic simulation is determining the drivers' behavior while following other vehicles and changing lanes. There are also more elaborate car following models such as

M. Khodakarami (✉) · Y. Zhang · B.X. Wang · M. Shirazi · M.S. Dastgiri
Zachry Department of Civil Engineering, Dwight Look College of Engineering, Texas A&M University, College Station, USA
e-mail: adel.khodakarami@gmail.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_25

stimulus response models that similarly use relative speed between the following and leading vehicles to calculate the exact reaction of the follower.

Although the aforementioned models seem to work very accurately, they have several limitations. First, they usually depend on some sensitivity parameters that are difficult to measure, and also are likely to vary by traffic fluctuation. Second, car following as the common experience of driving is different in reality from the models that have been developed so far. In other words, drivers are not aware of their exact distance and speed difference from the predecessor vehicles. Instead, they act based on a more general perception of the leading vehicle; moreover, they choose a range of attitudes instead of a definite predictable reaction. Hence, as Ceder [1] mentioned: "Drivers do not completely follow any deterministic behavior". The problem of unpredictability was the main motive for the current study to enable simulating the driver decision making process as close to reality as possible. Accordingly, the authors adopted the Prospect Theory (PT) that originated as an econometric model for modeling decision making under risk to model car following.

Since car following is a humanistic decision making process in which the driver should always be concerned about the space from the lead vehicle, it contains some risk related to driver behavior. Driving close to the lead vehicle increases the risk of collision while driving far behind the lead vehicle increases the travel time. Therefore, each driver should balance her utilities from travel time reduction and safety increase by selecting her position in a range of distance behind the lead vehicle where the former utility results in risky driving and the latter in conservative. This study investigates the application of PT concept to model car following behavior. The developed model uses PT to extract the probability distribution of distance gap that drivers adopt when following another vehicle. To assess the model's result and its approximation to real drivers' behavior, it should be compared with real case distance gap distribution. However since authors did not have access for this comparison, some simulations were performed to replace the lacking real data. Several simulation scenarios were tried to mimic different conditions that might happen in real car fallowing situations. In addition, although one could use the original parameters that were originally derived for economical decision making purposes, it is needed to calibrate these parameters for transportation specific applications. To this end, we conducted several sensitivity analyses and calibrated PT parameters to obtain an acceptable match between model and the simulation results.

The rest of study is organized as follows: section two provides a literature review on car-following models as well as PT application in transportation. Section three describes the methodology of PT application to model car-following. Section four and five represent the results and discussion, and finally section six provides the study conclusion and suggestions for future studies.

2 Literature Review

In this section, a brief history of car-following models and PT applications in transportation-related literature is reviewed. This history constitutes the background for this paper, and thereby enabling the construction of the core concepts for the car-following model proposed in this paper.

2.1 *Car Following*

One of the famous earliest car following models was defined by Pipes [2]. He proposed a straight forward model by assuming a safe minimum distance between follower and leader vehicles. Later Chandler et al. [3], working at the General Motors Research Laboratories, proposed one of the more famous models for modeling the follower driver's acceleration. This model calculates the acceleration of the following vehicle based on the driver's perception of his/her relative speed to the leading vehicle. This model was the base for the next well-known model proposed by Gazis et al. [4] which is known as Gazis–Herman–Rothery (GHR) or the General Motors (GM) model. In fact, determining these parameters within the GHR model was the main cause of proposing different models by later researchers. This class of car-following models could be categorized under stimulus-response car-following models. Brackstone and McDonald [5] described the GHR model as having lost its popularity in the early nineties according to the large number of contradictions observed as to the correct values of m and l.

One car following model which appeared late and was considered state-of-art is a fuzzy logic based model, which holds the most similarity to the adopted concept in current study. This model usually starts with dividing the inputs into several overlapping subsets (fuzzy sets) to determine what variables belong to which description of a term. Accordingly, in this model some of the terms used are "very close", "close" or "far" based on the space or time headway. Based on these defined terms, logical operators could be used to equivalent fuzzy outputs with the reality of driving action [5]. Kikuchi and Chakroborty [6] implemented the GHR model in fuzzy-logic framework.

In this paper, the authors use the same concept of dividing the space headway into different regions with different terms. Each region represents some risk and shows the extent of the utility to be investigated based on the prospect utility.

2.2 *Prospect Theory*

Prospect theory which was first proposed by Kahneman and Tversky [7] is a decision making model that mainly accommodates the violations of Expected

Utility Theory (EUT) in handling risk situations. Four major issues that expected utility theory fails to address dealing with risky decisions are listed as below:

- Reference dependency: EUT assumes that people make decisions about different options based on their state of wealth. However, in reality people tend to consider the difference in wealth relative to a reference point.
- Loss Aversion: EUT considers the same weight for loss and gain however the decision makers' value function for loss is steeper than gain.
- Gain and loss diminishing effect: The outcomes' value in EUT is considered as a linear trend, whereas both positive and negative have diminishing patterns.
- Small probabilities: EUT assigns the same weight to different probabilities, while in fact people usually overweight small probabilities.

Accordingly, EUT fails to accurately count for uncertain and risky decision making. Thus the individual's choices from EUT deviates from what they choose in reality. Prospect theory, was motivated to address the above mentioned shortcomings with EUT. Tversky and Kahneman [8] later applied some minor changes to a PT model and developed Cumulative Prospect Theory (CPT). The changes are mostly related to the probability weighting function and do not affect the basis of PT. In continuation of this paper the CPT is used instead of PT.

Employing CPT has shown a superior performance when predicting the decision maker's behavior in risky conditions compared to EUT in many research areas especially in economics. This method has recently been used in the transportation engineering area and has proven capable of addressing many problems concerning travelers' decision making under uncertainty. For instance, Xu et al. [9] used CPT to model travelers' route choice behavior under uncertainty. After analysis of a survey, they determined that even when using original values of CPT parameters, CPT can provide much better results than EUT. In addition, they tried to estimate CPT parameters in the context of travel behavior, and the results achieved after using estimated parameters showed significant improvement. Schwanen and Etta [10] used CPT to analyze the employed parent behavior when faced with an uncertain network while collecting children from the nursery at the end of a weekday. A series of results obtained from a stated preference survey revealed some violation of EUT.

Hamder et al. [11] investigated the application of PT to develop a car following model. In their model, the following driver is assumed to have a normally distributed perception of the leader's speed. They linked the following driver's behavior to the objective collision probability and based their research on a nonstationary sub-model with mixed behavior probability that basically assumes different mechanisms for the follower. Gain and losses in this research are assumed to be the gains and losses in terms of speed that is calculated from their previous acceleration state. They also limited the gains corresponding to the maximum desired velocity and the maximum loss to nonnegative velocity. In a later study [12], they improved their previous work by expressing loss and gains as functions of acceleration and deceleration. Then they modeled the acceleration as a function of three factors: the expected acceleration value, the variance of the acceleration distribution, and a correlation time. They assumed that the acceleration being distributed as a continuous logit model with a distribution conditioned on: follower speed, relative speed between the leader and follower, and the space gap between the two. Having access to vehicles trajectory data from Next Generation Simulation (NGSIM) project, they used nonlinear optimization process based on genetic algorithm as a tool to calibrate model parameters. Although they were not able to derive the parameter value distributions due to the lack of data, their results show high heterogeneity between the drivers' behavior. Also because their proposed model is a stochastic process, compared to Intelligent Driver Model (IDM) [13], the PT model experienced higher levels of fitting error. In the current study, the loss and gains are described in term of distance gaps instead of velocity or acceleration/deceleration.

CPT is comprised of three main components: (1) a reference point, (2) value function and (3) weighting function. The reference point represents the cutoff point dividing loss and gain domains. The value function is an S-shape function that shows a diminishing pattern for the values in both loss and gain, and the weighting function is an inverse S-shaped function which overweighs the low probabilities and under-weighs the high probabilities. Figure 1 shows the general shapes of value and weighting functions.

The right and left side of value function represent the gain and loss domain, respectively. The equations of these functions are shown in Eqs. 1–3:

$$\Delta x_i = x_i - x_0 \tag{1}$$

$$\Phi(x_i) = \begin{cases} x_i^{\alpha}, & x_i \geq 0 \\ -\lambda(-x_i)^{\eta}, & x_i \leq 0 \end{cases} \tag{2}$$

$$w(p_i) = p_i^{\gamma} / [p_i^{\gamma} + (1 - p_i^{\gamma})^{\gamma}]^{(1/\gamma)} \tag{3}$$

where:

x_i decision variable of component i,
x_0 : reference point,
$\Phi(x_i)$ value function,
p_i probability of component i which should be cumulated toward the highest value for loss and cumulated toward zero for gain,

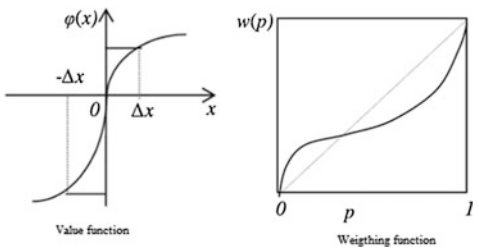


Fig. 1 Value and weighting functions

$w(p_i)$ weighting probability,
α, η, γ Prospect Theory parameters
λ loss aversion coefficient

In above relations α, η show the diminishing sensitivity of the value function and λ represents the amount of people sensitivity to loss rather than gain. In weighting function γ reflects how people overweigh or underweight the probabilities. All four parameters of α, η, γ and λ should be calibrated based on different context of application; but, the result of other researchers show that the estimated new parameters typically are not very far from the original values. For applying the CPT, the first alternative should be divided into several components and then the probability of occurrence for each of these components should be determined. To further the proposed relation for CPT as a final utility for w whole alternative would then be available. Therefore, the final utility for each alternative composed of (m + 1 + n) component is calculated based on below relations (m and n are, respectively, the number of components in loss and gain):

$$\pi^+(p_i) = w(p_i + \ldots + p_n) - w(p_{i+1} + \ldots + p_n), 0 \le i \le n \tag{4}$$

$$\pi^-(p_{-j}) = w(p_{-m} + \ldots + p_{-j}) - w(p_{-m} + \ldots + p_{-j-1}), -m \le -j \le 0 \tag{5}$$

$$U(x,p) = \sum_{i=0}^{n} \varphi(\Delta x_i)\pi^+(p_i) + \sum_{i=-m}^{-1} \varphi(\Delta x_i)\pi^-(p_i) \tag{6}$$

Now by having each alternative utility the probability of each utility selection can be calculated based on the appropriate logistics model, that depends on the distribution of errors.

3 Methodology

This section presents the main idea of the modeling approach. The general scheme of this model is based on the quality of driver perception of the car-following process and the reaction that she/he adopts in response to the leader. In this study, the reaction is determined based on the relative distance from the leading vehicle instead of other stimuli-like acceleration.

The main assumption of this model is that drivers do not have an exact perception with respect to speed, acceleration and space headway. Therefore, their response is different from what traditional models like GHR [4] predict. Thus, this study generates a distribution of responses instead of an exact response. For example, common sense dictates that drivers cannot know the exact distance between their vehicle and the lead vehicle nor can they ascertain the exact speed of the leader. Hence, drivers attempt to choose a distance as close as possible to the lead vehicle, while maintaining a safe distance. There is a perception of safe distance for drivers to keep their clearance with the lead vehicle that in this study is assumed as a ratio of their vehicle length. Therefore, the distance between two vehicles follows a region or extent instead of a certain point and subsequently the distribution over this region replaces the equation of a point location. Now the problem changes to finding this probability distribution and evaluating how the proximity to the leader vehicle affects distance distribution.

According to the described assumption in the above paragraph, people generally desire to lessen their travel time even if only to a very tiny extent; therefore, in an ideal situation where all vehicles could drive at the same speed and all the dynamics of leader and follower are the same, all drivers like to drive bumper to bumper. However, considering the real-world situation, drivers always need to keep a minimum safe distance to avoid any kind of unexpected collisions (one car length for every 10 mph of car speed in the US). Clearly, there is a contrast between the desire to shorten travel time and to maximize safety. The combination of these two components gives the problem a risky decision making nature addressed by CPT.

To apply CPT, first the terms of risk and risky decision must be defined. The distance between follower and leader has been divided into three segments. The first segment with length of "L1", represents the minimum space needed between the two vehicles, i.e., the "close" region. This means that exceeding this distance could increase risk of collusion while an unwanted event can cause the leader to stop or decelerate quickly. The second segment with length of "L2" represents the region that the majority of drivers prefer when following a lead vehicle; hence, it is known as the "desired" region. Finally the third segment which can have an unlimited length of "L3" starting from the last point of the desired region and is called the "far" region. By introduction of these three regions, one can conclude that all drivers like to keep themselves in the desired region and avoid the close (L1) and far (L3) regions. These regions reflect the property of reference points in CPT that distinguishes gain domain from loss domain. Accordingly, in current model, we are faced with twos reference point in contrast to typical problems that work with only one reference point. The reference points are the junction points of desired region with close and far regions. Figure 2 describes these regions.

Based on this classification, the ideal desire of each driver, as shown in Fig. 2, is staying at the first point of desired region or where d = L1. The problem comes from the rough ability of the drivers to adjust their location exactly at the desirable point; therefore if they unknowingly enter the distance equal to d, they are in the close region that is not safe. Assuming the range of driver error in adjusting his/her exact location defined by a length "L4" and named as the "tolerance" extent, drivers will choose the policy which maximizes their utility or level where they feel safe.

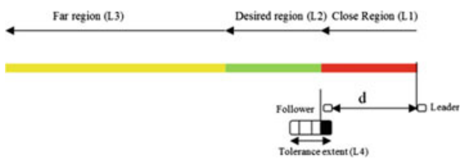


Fig. 2 A sketch of the three different regions

According to CPT the length of tolerance extent can be divided into several different parts or cells. Each of these cells serve as a CPT component and their probability is assumed to follow a uniform distribution. For example in the case shown in Fig. 2, there are 4 cells, where cell 1, colored black, represents the part of a tolerance extent placed in the close region and possess a negative utility with 25 % occurrence probability (the ratio of cell to the tolerance extent length). Hence, based on the number of cells, these probabilities can be determined and would be equal for all cells. In further CPT applications, both close and far regions are considered as domain of loss. The desired region is understood as the gain domain; therefore, the two endpoints of the desired region are taken as the reference points where in the desired region, utility increases by reducing the distance between the leader and the follower.

Subsequently, the problem can be modeled like a decision-making problem (on the desirable locations), in which the presence of tolerance affects the probability and risk of each alternative. To model this decision making a CPT method was used and after calculation of the utilities, a logit formula was employed to obtain the probability of different alternatives (positions behind the leader vehicle). The results that are distribution of follower distance from leader vehicles are presented in next section.

4 Developing a CPT Model for Space Headway

This section provides the results of the space-headway distribution estimation based on the CPT model.

The parameters that should be calibrated for CPT application include: α, η, γ and λ. In other transportation studies, different values for these parameters have been estimated. Despite this variation, these parameters except for λ kept their proximity to original values. Table 1 shows the original values and those adopted by Xu et al. [9].

Since microscopic headway data are hard to get, these parameters may not be easily estimated. In this study, we investigated space headway distribution first by adopting the values by T&K and Xu. Sensitivity analysis of the parameters is used also to test the sensitivity of the model to other values of parameters like L1 and L2 in order to analyze their effect on the car following behavior.

Table 1 CPT parameters values

Parameter	Original value	Transportation value
α	0.52	0.37
η	0.52	0.59
γ	0.74	0.74
λ	2.25	1.51

The other parameters that are used in this model with a variety of values are L1 and L2, which incorporate the reference values of the model. L1 comprises the concept of minimum safety distance that is approximated through Pipes [2] model as follows:

$$d_{MIN} = L_n \left[\frac{\dot{x}_{n+1}(v)}{10} \right] + L_n \qquad (7)$$

where:

L_n average length of vehicles,

$\dot{x}_{n+1}$ speed of leader vehicle in mph at time t,

In addition to determining L1, the desired region length L2 should also be determined. These factors depend on the real behavior of travelers and could be estimated from separate study that surveys the drivers' behavior. However, in this paper a set of values have been considered and tested to investigate the sensitivity of parameters. Figures 3 shows the graphs of distance gap distribution obtained from different combinations of parameters.

The above results represent the distribution of space headway between two following vehicles achieved for different combinations of CPT parameters. The value of each parameter is shown on the plots except for the values for L1, L2 and L3 in top plots that are L1 = L2 = L3 = 80 ft. As it is clear from the above plots, the original parameters and transportation related parameters show some difference. The most significant difference is caused by varying α that can change the distribution shape. However, two other parameters λ and η do not influence the distribution. Parameter α largely influences the gain domain that is interpreted by being located in desired region. The weak effect of λ and η is also related to this

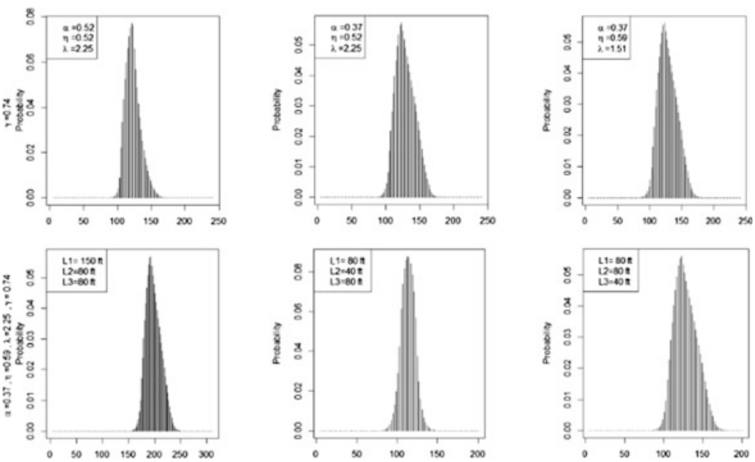


Fig. 3 Vehicle-spacing distributions between a following and leading vehicle. The graphs represent values of CPT parameters and changes in space headway distribution

significant effect. The bottom row of plots illustrates the change in distribution due to change in L1, L2 and L3. It looks like increasing L1 and L2 could only shift the distribution to right and left along *x* axis without any meaningful change into the distribution shape. Thus, those could be seen as a location factor. On the other hand, L2 changes the shape of distribution and acts as a shape factor. Among the above plots, the fifth plot seems to be the most appropriate distribution according to comparison with the simulation results in next section where the lower L2 length shows a better fit.

After identifying the model's parameter, we need to validate if the obtained distance-headway distribution matches with the reality. To this end, we need to use field data and find the distribution, or a implement a simulation. In this study, since we did not have access to the field data, CPT model results were validated against some microscopic simulation tests where the results are presented in the next section.

5 Simulating the Space Headway Distribution

This section presents the simulation results used to develop a space headway distribution in order to compare and assess CPT results. For this purpose, several simulation scenarios were generated in VISSIM. In each scenario the space headway of two consecutive vehicles was recorded for an hour of simulation. Then, generating a large number of headways, their probability density functions (pdf) were estimated and plotted using Kernel Density Estimation [14]. Six different scenarios were tested for simulation wherein each vehicle moved in a single lane loop and their space headway was recorded for an hour of simulation. The first two minutes of observation in each case was considered as a simulation warm up and was excluded from simulation results. The six simulated scenarios were as follows:

- *Scenario 1*—Two vehicles traveling at 70 mph desired speed where one is following the other one with initial headway of 1 s.
- *Scenario 2*—Two vehicles where the leader is traveling at 70 mph and the follower at an 80 mph desired speed.
- *Scenario 3*—Fifteen vehicles moving at desired speed of 70 mph with initial headway of 2 s.
- *Scenario 4*—Fifteen vehicles moving at desired speed between 60 and 70 mph with initial headway of 2 s.
- *Scenario 5*—Fifteen vehicles moving at desired speed of 70 mph with initial headway of 1.5 s.
- *Scenario 6*—Fifteen vehicles moving at desired speed between 60 and 70 mph with initial headway of 1.5 s.

Using the kernel estimation the distance distributions for all scenarios are plotted in Fig. 4. We used the method introduced by Scott [12] to find the optimal

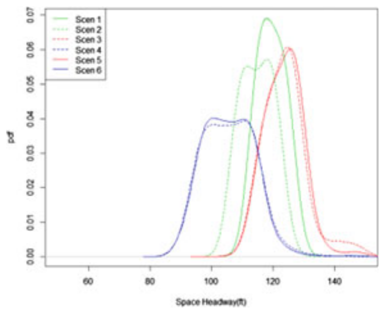


Fig. 4 The PDF comparison of different scenarios

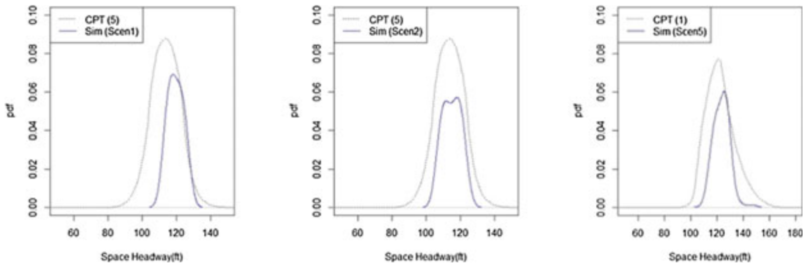


Fig. 5 Comparison between the CPT method and simulation results

bandwidth for the kernel estimation. This plot shows scenarios at a fixed 70 mph are slightly different from the ones between 60 and 70 mph as the desired speed.

At this point, the plotted distribution density from simulation, and the distribution mass from the CPT could be used for comparing these two methods and evaluating CPT results. Figure 5 shows the plot of distributions form CPT against simulation.

The legends in Fig. 5 represent simulation scenarios and CPT parameter sets that relate to the distributions. Figure 5 represents significant similarity between CPT results and simulation tests. The slight difference could be attributed to the discrete nature of CPT probability distribution against simulation that is a continuous pdf. These results state that the distribution from CPT, a psychological-based model, mimics the results from the simulation that with a physical logic. The slight scale differences could be resolved by tuning up the CPT parameters using field study results. A parameter calibration is conducted and the result is presented in Fig. 6.

Figure 6 illustrates a fine match between calibrated CPT and simulation space headway distribution. To find this match several CPT parameters were examined and the combination of $\lambda = 2.4$, $\eta = 0.78$, $\alpha = 0.52$, and $\gamma = 0.75$ was the final selection of parameters. Along with this parameter setting, to obtain distribution that is more continuous, the step size of CPT algorithm was decreased by 50 %, so it provides more smooth distribution. We also conducted a chi-square test to

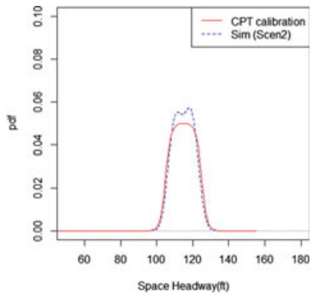


Fig. 6 Match of calibrated CPT model with simulation results

evaluate the similarity of these distributions more exactly. The chi-square p-value was very close to one indicating almost an identity between two distributions.

In spite of very solid solution from chi-square, the calibration of parameters would be more meaningful while using the real data, this example only shows the significant potentials of CPT model to be adopted in real world modeling.

6 Evaluating the Performance and Application of CPT

According to the obtained results in the previous section, it is clear that CPT is capable of describing the car following distribution, however it is sensitive to its parameters and there is a need for calibration. As it is mentioned before, CPT works according to the logic of decision makers while facing risky situations. Thus, it is reasonable to assume, due to different mindsets of the drivers to be risk averse or risk seeking, the parameters of CPT need to be calibrated for different personalities and conditions. Indeed, this is a valid assumption and whenever facing humanistic situation the effect of individual behavior should be taken care of. Thus, to apply the CPT there is a need for another study that categorizes drivers according to their risk taking attitude and then calibrates the parameters for each group. In other words, to apply CPT a set of parameters should be calibrated to fit different users characteristic.

The other issue related to CPT is the method of its application in micro simulation. To that end, after determination of different parameter sets, one can generate different probability distributions for different parameter sets. Now to simulate the car following distance, the micro model can take a random sample from the intended distribution for the specific parameter set. Thus, the car following model transforms into random samples of space headway distribution. To involve characteristics of different drivers, another sampling should be done on driver type distribution and then headway samples be taken from the distribution for each type.

7 Conclusion

This study studies a new car-following model using Prospect Theory to take care of drivers' decision-making under risky conditions. Traditional car-following models, develop their relations based on drivers' perception of the exact relative speed or space; however this perception is not completely valid in reality and drivers do not have exact perception of other vehicles location and speed. The proposed model in this paper instead uses a more general concept of spacing and offers the space division between leader and follower vehicles into three regions close, desirable, and far. These regions represent the desirable distance between follower and leader vehicles. Using Prospect Theory, this study's model calculates the probability distribution of distance between follower and leader vehicles based upon the utility of placing in each region. This study conducted several tests to compare the space headway distribution from CPT with micro-simulation distribution. The results display a fine similarity between these distributions. Calibrating the CPT parameters, the results illustrate a more significant match between CPT and simulation distribution. The achieved result validates the application of CPT for modeling the space headway.

CPT parameters are highly correlated to the drivers' characteristics like age, sex, education and other socio-economic attributes. In addition to drivers' characteristic, weather and traffic condition change the drivers risk taking behavior as well. Hence, as recommendation for further studies, the authors suggests calibrating CPT parameters considering different personal and environmental factors affecting drivers' behavior. Finding a range of parameters for different conditions will allow for applying CPT in micro simulation models.

References

1. Ceder, A.: A deterministic traffic flow model for the two regime approach. Transp. Res. Rec. **567**, 16–30 (1976)
2. Pipes, L.A.: An operation dynamics of traffic dynamics. J. Appl. Phys. **24**(3), 274–281 (1953)
3. Chandler, R.E., Herman, R., Montroll, E.W.: Traffic dynamics: studies in car following. Oper. Res. **6**, 165–184 (1958)
4. Gazis, D.C., Herman, R., Rothery, R.W.: Nonlinear follow the leader models of traffic flow. Oper. Res. **9**, 545–567 (1961)
5. Brackstone, M., McDonald, M.: Car-following: a historical review. Transp. Res. Part F **2**, 181–196 (1999)
6. Kikuchi, C., Chakroborty, P.: Car following model based on a fuzzy inference system. Transp. Res. Rec. **1365**, 82–91 (1992)
7. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. Econometrica **47**(2), 263–292 (1979)
8. Tversky, A., Kahneman, D.: Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertainty **5**, 297–323 (1992)
9. Xu, H., Zhou, J., Xu, W.: A decision-making rule for modeling travelers' route choice behavior based on cumulative prospect theory. Transp. Res. Part C **19**, 218–228 (2011)

10. Schwanen, T., Ettema, D.: Coping with unreliable transportation when collecting children: examining parents' behavior with cumulative prospect theory. Transp. Res. Part A **43**, 511–525 (2009)
11. Hamdar, S., Treiber, M., Mahmassani, H., Kesting, A.: Modeling driver behavior as sequential risk-taking task. Transp. Res. Rec. J. Transp. Res. Board, 208–217 (2008)
12. Hamdar, S.H., Mahmassani, H.S., Treiber, M.: From behavioral psychology to acceleration modeling: Calibration, validation, and exploration of drivers' cognitive and safety parameters in a risk-taking environment. Transp. Res. Part B Methodol. **78**, 32–53 (2015)
13. Treiber, M., Hennecke, A., Helbing, D.: Congested traffic states in empirical observations and microscopic simulations. Phys. Rev. E **62**, 1805 (2000)
14. Sheather, S.J.: Density estimation. Stat. Sci. **19**(4), 588–597 (2004)

Enhancing Drivers' Situation Awareness

Andreas Gregoriades and Maria Pampaka

Abstract This work describes the evaluation of a prospective situation awareness enhancement system (SAES) that exploits augmented reality through a head up display. The specification of the proposed SAES is based on domain knowledge from the literature. Herein, a method is described that utilizes the benefits of a modular simulator to model the design of a prospective SAES. The situation awareness assessment method used is based on the Situation awareness global assessment technique (SAGAT) and utilizes data obtained from the simulator that replicates the infrastructure of a road network in Cyprus. The simulator mimics the functionality of a prototype SAES using a number of candidate visualization metaphors, that simulate a number of candidate head up display designs. The paper describes the process of assessing the situation awareness of drivers that use a prototype smart SAES, through a series of experiments in a virtual reality CAVE. The effectiveness of the SAES is tested in a between-groups research design.

Keywords Situation awareness · Situation awareness enhancement system · SAGAT

1 Introduction

Road accidents are a daily hazard worldwide [1]. According to [2], around 1.2 million fatalities and more than 50 million injuries occur in roads worldwide every year. Given the current trends, the accident fatalities are projected to become the second most common cause of death in 2020 if no drastic measures are taken. Road accidents have many causes including demographic, infrastructural and political

A. Gregoriades (✉)
Cyprus University of Technology, Limassol, Cyprus
e-mail: andreas.gregoriades@cut.ac.cy

M. Pampaka
University of Manchester, Manchester, UK
e-mail: maria.pampaka@manchester.ac.uk

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_26

factors [2]. From another perspective, accidents are associated with human error, that is the human activity (or absence of activity) that leads to incorrect system behavior [3, 4]. This may occur due to human beings' physical, perceptual and cognitive limitations [5] and is directly related to visual attention [1] and workload [4]; all these factors, in turn, affect situation awareness (SA). The analysis of accidents due to human error can be carried out from two perspectives: the designers' and the users'. The former addresses the system designing flaws that hinder human activity due to usability problems whereas the latter analyzes internal cognitive processes of human operators to identify decision making bottlenecks caused by reduced SA [6, 7] induced by increased workload. High levels of workload degrade drivers' concentration, information processing, SA and decision making, leading to increased errors [8]. Endsley et al. [7] warn socio-technical system designers of the importance of maintaining SA in complex systems design and draw the attention on issues that could inhibit SA. One of the most important strains of SA is information overload. Too much information at any point in time hinders human operators' SA. Overloading divides the decision maker's attention among numerous stimuli resulting in increased demand for cognitive resources. When too much information is available information scanning capability is reduced, because of attentional tunneling [6, 7], resulting in decision makers locking their attention on certain aspects of the environment whilst ignoring other important stimuli.

SA constitutes a critical factor in road safety. In particular, it provides the driver with the ability to anticipate events given perceived driving and environmental conditions. Smart driver assistive technologies belong to a class of systems that are used to alleviate accident risk by either reducing driver workload or enhancing driver situation awareness. Such systems aim to draw drivers' attention on critical information cues that help maintain a good level of situation awareness that in turn improve decision making. However, in some cases, these systems could have a negative effect due to the extra information load they incur to the driver which could lead in an erroneous act. The focus of this study is on the evaluation of a proposed user interface designs of smart SAES that aim to improve SA of drivers. The SA analysis is based on phenotype driver behaviour data analysis through experiments in simulated Virtual Reality (VR) settings [9]. The use of a driving simulation is inevitable since it is difficult to eliminate confounding effects on control measures in field experiments. Hence, the ability to test driving behaviour under controlled settings. However, low cost driving simulators do not provide a sufficient level of realism to analyse SA, and unrealistic conditions may affect the driving behaviour which could influence the validity of the study. The method proposed herein demonstrates the use of a custom made driving simulator that exploits 3D modeling tools in a VR CAVE facility. The research questions are, thus: Does any of the two SAES designs improve driver SA? Which SAES design is better among the two? Which features of the SAE design need refinement?

The paper is organized as follows. Next section addresses the literature in the areas of SA and SA assessment. Next the design of the experiment for the assessment of drivers' SA is described. The analysis of the experimental data is then described and the emerging results are explained. The paper concludes with a brief discussion of the implications of the findings.

2 Literature Review

2.1 Situation Awareness

SA constitutes a critical factor in safety critical systems. In transportation, it provides the driver with the ability to anticipate events given perceived driving and environmental conditions. SA defines the process of perceiving information (level 1) from the environment, comprehending its meaning (level 2) and projecting it into the future (level 3) [6, 7]. From the perspective of driving, SA is categorized as spatial, identity, temporal, goal and system awareness [10]. Spatial awareness, addresses the appreciation of the position of all relevant environmental features; identity awareness addresses knowledge of salient environmental items; temporal awareness addresses knowledge of the changing spatial picture over time. Goal awareness addresses the navigation plan to the destination or at a lower level, the maintenance of speed and direction to conform to the navigation plan, or the need to maneuver the vehicle within the surrounding traffic. Finally, system awareness addresses information within the larger driving environment which is the road network. Therefore, a goal-oriented model of driver behavior, encompasses strategic, tactical, and operational goals of driving [10].

Several studies examined ways to enhance SA in driving context. Smart driver assistive technologies have been developed to alleviate accident risk by either reducing driver workload or assisting driver attentiveness. Examples include adaptive cruise control, collision notification, driver monitoring, traffic signal recognition, night vision, lane departure warning systems and blind spot monitoring. Such systems aim to draw drivers' attention on critical cues that improve their decision making. However, they only provide limited support to SA since they address isolated factors affecting it and in some cases with negative effect due to the extra information load they incur to the driver. The first step in improving drivers' SA is to enhance their capability of perceiving and interpreting traffic and environmental conditions, which constitute level 1 and 2 SA processes. However, such smart systems facilitate level 3 SA for navigation, which might decrease drivers' attention, due to secondary task execution, that could lead to reduced level 1 SA. This could undermine attention to operational or tactical driving activities (e.g. braking, lane changing, gap acceptance etc.). These issues have been examined by various researchers. Walker et al [11] examined the effects of different forms of non-visual smart driver assistive technology feedback on driver SA using a probe-recall method and suggested that adding auditory feedback increases driver SA compared to vision only feedback. Other studies identified associations between SA and driving performance using secondary tasks analysis. Specifically, [12] showed that an adaptive cruise control system reduced workload and improved overall SA when the driver was using a mobile phone under typical driving conditions. However, the cell phone conversation decreased level 2 and 3 SA while level 1 SA was not affected significantly [12]. A follow up study by [13] examined the effects of in-vehicle navigation aids and reliability on driver SA and performance in a simulated navigation environment. Results revealed that navigation information improved driver SA and performance with improved operational and strategic behaviors, while unreliable and task-irrelevant information degraded SA and performance.

However, to be able to evaluate the effect of a smart driver assistive technology on SA it is essential to employ a suitable SA assessment technique. The following section addresses this subject.

2.2 SA Assessment Methods

According to Endsley [6, 7] the methods for assessing SA in dynamic systems, is categorized into subjective and objective. The former uses subjective rating of drivers after their engagement with a situation, e.g. the Situation Awareness Rating Technique (SART) [14] also includes observer subjective rating techniques. SART focuses on generic, overall task characteristics rather than on specific task elements [15], involving a series of questions on three dimension of SA, namely demand on attentional resources, supply of attentional resources and the understanding of a situation.

Objective measures, on the other hand, use actual assessments of the situation and can be classified into: freeze probe recall techniques (e.g. SAGAT: Situation Awareness Global Assessment Technique), real-time probe (e.g., SPAM: Situation-Present Assessment Method), and real-time probes based on SAGAT, process indices (e.g., eye-tracking), or performance measures. SAGAT [6, 7] is a dynamic query technique that inquires participants' recent memory of the situation. In SAGAT, driving situation is freezed, all screens become blank and questions are presented to the participant. The more correct answers, the better the operator's SA. Because the SAGAT freezes the display frequently, it is criticized as being intrusive to performance [7]. Alternatives such as Real-time probe techniques have been suggested for objectively and unobtrusively measuring SA in highly dynamic operating environments [16]. In SPAM [17], SAGAT-like queries are given to the operator, but information remains in view and response latency is used as the primary measure. Real-time probes based on SAGAT [7] only involve natural verbal communications between an experimenter and an operator without any prior warning signal, which is given in the SPAM. However, conversation with an experimenter during driving is likely to distract participants and could lead to deterioration of their SA due to overloading. Hence, SPAM measures are dependent on workload and spare capacity. The fact that subjects have access to the information and can refer to it to answer could be interpreted as an indicator of the spare workload capacity more than a SA indicator. Another limitation of SPAM and real-time probes based on SAGAT is that the questions and answers are verbal: addressing spatial representation of traffic is more difficult verbally than diagrammatically.

Hazard perception is considered as an alternative technique for evaluating SA [18]. In this method researchers have been using filmed traffic situations for a hazard perception test and asked participants to detect a traffic hazard, using a button [19], or touch screen [20]. However, requiring such an active response from participants is different from the natural driving environment. Therefore, implicit performance measures are considered more appropriate. If participants maintain good SA, they are expected to cope with the situation effectively. Empirical research has shown that drivers' hazard perception has been found to be negatively associated with drivers' accident records [8]. This, and the finding that SA assessment is also associated with the coping level with hazard events [17], makes the intervention during a hazardous event, such as those in SPAM and real-time SAGAT, impossible since it affects driver performance through increased workload induced by the distraction.

Therefore, in this study due to the richness of information provided by the proposed SAES, the SAGAT technique was used to avoid driver distraction and to improve the scope of the inquiries during interventions by addressing issues relating to projection and surrounding traffic conditions.

3 SAES Evaluation Method

The SA assessment method used in this study is based on SAGAT and is undertaken in a VR cave facility. The goal of the proposed SAES is to improve road safety by reducing the likelihood of accident through improved SA. The specification of the SAES designs based on domain knowledge from the literature and input from subject matter experts. Therefore, SA guidelines as specified by [6] are expressed in information requirements, visualization metaphors and interaction styles which are specified in terms of the functional requirements of the SAES. The overall design of SAES is based on Design Science methodology [21] which embraces the use of feedback loops until the desired artefact qualities are obtained. Initially a high level design of the system is grafted and subsequently specification is refined into functional specifications. Therefore, for any of the two SAES designs to be useful, the level of SA improvement had to be significantly better than the control conditions (no use of SAES). If this initial SAES requirement is not met then the specification of the candidate SAES is altered and the process is repeated until this SA requirement is satisfied. The next step in the process is the specification of the test scenarios, based on which the SAES is going to be evaluated.

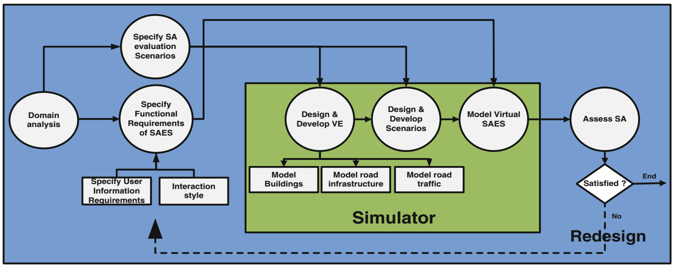


Fig. 1 Simulation-based SAES validation method based on design science

Grounded within the problem to be analyzed, the goals of the desired virtual environment (VE) are set. Accordingly specifications of the VE to be used for the evaluation of the artefact are also set. During this stage a tailor made VR driving simulator is customized based on the above goals, to model the research question. The customization of the simulator is performed in three steps (Fig. 1). The first step addresses the development of the test environment in terms of buildings, infrastructure and traffic flow. The second step concentrates on the modelling of the scenarios, including atypical events in the simulation that would stress test the subjects in the experiment. The third step includes the modelling of the virtual version of the SAES designs under scrutiny. Prior to its use, the VR simulator needs to be validated against a number of factors such as realism, to guarantee the correctness of the SA assessment. The assessment of SA is refined into phenotype behaviors that can be monitored in a driving simulator. Phenotype driving behaviors are monitored and logged into the simulator's database. The logged observations from the simulation are collated into a single metric that corresponds to the assessed SA. The SA level is compared against the desired SA level. If the minimum level of SA is not satisfied then the virtual artefact under scrutiny needs to be redesigned. The process is repeated until the SA level attained by the drivers is satisfactory. Two SAES designs were evaluated using this methodology. The first SAES design uses the paradigm of radar on a head-up display and aims to provide peripheral vision and threat prioritization, with the capability of informing drivers of traffic situation in surrounding roads. Hence it supports level 1 and 2 of the SA model. The second design addresses the need to prioritize information based on risk level and aims to warn drivers of hazardous events that are expected to emerge from side roads and are not yet visible, using the metaphor of pointing arrows on the HUD. This design is similar to the Mercedes blind spot assist system. Throughout the experiment the simulator monitors and logs the phenotype observations in a database. These are map-matched based on road section that the observation was made [4].

4 Data Collection

During the SAES evaluation stage, 17 participants were involved, each spending on average 90 min to complete the experiment in the VR CAVE. Participants were from the local population and had a valid driver's license and 20/20 vision or wore corrective glasses or lenses. Throughout the experiment, physiological, objective and subjective information is obtained from the driver, such as, electroencephalography, driving data, and questionnaires. During the experiment, participants had to drive on a road divided into 63 sections, under the three conditions. The simulation was stopped at several points to inquire participants' understanding of the situation according to the SAGAT method. Collected data from the experiments underwent cleansing and pre-processing prior to being analyzed statistically.

SAGAT related observations were made using the simulation freezing method. Participants had to answer question regarding their current situation just before the simulation freezing. Throughout the experiment, each participant encountered 3 freezing events that corresponded to hazardous events that required high SA. These events occurred in road Sects. 17, 32 and 53. In each event, perceived-user-situation was compared against the actual-situation using a number of features that described the situation, such as vehicles in front or behind the host vehicle, current speed, current vehicle's lane, peripheral situation, imminent threats etc. The higher the dissimilarity between the two the lower the SA level of the participant for that simulation freezing point. Drivers' workload at each time step of the simulation was measured using a combination of physiological and phenotype measures, such as Electroencephalography (EEG) and lateral deviation. Hence, the higher the workload the lower the SA of participants. Similarly, the higher the lateral deviations the lower their SA. Workload is assessed using a non-invasive EEG (Fig. 2) such as NeuroSky's brain computer interface. The algorithm implemented in the device measures the participant's attentional resources consumed while performing a task. The assumption here is that if the participants' attention readings are high near the points where the situation is intense, such as in hazardous events, this indicates that the user is utilizing high cognitive resources to deal with the situation, hence the workload is high. Therefore, this could indicate that the participant will not be able to perceive the situation correctly, due to limited available cognitive resources. This is manifested in missed informational cues, as a result of information tunneling. The underlying method of assessment is the measure of raw EEG brainwaves (Alpha,

Fig. 2 EEG workload assessment using non-invasive brain computer interface

Beta, Theta, Gamma and Delta) using three dry electrodes located on the left ear and forehead. Attention is calculated using an algorithm that takes into account all five bands of brainwaves. Collected data from the EEG was mapped with the rest of the data from the SAGAT questionnaires in order to make the comparison between the actual and perceived situation. The mapping was performed manually using road section ID at the freezing location as a reference point.

5 Subjective Evaluation of the Two Designs

In the subjective analysis, participants had to complete a questionnaire after their engagement with the two designs in the experimental settings. The evaluation of the two designs was based on the constructs listed in Table 1. Analysis of the results revealed that both SAES designs were perceived by the users as better than the control condition (i.e. without SAES). Specifically the post-experiment questionnaire addressed the following dimensions of each candidate design: features, user interface, ease of learning, system capabilities, usefulness, ease of use, and situation awareness. Each dimension was assessed on average by 5 questions on a 7 point response scale with a 1 (negative effect) to 7 (positive effect). To increase the

Table 1 The dimensions of the post-questionnaire constructs

ID	Questionnaire constructs
1.	Overall rating (general features) of the system (Rada/Arrows)
2.	User interface (Rada/Arrows)
3.	Learning (Rada/Arrows)
4.	System capabilities (Rada/Arrows)
5.	Usefulness (Rada/Arrows)
6.	Ease of use (Rada/Arrows)
General	
7.	The driving simulator is realistic
8.	The radar is better than the arrow system in helping me understand the situation around my vehicle
9.	The arrow system is better than the radar helping me understand the situation around my vehicle
10.	None of the two systems can help me have a better understanding of what is the situation around my vehicle
11. 12.	During the experiment, in which cases did you consult the radar system?
13.	During the experiment, in which cases did you consult the arrow system?
14.	What other features would you like to see in the radar system?
15.	What other features would you like to see in the arrows system?
16.	Which point(s) on the road was (were) the most risky? (Show these on the diagram below)

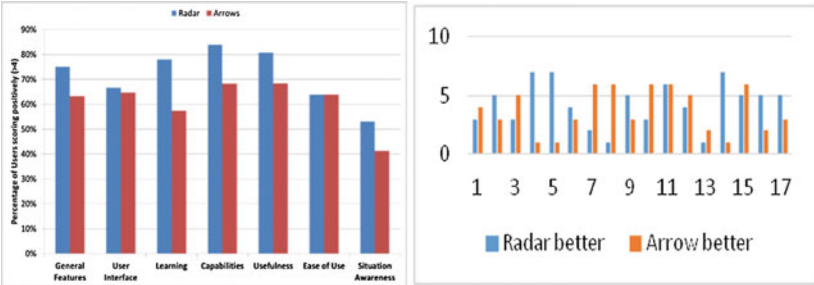


Fig. 3 Percentages of positive responses in each of the measured dimensions, by design (*left*) Question on whether radar design better that arrows and vice versa (*right*)

discrimination in the evaluators' judgment, participants' were asked to report the reasons for their choices and any interaction problems they had experienced with each design.

Responses regarding participants' perceptions on the different dimensions of the SAES were considered as a percentage of positive scores (i.e. scores of 4 and above). Figure 3 shows the percentages of positive responses (i.e. >4, or <4 for negatively worded statements) for each of the measured dimensions on which the two designs were evaluated. Based on this analysis, there do not seem to be noticeable differences in regards to user interface and ease of use. However, overall the radar design seems to have been perceived more positively than the arrows, especially in relation to learning, system capabilities, and usefulness. This might be attributed to the small size of the arrows that were popping up on the HUD windshield. Hence, participants did not perceive the system's response in some cases, which leaded in perceiving the arrows design less useful in comparison to the radar.

Participants' post-test response on radar or arrows design superiority questions reveal that radar was perceived as superior to arrows as shown in Fig. 3. This difference was statistically significant. Moreover paired sample t-test of the differences between the constructs of the questionnaire listed in revealed that the user interface of the radar design is superior to that of the arrows and this is statistically significant. This justifies the overall preference of participants towards radar design. Other constructs of the questionnaire highlight that the radar is better but this is not statistically significant. Finally, based on open responses from participants, in certain occasions the number of arrows that were present on the windshield exceeded two and hence the cues were becoming destructing rather than informative. On the other hand, the Radar design also had its shortcoming in terms of visualization of the threats. Specifically, the coloring and size of threats were considered insufficient.

6 Results

The analysis of the SAGAT data was performed using ANOVA in a within-subjects model for the 3 freezing points of each condition. Based on the results the use of both SAES designs in an augmented reality overhead display demonstrated a superior performance to no-design. Results from the SAGAT analysis using SPSS also revealed that design 1 (radar) was superior to design 2(arrows) and no design. This was identified as significant. Despite having different traffic conditions and SAES design phase 3 still shown significant results with other 2 phases. However, since SAES design were used interchangeably in phase 3 the effect of learning was equally distributed among the two designs.

Additionally, a within group chi square analysis for the road sections in which the freezing events occurred revealed a significant difference between the two SAES designs. Specifically, 100 data points were recorded by simulator for each participant in each road section. Therefore for the road section in which the events occurred, the data points used for the analysis were 100data points*3 road sections*3 designs*17 participants. Prior to the analysis the data were preprocessed based on assumptions regarding the driving behavior of low SA drivers. Specifically, high workload, speeding or high lateral deviations reduce SA and these 3 rules have been used to transform the data into a format that indicated the positive or negative effect on SA. A collated results of SA per design and road section was performed using a function that related the three factors. The chi square analysis of the collated results, showed that the Radar design was significantly better than the Arrows design which verifies the results of the SAGAT method.

7 Conclusions

The SA evaluation method presented herein provides a novel cost effective solution to evaluate prospective SAES. The method is based on design science and encourages the redesign of the artefact until it satisfies the designer needs. Results indicate that what the users experience during their interaction with the SAES and what they perceived of this experience as reported in the post-test questionnaire point to the same conclusion. Specifically, statistical analysis of the data collected during SAGAT and Log files indicated that the radar design is superior to arrows and no design. Similarly, subjective evaluation of the candidate designs also revealed the same results. Limitations of this work concentrate to fact that simulated settings do not currently offer the resolution of the real world, and so they may affect driving behavior.

References

1. Konstantopoulos, P., Chapman, P., Crundall, D.: Driver's visual attention as a function of driving experience and visibility, using a driving simulator to explore drivers' eye movements in day, night and rain driving. Accid. Anal. Prev. **42**, 827–834 (2010)
2. Eksler, V., Lassarre, S., Thomas, I.: Regional analysis of road mortality in Europe. Public Health **122**, 826–837 (2008)
3. Fuller, R., Santos, J.: Human Factors for Highway Engineers. Pergamon, New York (2002)
4. Gregoriades, A., Sutcliffe, A., Papageorgiou, G., Louvieris, P.: Human-Centred safety analysis of prospective road designs. IEEE Trans. Syst. Man Cybern. Part A **40**(2), 236–250 (2010)
5. Montella, A., Ariab, M., D'Ambrosiob, A., Galantea, F., Maurielloa, F., Pernettic, M.: Simulator evaluation of drivers' speed, deceleration and lateral position at rural intersections in relation to different perceptual cues. Accid. Anal. Prev. **43**, 2072–2084 (2011)
6. Endsley, M.R., Jones, D.G.: Designing for Situation Awareness: An Approach to Human-Centered Design, 2nd edn. Taylor and Francis, London (2012)
7. Jones, D.G., Endsley, M.R.: Use of real-time probes for measuring situation awareness. Int. J. Aviat. Psychol. **14**(4), 343–367 (2004)
8. Williams, A.F., O'Neill, B.: On-the-road driving records of licensed race drivers. Accid. Anal. Prev. **6**, 263–270 (1974)
9. Davenne, D., Lericollais, R., Sagaspe, P., Taillard, J., Gauthier, A., Espié, S., Philip, P.: Reliability of simulator driving tool for evaluation of sleepiness, fatigue and driving performance. Accid. Anal. Prev. **45**, 677–682 (2012)
10. Matthews, M.L., Bryant, D.J., Webb, R.D.G., Harbluk, J.L.: Model for situation awareness and driving. Transp. Res. Rec. **1779**, 26–32 (2001)
11. Walker, G.H., Stanton, N.A., Young, M.S.: The ironies of vehicle feedback in car design. Ergonomics **49**(2), 161–179 (2006)
12. 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)
13. Ma, R., Kaber, D.B.: Situation awareness and driving performance in a simulated navigation task. Ergonomics **50**(8), 1351–1364 (2007)
14. Taylor, R.M.: Situation awareness rating trechnique (SART): the development of a tool for aircrew systems design. AGARD-CP-478. In: Situation Awareness in Aerospace Operations. 3-1-3-17 Advisory Group for Aerospace Research & Development (1990)
15. Salmon, P.M., Stanton, N.A., Walker, G.H., Jenkins, D., Ladva, D., Rafferty, L., Young, M.: Measuring situation awareness in complex systems: comparison of measures study. Int. J. Ind. Ergon. **39**(3), 490–500 (2009)
16. Durso, F.T., Dattel, A.R.: SPAM: the real-time assessment of SA. In: Banbury, S., Tremblay, S. (eds.) A Cognitive Approach to Situation Awareness: Theory and Application, pp. 137–154. Ashgate Publishing, Furnham (2004)
17. Durso, F.T., Hackworth, C.A., Truitt, T.R., Crutchfield, J., Nikolic, D., Manning, C.A.: Situation Awareness as a Predictor of Performance in En Route Air Traffic Controllers. U.S Department of Transportation-Federal Aviation Administration, USA (1999)
18. Horswill, M.S., McKenna, F.P.: Drivers' hazard perception ability: situation awareness on the road. In: Banbury, S., Tremblay, S. (eds.) A Cognitive Approach to Situation Awareness: Theory and Application, pp. 155–175. Ashgate Publishing Limited, UK (2004)
19. McKenna, F.P., Crick, J.L.: Hazard perception in drivers: a methodology for testing and training (final report). Behavioural Studies Unit, Transport and Road Research Laboratory, Corwthorne, UK (cited in Horswill and McKenna 2004) (1991)

20. Hull, M., Christie, R.: Hazard perception test: the geelong trial and future development. In: Proceedings of the National Road Safety Seminar (1992)
21. Hevner, A., Chatterjee, S.: Design Research in Information Systems, Integrated Series in Information Systems, vol 22. Springer Berlin (2010)

Driving Behavior in Weaving Maneuver: A Driving Simulator Study

Maria Rosaria De Blasiis, Chiara Ferrante, Antonella Santilli and Valerio Veraldi

Abstract In terms of road safety, as well demonstrated in previous studies, the combination of concentrated flows in one specific area might induce more critical maneuvers. In this contest, one of the crucial points for road safety are weaving lanes where different flow can perform different maneuver merging and diverging, producing some conflict points. Bearing in mind this facet, this particular elements, have to be designed with a specific geometry, in order to ensure both maneuver, deceleration for exit and acceleration for entrance in the main flow. The main goal of this paper is the analyze the exchange maneuver of vehicles by means of a real time driving simulator, which allows to evaluate the performance of car users when approaching the exchange of lanes under the same boundary conditions. More specifically, the present paper deals with the extension of previously performed research on this type of element. Therefore, same indicators, same analysis technique and virtual reality scenarios were adopted, in order to establish a comparison between two different conditions of approach to the weaving lane: the first with same speed between the main and the secondary flow, the second one, with a significant speed difference between the flows. The analysis of the maneuvers has been addressed at first under a geometrical point of view and then estimating the related risk parameters in line with the same indicators presented in the previous research (deceleration and risk area evaluated as longitudinal ant transversal distance between vehicles). The main results highlight that, under a geometrical point of view, in different speed condition between the two flows, car users tend to reduce the length of the maneuver, performing the weaving earlier and then occupying a reduced length of the lane in comparison at same speed condition. Under the point of view of risk analysis, results are more significant in terms of traffic flow management rather than in terms of lengths of the lanes. In case of flows with different speed, car users tend to use brakes more often, probably because they are calibrating the speed of the vehicle approaching the weaving lane.

M.R. De Blasiis (✉) · C. Ferrante · A. Santilli · V. Veraldi
Department of Engineering, "Roma TRE" University, Via Vito Volterra 62, 00146 Rome, Italy
e-mail: mariarosaria.deblasiis@uniroma3.it

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_27

Keywords Drivers' behavior · Weaving lanes · Driving simulator · Perception of risk · Human factor

1 Introduction

1.1 Background

In term of road safety, it is well known that drivers' behavior and subjective perception of risk play a key role. As a matter of fact, the driver, adapts its behavior according to the stresses caused by road he is traveling on.

The parameter typically employed for verifying this direct correspondence is the speed of travel. Indeed, as suggested by [1, 2] the driver adjusts his speed in function of the subjective perception of risk accident [3], so that. The speed variation can represent a significant measure of risk perceived by the driver [4, 5]. In addition, experimental analysis [6] have shown that the assessment of the risk threshold by users, as well as depending on the discomfort induced by the road, is a function of more psycho-physiological variables, namely:

- the driving memory;
- the driver's behavior related to road conditions and driving expectations.

As demonstrated by [7], drivers' behavior is related to road safety and consequently to the changes in user's behavior result in a variation of the road safety conditions, and therefore on the incidental risk [8], [9]. In particular, under these points of view, the case of weaving lanes appears as worthwhile of interest, already since seventies [10, 11]. According to Highway Capacity Manual [12] weaving lanes are defined as the crossing of two or more traffic streams traveling in the same direction along a significant length of highway without the aid of traffic control devices.

In practice, weaving maneuvers occur when an acceleration lane is connected to a deceleration lane through an auxiliary lane that produces a roadway enlargement with respect to its current section. Therefore, this area, being crossed by different flows (namely weaving and straight traveling), and so holding a crucial importance as far as road safety is concerned [10], is taken as objective in this work.

The key aspect of this research is that the entry into the main flow, does not depend on the user's will, but is rather delayed by a time (gap-acceptance) which corresponds to an interval of space that must be large enough to allow the vehicle to join the main flow. The definition of this gap affects the geometrical design of the weaving lane, and broadly depends on three variables [13]:

- The traffic on the main lane;
- The insertion frequency of other vehicles into the main lane;
- The speed difference between the vehicle that is entered and the flow on the main lane.

First scientific contributes, testing several traffic and geometric conditions of the weaving lane [11, 14], suggested that the solution to such an issue may lay in the increase of the length of the weaving lane, in order to not affect the user's psychology in terms of quickness of the maneuver. More recent experimental studies conducted through a virtual reality driving simulator [15] demonstrated, on the contrary, that the length of the lane is not an influential parameter on the dynamics of the maneuver, which is rather conditioned by other parameters, some of which psycho-physiological. In fact, by testing a sample of users engaged in a weaving maneuver on a lane dimensioned with the maximum length defined by the Highway Capacity Manual [12] (600 m), it was found that the users tended to exploit about half of the available lane only. Therefore, it is evident that the lane design should not only refer to geometric variables, but should also refer to the parameters that influence the users' behavior.

Particularly, the difference in speed between the user and the vehicles on the main lane is known to be one of the most significant variable defining the gap-acceptance dimension [16]. In fact, this difference of speed affects the fluidity of the maneuver, which is quicker when the speed difference is low [17]. Conversely, when the speed difference between the two flows is significant, the entering vehicle will slow the main flow down, which involves a decay of the weaving lane functionality [18].

1.2 Objective

This research aims at analyzing the user's behavior, while performing the weaving maneuver, by means of a virtual reality driving simulator. Main goal is to determine the risk perceived by the users and the safety level of the maneuvers [19].

Authors focused their efforts on the study of the weaving maneuver conditions, in terms of speed difference between the flow on the main lane and the entering flow, by imposing a significant speed reduction. Lastly, results were compared with those deriving from of a previous experiment [15], in order to investigate different conditions of the approach maneuver.

2 Methodology

2.1 Equipment

In order to characterize drivers' behavior during driving STISIM Driving Simulator of the Virtual Reality Laboratory of University Research Centre for Road Safety (LASS3) has been used (Fig. 1).

Fig. 1 STISIM driving simulator at the virtual reality laboratory of university research centre for road safety (LASS3)

The reliability of full instrumentation has been fully validated [20, 21]. The simulator is a real vehicle, so that a better perception of the real conditions by the user during the tests is provided. The image are projected in front of the car and sideways in order to cover a visual angle of 135°, and sound speakers are located in the hood of the car, for emulating the acoustic environment at the best.

To assess the outcome of the simulations thirty-five parameters are recorded with a frequency equal to 0.25 s. Through a driving simulator study it is possible to make an objective comparison between the simulation output of different drivers, because this tool is able to repeat the same test conditions for all drivers.

2.2 *Sample of Drivers and Its Validation*

An homogeneous sample of subjects was selected and the same driving conditions were generated for each driver, in order to avoid biasing of results induced by driver attitude, experience in driving, age, stress phenomena, emotional state or neuro–cognitive status or by other factors. Twenty subjects (8 women and 12 men, 35 years old on average, with age ranging from 20 to 50) were recruited via direct contact as volunteers from the Department of Engineering at the University Roma Tre.

All participants had a valid Italian driving license and reported to have driven, on average, 13050 km in the previous year (range: 2000–25,000 km). None of the subjects had previous experience with driving simulator.

Through Chauvenet criterion [22] we assessed that the number of participants is significant from a statistical point of view, so assuring a correct statistical data interpretation. For each generated scenario, some drivers were excluded because they showed anomalous behavior during driving, typically in terms of average speed registered on the weaving lane. Outliers were considered as the speed values strongly higher/lower than mean value. According to the Chauvenet criterion, no data were rejected.

2.3 Scenario and Traffic Flow Conditions

Starting from the assumption that several findings in literature indicate that the length of the weaving lanes combined with traffic volumes is a crucial factor for accident rate along interchange zones [11, 23] the authors wanted to analyze drivers behavior in these different conditions.

Four simulation scenarios have been carried out, the cross-section has two lanes (3.50 m wide), designed according with the Italian regulations. The only difference between scenarios was the length of the weaving lanes. Starting from the maximum value of length provided by Highway Capacity Manual [12] which is 600 m called, three more lengths of the weaving lane were calculated by employing two different probabilistic methods, according to the Italian regulations [24]. Therefore, each simulation scenarios were characterized by a different length of the weaving lane, namely: Very Long 600 m (VL), Long 220 m (L), 185 m Medium (M), Short 140 m (S).

Four weaving flows representative of the wider range of road traffic conditions have been chosen. In particular four different classes have been defined: Very High Flow (VHF) 2200 veh/h, High Flow (HF) 1600 veh/h, Average Flow (AF) 800 veh/h, Low Flow (LF) 400 veh/h. Some vehicles, out of the flow, are expected to change lanes and some others to continue along their lane. Which kind of vehicle the driver meets on the weaving section depends on the velocity that he adopts when he faces the lane change. Each traffic flow condition has been combined with each simulation scenarios. In this way 16 scenarios with different driving situations have been carried out. Each driver drove all scenarios, so 640 weaving maneuvers have been analyzed.

2.4 Procedure

At the beginning participants were informed about the test procedures, in terms of duration, use of the steering wheel, pedals and gear. According to the procedure of simulation experiments, participants were required to complete a training simulation scenario for at least 10 min driving. The drivers were requested to drive along four simulated road stretches, in four different runs, without speed limitations. Subjects could see their speed on the speedometer visualized on the screen and they were free to choose the velocity, according to what the road scenario suggested them. Between each scenario participants were allowed a short break. This break was intended to avoid as much as possible the fatigue effect of each driving period.

2.5 Road Safety Indicators

The outcomes of simulation will be analyzed, in order to investigate drivers' behavior along the weaving lanes characterized by different lengths and in different traffic flow conditions, through two indicators:

- Deceleration;
- Risk area.

The first one represents how drivers change their speed when they start the weaving maneuver. The second one is based on the concept of safety distance.

Deceleration. Assumed that the speed variations is a reliable indicator of drivers risk perception [25], three different classes of deceleration, representative of drivers behavior, have been defined. The first class includes drivers that did not decelerate; the second one takes into account drivers that decelerate just by lifting. This allow the drivers to adjust their speed to coordinate the flow into which they have to get; the third class represents situation where the use of the brake is representative of a condition of risk perception that induced drivers to brake.

The three classes have been defined as following:

- Class 0—drivers do not decelerate approaching weaving lane;
- Class 1—maneuvers carried out only by lifting throttle;
- Class 2—maneuvers carried out braking;

In order to confirm these results has been analyzed also how each driver reduced his speed. With this aim the deceleration curves, performed in the weaving areas, have been studied (Fig. 2). Through the analysis of this curves it was possible to obtain a value which takes into account the deceleration value and how long drivers brake.

This curve is defined as:

$$\int_{Pi}^{Pf} a\,dt \tag{1}$$

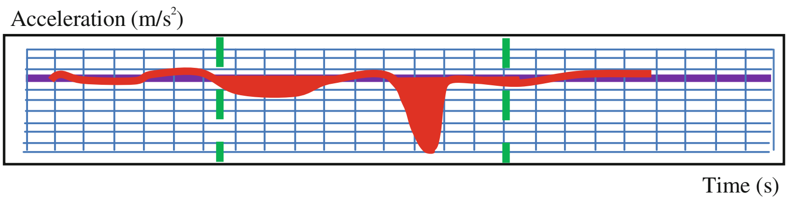


Fig. 2 The figure shows the curve of deceleration area

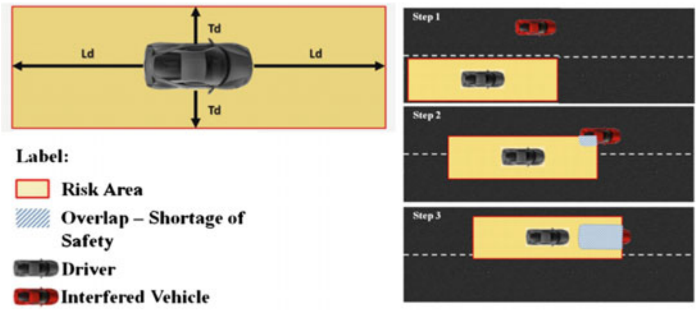


Fig. 3 The figure shows risk area and an example of shortage of safety during weaving maneuver

where:
P_f Final point of maneuver
P_i Initial point of maneuver
a Acceleration
t Time

Risk Area. Risk area is an indicator based on the concept of safety distance and it takes into account both the longitudinal distance (Ld) and the transversal distance (Td) between two vehicles that interfere on the weaving area (Fig. 3).

Usually to investigate safety audit during stopping maneuver, only longitudinal distance has been considered. Regarding weaving lanes, it is necessary to analyze not only longitudinal distance but also transversal distance because two vehicles can perform weaving maneuver at the same time.

The traditional longitudinal safe distance between vehicles D_s is defined as shown [26]:

$$Ds = \tau \times v \tag{2}$$

where:
t Reaction time
v Flow speed

Regarding transversal distance there are no findings in literature and because of this the authors decided to consider the dimension of the vehicle limit size (2.5 m). Whenever during the weaving maneuver an overlap occurs between the drivers' risk area and an interfered vehicle, a shortage of safety has been determined.

Moreover relationship between the overlap and the time in which the phenomenon occurs has been analyzed. This result allows evaluating not only the quantitative assessment of the shortage of safety but also for as long as this condition persists.

3 Results and Discussions

According to findings in literature [27], the Analysis of Variance test (ANOVA) to research the statistically significant differences on indicators values among analyzed configurations has been used. For each parameter two analysis were performed. The first one to evaluate the effects due to the length of the weaving lanes; the second one was performed to investigate the effects of different traffic flow conditions. The null hypothesis was that the average of the dependent variable was the same for the configuration investigated. Rejecting the null hypothesis would mean that independent variable, the traffic flow of the weaving lane, influences the dependent variable. Table 1 shows the results of the ANOVA test for each indicator performed on the four different lengths of the weaving lane combined with the four different traffic flow conditions. The null hypothesis was rejected almost for the whole set of data, in fact more than 80 % of cases have a statistical significance higher than 90 %.

3.1 Deceleration

The results about the deceleration indicator show an increasing amount of users who reduce their speed when interfering flows increase. Indeed, in each flow conditions, more than half of drivers slow down. In detail, the amount of drivers who reduce their speed ranges from 60 % in low flow conditions, to 90 % in very high flow conditions.

The decelerations using brake, which are the most significant in terms of drivers' behavior and therefore allow studying the real risk conditions, follow this trend. In fact, they show greater values for very high flow conditions. In particular (Fig. 4), the maximum value is in short lane in very high flow, equal to 22 %. This fact demonstrate that a too short lane and several vehicle interferences lead to a growing of risky maneuvers, caused by an high level of interferences that increase the risk perceived by the driver during the weaving maneuver.

In order to confirm deceleration results Deceleration Area has been analyzed. This area is the area under the curve which represent acceleration. Regarding this area the results confirm the trend of deceleration and show that the area is larger as the flow increases (Fig. 5).

3.2 Risk Area

Regarding risk area is it necessary a preliminary consideration: in accordance to previous study, the risk area indicator show a bad perception of safety in the very long lane, by the sample of driver. This phenomena is highlighted by the high value of Risk Area show in Table 2.

Table 1 Results of the analysis of variance

Indicators	On traffic volume factor				On lane length factor			
	Short lane	Medium lane	Long lane	Very long lane	Low flow	Average flow	High flow	Very high flow
Deceleration area $[(m/s^2)/s]$	F(3,156)	F(3,156)	F(3,156)	F(3,156)	F(3,156)	F(3,156)	F(3,156)	F(3,156)
	5.6 $p < 0.01$	2.0 $p = 0.12$	5.1 $p < 0.01$	0.8 $p = 0.48$	12.7 $p < 0.01$	17.9 $p < 0.01$	13.1 $p < 0.01$	3.4 $p = 0.02$
Risk area $[m^2]$	F(3,156)	F(3,156)	F(3,156)	F(3,156)	F(3,156)	F(3,156)	F(3,156)	F(3,156)
	1.4 $p = 0.25$	3.2 $p = 0.025$	5.7 $p < 0.01$	4.5 $p < 0.01$	14.6 $p < 0.01$	23.0 $p < 0.01$	31.4 $p < 0.01$	8.4 $p < 0.01$

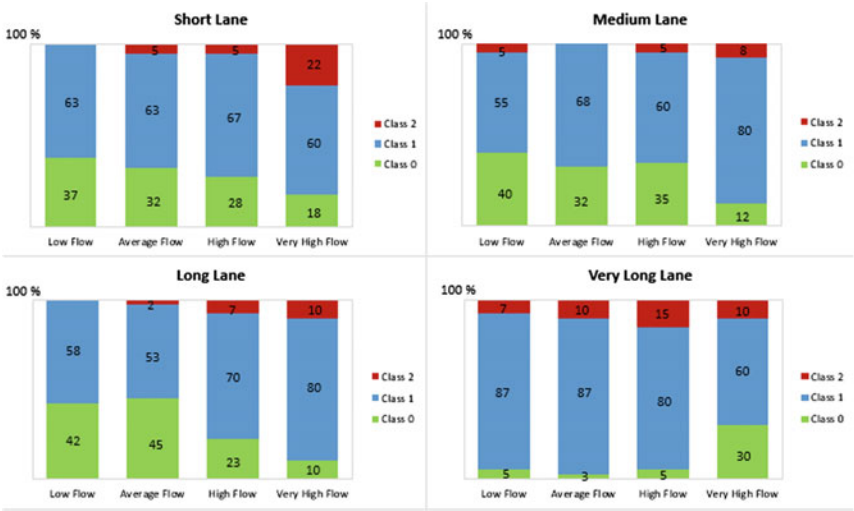


Fig. 4 Deceleration classes

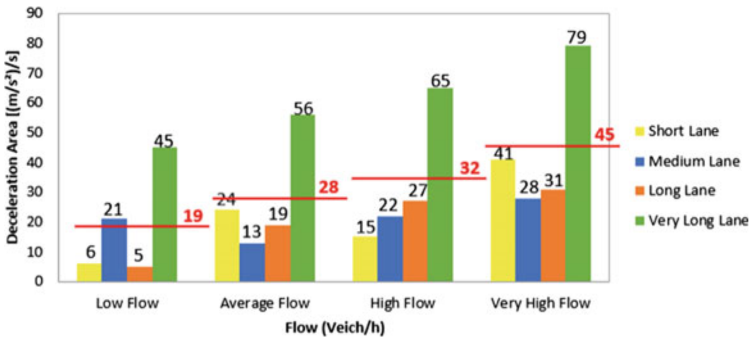


Fig. 5 Deceleration area

Excluding the anomaly of the very long lane the results show a significant variation in function of the flow, but not in function of the length of the lane. In particular, the value of risk area increase when interfering flow increase. On the contrary, varying the length of the lane, the values of risk does not show a significant variation. The maximum value is in the low flow between the short and the long lane, and is 28 (Fig. 6). This result show that the increased interference increase the levels of risk perceived by the driver, who pays more attention to the flow rather than to the length of the weaving lane. With a very high flow, the maneuver is less regular, increase the decelerations and reduces the safety distance between vehicles. This ever results in an increase in the drivers' discomfort and to a lowering of the risk acceptance threshold.

Table 2 Average values of risk area

Lane	Flow			
	Low	Average	High	Very high
Short	59	104	171	199
Medium	46	91	167	202
Long	31	92	159	192
Very long	93	344	401	311

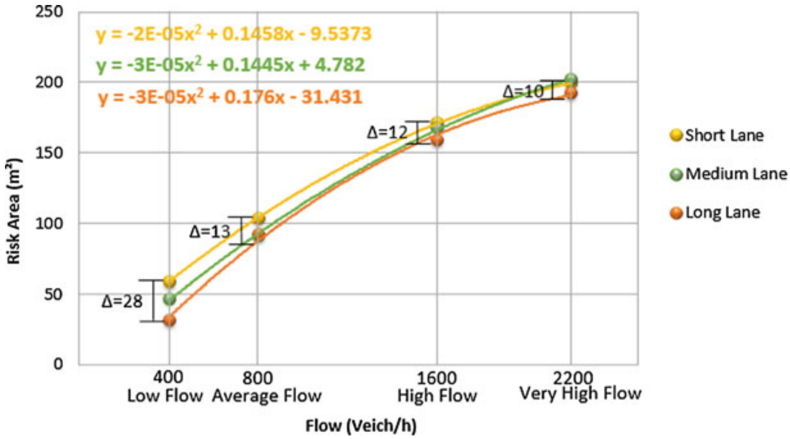


Fig. 6 Risk area

4 Conclusions

The paper deals with the analysis of drivers' behavior, through the use of a driving simulator, in approaching the weaving lane with a difference in speed respect to the main lane. The contribution of this research is the analysis of drivers' behavior in these situations in order to suggest a method for weaving lane design. For this purpose the effects of different length of weaving lanes and different traffic volumes have been studied and statistically validated. Results showed that the drivers' behavior is influenced significantly more by the flow on the main lane rather than by the length of the weaving lane. Therefore, the driver perceives more the risk of not finding the gap acceptance to take the weaving lane, because of the high interference, than the risk of not having enough time to finish the maneuver.

This topic is faced by means of two indicators: deceleration and risk area. Regarding deceleration, a high percentage of braking is representative of a high risk perception by drivers. In order to validate these results, the authors analyzed another indicator defined as risk area and based on the concept of safety distance.

The results obtained by the analysis of this indicator showed that driver's behavior changes significantly under high and very high flow conditions, accordingly to previous results.

About length of lane, which is the key element of the weaving lane design, this research reaches interesting results about behavior of drivers traveling along short weaving lanes: tests pointed out that a too short lane and significant vehicle interferences lead to an increase of risky maneuvers. Conversely, it has been shown that a very long weaving lane does not increase the safety level, but generates an uncertainty of behavior of the driver in the approach maneuver. All results, in fact, showed anomalous values related to the deceleration and to the risk area for very long lanes, so confirming that a too long weaving lane does not lead to higher safety conditions. Conversely, this study unlighted that a too high increase of weaving lanes length leads to an uncertainty in the behavior of drivers, who do not correctly interpret road geometry, and are so forced to adjust their trajectory by braking, with high adhesion values.

Despite the promising results, the present research can be expanded by testing a larger sample of drivers and analyzing different flow conditions, with the aim of developing a comprehensive analysis of the drivers' behavior.

References

1. Nobuhiro, U., Yasunori, I., Itsubo, S., Yasuhara, S.: A microscopic analysis of traffic conflict caused by lane-changing vehicle at weaving section. In: Proceedings of the 13th Mini-Euro Conference Handling Uncertainty in Transportation Analysis of Traffic and Transportation Systems, pp. 143–148 (2002)
2. Guppy, A.: Speeding in relation to perceptions of risk, utility and driving style by British company car drivers. Ergonomics (1995)
3. Horswill, M., McKenna, F.: The development validation and a application of a video-based technique for measuring an every day risk-taking behavior: drivers speed choice. J. Appl. Psychol. **84**(6), 977 (1999)
4. Winsum, W., Heino, A.: Choice of time-headway in car-following and the role of time to collision information in braking. Ergonomics. **39**(4), 579–592 (1996)
5. Farah, H., Polus, A., Bekhor, S.: Drivers' irrationality in evaluating risks on two-lane highways. J. Transp. Saf. Secur. **4**(1), 67–82 (2011)
6. Benedetto, C., Calvi, A.: Driver's perceived discomfort and road safety. Adv. Transp. Stud. **14**, 37–48 (2008)
7. Benedetto, C., De Blasiis, M.R.: La considerazione della "memoria breve" per le verifiche di qualità del progetto stradale: Gli effetti della geometria d'asse sulla dispersione delle traiettorie, vol. 1, pp. 340–350, Parma (2002)
8. Wang, B., Hensher, D.A., Ton, T.: Safety in the Road Environment: A Driver Behavioural Response Perspective. Springer, Berlin (2002)
9. Iversen, H., Rundmo, T.: Personality risk driving and accident involvement among Norwegian drivers. Pers. Individ. Differ. **33**(8), 1251–1263 (2002)
10. Calvi, A., De Blasiis, M.: Driver behavior on acceleration lanes: driving simulator study. Transp. Res. Board, 96–103 (2011)
11. Cirillo, J.A.: The relationship of accidents to length of speed-change lanes and weaving areas on interstate highways. Transp. Res. Board (1970)

12. Highway Capacity Manual. In: Transportation Research Board, National Research Council, Washington DC (2000)
13. Kesting, A., Treiber, M., Helbing, D.: General lane-changing model MOBIL for car following models. Transp. Res. Record (1999)
14. Golob, T.F., Recker, W.W., Alvarez, V.M.: Safety aspects of freeway weaving sections. Transp. Res. **38**(1), 35–51 (2003)
15. De Blasiis, M.R., Diana, S., Veraldi, V.: Safety audit for weaving maneuver: a driver simulation safety analysis. In: 91th Transport Research Board Annual Meeting, RSS (2015)
16. Marczak, F., Daamen, W., Buisson, C.: Key variables of merging behaviour empirical comparison between two sites and assessment of gap acceptance theory. Procedia Soc. Behav. Sci. **80**, 678–697 (2013)
17. Esposito, T., Mauro, R., Corradini, M.: Fondamenti di infrastrutture viarie, vol. I. La geometria stradale, Hevelius Edizioni (2003)
18. Roess, R.P., Prassas, E.S.: Overview of uninterrupted flow methodologies of the highway capacity manual. In: The Highway Capacity Manual: A Conceptual and Research History. Springer, Berlin (2014)
19. Benedetto, A., Benedetto, C., De Blasiis, M.R.: A new effective approach to accidents prediction to improve roads' design and rehabilitation. Transp. Res. Board (2004)
20. Bella, F.: Validation of a driving simulator for work zone design. Transp. Res Board (2005)
21. Benedetto, C., De Blasiis, M.R., Benedetto, A.: Driving simulation based approach for quality control of road projects. Transp. Res. Board (2003)
22. Bella, F.: Driving simulator for speed research on two-lane rural roads. Accid. Anal. Prev. **40** (3), 1078–1087 (2007)
23. Cirillo, J., Dietz, S., Beatty, R.: Analysis and modelling of relationship between accidents and the geometric and traffic characteristics of the interstate system. In: Federal Highway Administration, Washington DC (1969)
24. Da Rios, G.: Progetto di intersezioni stradali, UTET Edizioni (1985)
25. Benedetto, A., De Blasiis, M.R., Calvi, A.: Risk of vehicle rear collision in function of traffic flow (2007)
26. Ministry of Infrastructure and Transport, "Norme funzionali e geometriche per la costruzione delle intersezioni stradali," (2006)
27. De Blasiis, M.R., Calvi, A.: Drivers' behaviour on acceleration lanes: a driving simulator study. In: TRB Annual Meeting, Washington DC (2011)

Observations of Drivers' Behavior When Opening Car Door with Two Stages

Cheng-Yong Huang

Abstract Door crashes are common occurrences in motorcycle-car accidents in Taiwan. Since 2013, the two-stage door opening method has been an examination item when obtaining a driver's license in Taiwan to teach people to avoid these types of accidents. According to the investigative results of this study, half of the car drivers in Taiwan still do not know the two-stage door opening method. Furthermore, observations of driver's two-stage door opening behavior in this study showed that men habitually use their left hand and women habitually use their right hand to open car doors. The one-way analysis of variance (ANOVA) showed that the older the driver, the shorter their first stage car door opening distance is. This means that elderly drivers are more careful about opening their car doors. The correlation analysis showed that the body weight continuous variable has a significant and negative correlation with car door opening distance, and a significant and positive correlation exists between car door opening distance and driver's field of vision. The results of this study can serve as a reference for future car door opening human factor design.

Keywords Door crash · Motorcycle-car accident · Behavior observation

1 Introduction

Door crashes are a type of vehicle and motorcycle traffic accident that is unique to Taiwan. The cause for this is Taiwan's special human geography and traffic patterns. Taiwan is an island with an area of 36,000 km^2 and a population of approximately 23 million people. Taiwan ranks 10th in population density among countries in the world and has an average of 645 people per square kilometer. Overall, 70 % of land in Taiwan is mountains and hills, and plains only account for

C.-Y. Huang (✉)
Department of Arts and Design, Natinal Dong Hwa University,
No. 1, Sec. 2, Da Hsueh Rd., Shoufeng, Taiwan 97401, People's Republic of China
e-mail: yong@mail.ndhu.edu.tw

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_28

30 % of land. Thus, the population is mainly concentrated in the urban areas of the plains. Population density in the most populated urban areas can reach 10,000 people per square kilometer [1]. Taiwan also has the highest proportion of motorcycles as transportation tools among all nations [2]. According to Taiwan government statistics, there were approximately 15 million motorcycles and 7 million cars in Taiwan as of 2012 [3]. That means there were twice as many motorcycles as cars. This phenomenon is related to Taiwan's economic development. Motorcycles have better fuel economy and are very suitable for short distance travel. In addition, motorcycles are convenient to drive in narrow alleys and crowded areas. The combination of narrow crowded areas and excessive number of motorcycles and cars means that it is difficult to find parking in urban areas. Consequently, parking spots are often made on the side of the road while motorcycles drive by the side of these parking spots. If drivers parked on the side of a road do not pay attention to incoming vehicles and suddenly open their doors, they can easily cause a moving motorcycle to crash into the car door and result in traffic accidents.

Because door crash accidents are common in Taiwan, the Taiwan Vehicles Supervision Station added a two-stage door-opening method item to driver's tests starting in July 2013 [4]. That is, drivers must first look whether there are vehicles passing by when they open car doors to get in or get out of their cars. However, the so-called two-stage door opening method does not actually have any clear specifications. Thus, we attempted to use a human factor engineering perspective to explore two-stage door opening behavior.

2 Literature Review

Door crash accidents are a type of automobile-motorcycle right-of-way (ROW) accident [5]. Because Taiwan cities are crowded and narrow, and there are a large number of cars and motorcycles, many parking spaces are located on the side of streets for temporary parking. However, motorcycles often drive on the outer lanes, and drivers of parked cars can interfere with a motorcycle's right of way when they open the car door. When car drivers do not pay attention while opening their car door, they can surprise passing motorcyclists and cause a door crash accident. The fallen motorcycles rider can also then be easily run over by cars behind them and suffer severe injuries. Illustrations of potential door crash accidents are shown in Fig. 1.

The term door crash was seen in Pei's [6] article in the Accident Analysis and Prevention journal. However, Dennerlein and Meeker [7] have also mentioned this type of traffic accident, which they refer to as bicycle riders colliding with the car door of cars parked on the side of the road. Severe injuries can be caused by collisions with the door (bicycle rider flipping) and by impact with the ground, especially for bicycle riders who are not wearing helmets. This type of research on bicycle riders colliding with car doors opened by drivers is rare. Although some

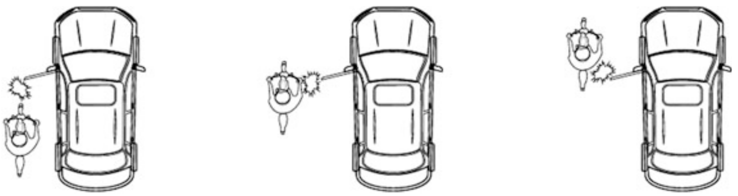


Fig. 1 Door crash illustration

unpublished articles discuss the inappropriateness of bicycle lanes being in close proximity to roadside parking spots and the associated risk of door crash, most typical bicycle lanes are still in the "door zone" of parked cars because car doors extend out about 90–105 cm when opened. Most bicycle lanes are often only slightly wider than 90–105 cm, which often increases the risk of bicycles colliding with open doors of parked cars even if the bicyclist is riding in the middle of the bicycle lane. Regardless of whether the bicyclist is riding on a bicycle lane, bicyclists are often instructed to ride one door distance away from roadside cars. This can improve the visual field of the car driver at the side of the road and make the bicycle rider more visible [8].

Another paper related to door crashes is by Johnson et al. (2013), which pro-posed that bicycle riders involved in collisions with open car doors often end up with severe injuries. However, few studies have been conducted on the nature of these accidents and the resulting injuries [9]. A "door crash" refers to the potential collision that can result when open doors of roadside cars are in the path of bicycle riders [7]. Fatal results often occur when bicycle riders collide unexpectedly with open car doors. From 1989 to 2012, four bicyclists in Victoria, Australia, died from door collision accidents [9]. Similar to other types of bicycle accidents, door collision accidents are uncommon, thus it is difficult to accurately determine a total number. Generally speaking, no records or data are kept, especially when the bicycle is not damaged [10].

The two previously described papers are related to bicycle and car door crash accidents. Huang [1] used a human factor engineering perspective to observe drivers' car door opening behavior. The test subjects wore an EEG and a camera on their heads while driving and parking at a designated parking spot. The results of the study showed the door opening behavior for four different types of car drivers. The types are as follows: (1) drivers that did not look before opening the door; (2) drivers that looked at the rear view mirror before opening the door; (3) drivers that turn their head and look out the rear left side window before opening the door; and (4) drivers that partially open the door and look first before fully opening the door. Brainwave focused data of drivers while opening the door also shows that drivers will directly open doors without looking when roads are wider and there is less traffic, and that their focus level while opening car doors are also lower.

The above-described study also showed that there is a 40 % probability that drivers in Taiwan use rear view mirrors to look for incoming vehicles. There is only a 20 % probability that drivers will use a two-stage door opening. This indicates that most drivers are not in the habit of using two-stage door opening. Thus, further pro-motions and studies are required to improve two-stage door opening awareness in drivers.

3 Experiment Design

According to Human Factor by Hsu et al. [11], human factor engineering studies can be conducted in the laboratory or in the field. Variables are easier to control in a laboratory scenario and the study will be less likely to be affected by non-topic related interference, which means that the collected data is also more accurate. The advantage of a field environment is that study related variables, environmental conditions, and test subject characteristics all conform to actual situations. Thus, the obtained results are more likely to be inferred on actual operations. The disadvantage is high cost, safety risks to test subjects, and experiment control is harder to achieve [11]. Behavioral observation is primarily used in this study to understand driver' two-stage door opening behaviors. Observing in a laboratory environment cannot realistically reveal the situation when drivers open the car door. Therefore, we chose to use the field study method because it is more appropriate.

3.1 Experiment Environment Setup

To understand drivers' two-stage door opening status, a Toyota Altis 1800 cc experiment car was used for the experiment. This is the number one selling car in Tai-wan. The experiment was divided into two parts. The first part was the car door opening distance and the second part was the driver's field of vision. The car door opening distance mainly measures the distance in the first stage of drivers' car door opening. That is, when the driver partially opened the door to look for incoming vehicles. A tape measure was fixed on the car with a 3D printed fixing device. When the car door opens, we could determine the opening distance. A camera was set up in the back to record the distance in the first stage car door opening. To test the driver's field of vision, a two-meter wide telescoping rail was set up in the rear of the car. The cloth was marked with scales in five centimeter increments. Test subjects were also asked to wear a camera to record the width that can be seen by the subjects when they open the car door to look backwards.

3.2 Experiment Procedure

First, we set up the experiment vehicle and the recording equipment rack. Then the subjects were asked to wear a Gopro head camera. We then asked the test subject whether they know what a two-stage door opening is. This can determine the general public's awareness level towards two-stage door openings. If the test subject does not know, then the experiment assistant will notify the subject of how to conduct a two-stage door opening. Next, the test subject was asked to get in the car, simulate shutting off the car engine, and use a two-stage door opening to get out of the car. At that time the experiment assistant will take note of which hand the test subject uses to pull the door handle. Subsequently, the test subject was asked to get in the car and repeat the same steps, with the exception of using a different hand to conduct the two-stage door opening. This can determine which hand the general public uses to open car doors, and the card door distance and visual distance when using their left/right hand to open car doors.

4 Experiment Result

In total, 22 test subjects participated in this experiment. The test subjects included 12 men and 10 women from 22 to 58 years of age. All are Republic of China citizens and have a valid Taiwan driver's license. Test subjects' car driving experience range from 1 to 30 years. Two of the test subjects were eliminated because their head cam angles were too high and did not record their visual distance when opening car doors. Thus, 20 subjects (10 men and 10 women) completed valid experiments. The following is a description of their two-stage door opening observation results.

4.1 Two-Stage Door Opening Awareness

Of the 20 test subjects, 10 did not know about the two-stage door opening. Of the 10 women, six did not know about the two-stage door opening. Of the 10 men, four did not know about the two-stage door opening. Therefore, 60 % of women and 40 % of men did not know about the two-stage door opening. Awareness results are shown in Fig. 2.

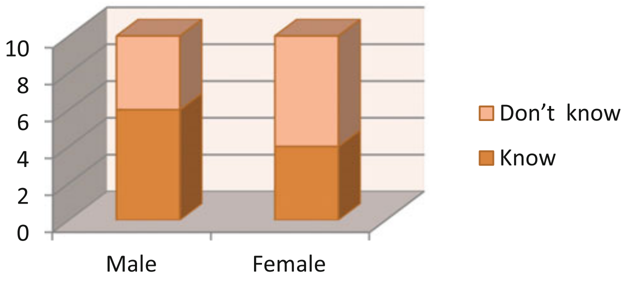


Fig. 2 Men and women's awareness of two-stage door opening

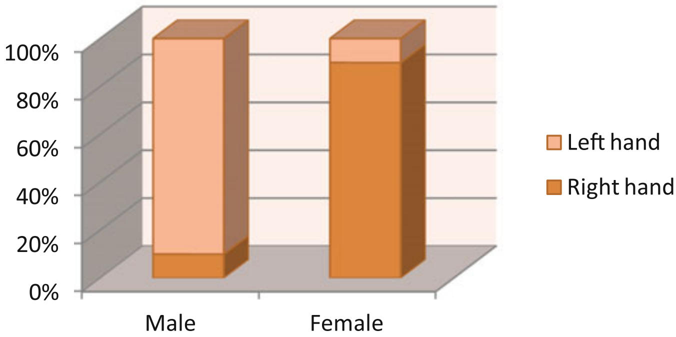


Fig. 3 Men and women's hand habit when opening car doors

4.2 Hand Habitually Used to Open Car Doors

The experiment assistant did not notify the test subjects of which hand to use in the two-stage door opening and the test subject followed their own habits in opening the car door. The results showed that out of 10 men, 9 habitually used their left hand to open car doors. Of the 10 women, 9 habitually used their right hand to open doors. This hand use habit was one major difference between men and women when opening car doors. Habitual hand used by men and women to open car doors is shown in Fig. 3.

4.3 Car Door Distance One-Way ANOVA

One-way ANOVA analysis was conducted on the first stage door opening distance. That is, one-way ANOVA was conducted based on gender, door opening hand, and

Table 1 One-way ANOVA summary

Item		Number	Average	Standard deviation	Standard error	F	Significance
Gender	Men	20	22.10	7.21	1.61	0.032	0.860
	Women	20	22.55	8.73	1.95		
Door opening hand	Right hand	20	20.30	8.65	1.93	2.742	0.106
	Left hand	20	24.35	6.69	1.50		
Age group	20–29	6	29.50	8.87	3.62	3.504	0.025
	30–39	8	19.38	8.18	2.89		
	40–49	14	23.85	7.42	1.98		
	50–59	12	18.92	5.30	1.53		

age group. The summary of the analysis is shown in Table 1. The results showed that both men and women had an average door opening distance of 22–23 cm, and there was no significant difference. Statistic results showed that the first stage car door opening distance when using left hand to open the door was 24.5 cm. This is four centimeters more than when using the right hand to open the car door, but statistically, this is not a significant difference. Finally, subjects were tested based on age (car driving experience). Four groups were formed in 10-year increments. Results showed that the youngest group (20–29 years old) had the largest first stage car door opening distance (average of 29.5 cm). The oldest group (50–59 years old) had the shortest first stage car door opening distance (average of 19 cm). Statistically, this is a significant difference. This indicated that the older the driver the more cautious the driver was when opening car doors. Car door opening distance for different age groups is shown in Fig. 4.

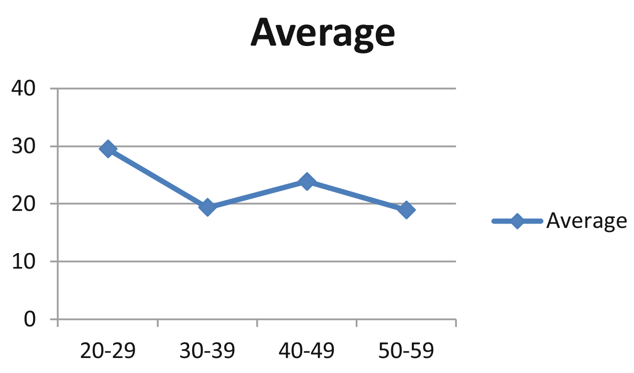


Fig. 4 Car door opening distance for different age groups

4.4 *Relationship Between Human Factor Size and Field of Vision and Car Door Opening Distance*

Before the test subjects proceeded with the two stage door opening experiment, they were measured for body height, body weight, and left/right arm length. Correlation analysis was conducted on these continuous variables in relation to door opening distance and field of vision during first stage door opening. The results were used to understand the linear relationship between continuous variables. Summary of the correlation analysis is shown in Table 2. The correlation analysis showed that body weight has a significant negative correlation (one-tailed test) with car door opening distance. That is, the heavier the body weight, the smaller the car door opening distance (r = −0.276*). However, car door opening distance in relation to the test subject's field of vision exhibited a positive correlation (r = 0.674*). This means that greater the car door opening distance, the broader the field of vision.

The above results show that body weight will affect car door opening distance. Next, we conducted further analysis of left/right hand use in correlation to car door opening distance. Results showed that in right hand car door opening distance, body weight did not have a significant correlation with car door opening distance. However, car door distance and field of vision showed an even higher positive correlation (r = 0.714**). This indicates that field of vision was wider when using the right hand to open doors. A summary of right hand correlation analysis is shown in Table 2. In left hand car door opening distance, body weight and left hand car door opening distance exhibited a significant negative correlation (r = −0.392*). This indicated that the heavier the body weight, he smaller the car door opening distance. A summary of left hand correlation analysis is shown in Table 3.

Table 2 Correlation analysis summary

	Body height	Body weight	Right arm	Left arm	Car door distance
Body height					
Body weight	0.463				
Right arm	0.881	0.322			
Left arm	0.889	0.356	0.922		
Car door distance	−0.073	−0.276	−0.047	−0.105	
Field of vision	−0.085	−0.177	0.057	−0.010	0.674

Table 3 Summary of right hand correlation analysis

	Body weight	Right arm	Car door distance
Body height			
Right arm	0.322		
Car door distance in right hand opening	−0.207	0.045	
Field of vision	−0.219	−0.040	0.714

Table 4 Summary of left hand correlation analysis

	Body weight	Left arm	Car door distance
Body height			
Left arm	0.356		
Car door distance in left hand opening	−0.392	−0.284	
Field of vision	−0.125	0.037	0.654

5 Discussion

The human factor engineering experiment in driver's two-stage door opening was conducted in this study. The following are the results of this study: (1) although the two-stage door opening has been required by the Taiwan MVDIS since 2013 when testing for a license, half of the car drivers in Taiwan do not know they need to use two-stage door opening. Overall, 60 % of women and 40 % of men drivers do not know about this regulation; (2) when testing subjects' habitual hand use in opening car doors, 9 out of 10 men used their left hand to open the car door. Out of the 10 women, 9 habitually used their right hand to open the car door; (3) results show that the older the driver, the more cautious they are about opening the door, and the shorter the car door opening distance; and (4) the correlation analysis results showed that body weight affected the left hand car door opening distance. The heavier the body weight of the driver, the shorter the car door opening distance. Car door opening distance also exhibited a highly positive correlation with field of vision (Table 4).

6 Conclusion

Taiwan has few plains and many mountains. Taiwan also has a high population density concentrated in urban areas, which creates unique traffic patterns. The motorcycle ratio among transportation tools is the highest in the world, and there are twice as many motorcycles as cars. Because not enough space is available for parking, a large number of parking spots have been established on the side of roads. Motorcycles driving past these parked cars can easily cause door crash accidents.

Although two-stage door opening has been added to the driver's test since 2013, half of the people with driver's licenses do not know about the two-stage door opening. More women do not know about this rule than men. The majority of men use their left hand to open car doors while women generally use their right hand to open doors. The older the driver, the more cautious they are when opening car doors, and have smaller door opening distances. Body weight and left-handed car door opening distance exhibits a significant negative correlation while car door opening distance exhibits a significant positive relationship with driver's field of vision. The above results can serve as a reference for future human factor design in car door opening.

Acknowledgments The author gratefully acknowledges the financial support from the Ministry of Science and Technology in Taiwan (Most 104-2410-H259-046).

References

1. Huang, C.Y.: Observations of drivers' behavior when opening car doors. In: 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015, pp. 399–406. Las Vegas (2015)
2. National Statistics, R.O.C. http://ebas1.ebas.gov.tw/pxweb/Dialog/statfile9.asp#
3. Pai, C.H., Hsu, J.J., Chang, J.L., Kuo, M.S.: Motorcyclists violating hook-turn area at intersections in Taiwan: an observational study. Accid. Anal. Prev. **59**, 1–8 (2013)
4. Motor Vehicle Driver Information Service. https://www.mvdis.gov.tw/
5. Pai, C.W.: Motorcyclists right-of-way accidents-a literature review. Accid. Anal. Prev. **43**, 971–982 (2011)
6. Pai, C.H.: Overtaking, rear-end, and door crashes involving bicycles: an empirical investigation. Accid. Anal. Prev. **43**, 1228–1235 (2011)
7. Dennerlein, J., Meeker, J.: Occupational injuries among boston bicycle messengers. Am. J. Ind. Med. **42**(6), 519–525 (2002)
8. Hunter, W., Stewart, R.: An evaluation of bike lanes adjacent to motor vehicle parking. University of North Carolina Highway Safety Research Center (1999)
9. Johnson, M., Newstead, S., Oxley, J., Charlton, J.: Cyclists and open vehicle doors: crash characteristics and risk factors. Saf. Sci. **59**, 135–140 (2013)
10. Harman, B.: Bicycle safety and law enforcement. In: Australasian Road Safety Research Policing and Education Conference, Melbourne (2007)
11. Hsu, S.S., Peng, Y., Wu, S.P.: Human Factor (in Chinese), 2nd edn. Tsang Hai Publishing, Taipei (1990)

Resurrecting Driver Workload, Multivariate Analysis of Test Track Data

Jack L. Auflick

Abstract This paper presents multivariate analyses of data collected from divers at a test track during the Driver Workload Metrics (DWM) project. As noted in a prior publication, the DWM project was a cooperative effort with the National Highway Transportation Safety Administration (NHTSA) and four automotive manufacturers. The DWM project defined workload 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. It was hypothesized that, depending on the type of secondary task performed while driving, measured workload and the correlated quality of driving should either remain the same or decline, but would manifest in degraded measures of lane keeping, longitudinal control, or eye glance behavior. Data for this new analysis was collected from test subjects who drove an instrumented car on a test track while performing various on-board tasks. These data also contain additional responses from several new visual manual task that were originally deemed to be too hazardous for test subjects while driving on a major four lane highway. It was therefore further hypothesized that the new task would demonstrate higher levels of visual-manual workload when compared to less demanding tasks. As in the prior DWM multivariate paper, test subject responses from the kinematic and eye glance behavior from the test track data were first analyzed using Maximum Likelihood Factor Analysis. This well-known statistical method attempts 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 comprehend the concept of driver workload. As in the DWM on-road analyses, these new 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. Furthermore, when secondary statistical analyses of the normalized factor scores were done using Multivariate Analysis of

J.L. Auflick (✉)
Engineering Systems Incorporated, 1174 Oak Valley Drive,
Ann Arbor, MI 48108, USA
e-mail: jalinflick@outlook.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_29

Variance (MANOVA) the results found highly statistically significant workload differences in age groups and task type.

Keywords Human factors · Driver workload · Driver distraction · Exploratory factor analysis

1 Introduction

In recent years, automotive manufacturers have increasingly been introducing new technologies into their cars and trucks often suggesting that the new technologies will provide enhanced safety for drivers. For example, these technologies range from advanced navigation systems, lane departure warnings, blind spot warnings and active braking systems designed to prevent rear-end collision. While ostensibly improving safety, these technologies, as well as handheld devices like cell phones or tablets, generate fears from researchers and regulatory agencies alike that use of these new technologies and devices within the automotive cockpit will at times overload and distract the driver. As a result, various regulatory agencies within the US and the European community have funded extensive research projects in attempts to understand the dimensions of driver distraction. These noted projects include the Human Machine Interface and the Safety of Traffic in Europe (HASTE) [1], the Naturalistic Driving Program at Virginia Tech [2], the European Advanced Driver Attention Metrics Program (ADAM) [3], the Crash Avoidance Driver Metrics, and the Driver Workload Metrics[1] (DWM) [4].

Of particular interest for this paper, the DWM project established correlates between the driving metrics, eye glance behavior, and demands placed on the driver from secondary discretionary tasks that had the potential to interfere with the primary driving task. Workload was 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. It was hypothesized that, depending on the type of secondary task performed while driving, 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. Within the DWM project, there were three separate data collection efforts: (1) Analysis of surrogate laboratory test measures, (2) on-road data collection, and (3) data collected from subjects while driving on a test track.

In its final report, the DWM project extensively reviewed driver workload literature and Multiple Resource Theory (MRT) [5], creating a set of conventional experimental tasks commonly performed in vehicles today. These tasks were

[1]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), a partnership established by Ford and GM to undertake joint precompetitive work in advanced collision avoidance systems.

defined by the input and output modalities needed to perform the task: either visual input and manual output, or auditory input and vocal output [6]. Tasks used on the test track were comprised of eight auditory-vocal and 13 visual-manual tasks with a baseline just drive and a combination auditory-vocal and visual-manual task, where subjects picked up a cell phone, pressed a preset button, and then interacted with an automated voice recognition flight schedule service. An example of the auditory-vocal tasks was a biographic task where the experimenter asked the subjects a series of questions intended to elicit a verbal response. Questions such as, "Where do you live?" and "How many children do you have?" An example of the visual-manual tasks was the CD7 task that directed the subjects to select a CD from the visor, place it in the CD player, and tune to track seven. However, on the test track several additional visual-manual tasks were included that had originally been determined by an institutional review board as being possibly too hazardous for the on-road data collection. They included map reading (easy and hard), maze tracing (easy and hard), read text (easy and hard) and a destination entry navigation task.

1.1 Multi-Dimensional Interrelationships

This project accomplished several goals including development of 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 in-vehicle systems, and the creation of a metrics toolkit that could be used by automotive engineers during all stages of the design process to assess the implications of driver workload while using future systems. When the DWM project collected its data, workload, the construct of interest, had no direct empirical measure available at the time. As a result, the existence of driver workload was inferred from an extensive string of bivariate correlation and regression analyses, observed across multiple measures of performance. These individual analyses provided brief glimpses into the complex, hidden interrelationships within the collected data. DWM had one unrealized goal, i.e. to apply exploratory factor analysis (EFA) methods in an attempt to uncover the latent, unobserved structure of correlated measures of workload. It is this hidden structure that must be examined to empirically comprehend the multidimensional concept of driver workload.

As explained in the DWM final report, driver performance data on the test track was collected from 69 test subjects using a canonical repeated measures experimental design where each test subject repeatedly performed the secondary tasks during real driving on the track. From this pool of subjects, 42 were selected based on their having nearly complete eye glance behavior data. The sample of participants was approximately balanced by gender and age. Task sequences and presentations of the secondary driving tasks were randomized for each subject. The driving condition selected for testing was at highway speed, approximately 55 mph, using a car-following scenario on a multilane paved test track surface with sharp banked curves followed by long straight, level road, under clear, dry, daytime conditions.

During testing, subjects drove the same instrumented car that was the center car of a three-vehicle platoon. This platoon operated as a single testing unit and provided a realistic car-following driving experience. During each task, an extensive array of sensors, cameras, and on-board instrumentation recorded kinematic data for longitudinal and lateral vehicle control, including task duration, speed, range, range rate, time headway, and time to contact. Mean, median, standard deviation, minimum and maximum distances, and time durations (at the minimum or maximum) were captured for each of the kinematic variables. In addition, steering wheel behavior was measured for three variables, e.g., how much time during a task the steering wheel was held beyond a zero-degree location and how long it was held at either a 15 or 20° offset from zero degrees, including the number of lane exceedances and durations, plus the number of center stripe touches and their durations.

In conjunction with 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. Eye glance data was manually scored through review of on-board video taken during each task. As in the on-road testing, due to timing and funding constraints, only 42 of the 69 test subjects had complete eye data at the time the final DWM analyses were completed.

2 Factor Analysis

This present analysis followed the same procedures as those reported for the analysis of on-road data [7]. The same 42 vehicle kinematic variables (like speed, range, range rate, longitudinal and later 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 were the independent variables. The analysis also included the eye glance data from the 42 subjects who had complete visual data. There were approximately 14 subjects in each of three age related groups, i.e. older, middle, and younger. Their eye glance data included median, standard deviation, and percent task duration, etc. from test track driving related to mirror, situation awareness, task-related, and not-road or not-assigned locations. The purpose of this new analysis 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. Specifically, this analysis used the maximum likelihood factor analysis (MLFA) method [8, 9] one of several commonly used exploratory methods that in general, seeks to: (1) reduce the number of variables and (2) to detect latent unobserved structure within the relationships between variables. MLFA accomplishes this by first estimating the loadings and communalities in the data set and subsequently maximizing the probability of the observed correlation matrix occurring. In MLFA, maximizing the likelihood function determines the parameters that are most likely to produce the observed data.

2.1 Hidden Structure

This analysis started with all 42 independent variables. So there could have been up to 42 orthogonal or independent factors. MLFA first identified that there were only seven factors, i.e. latent (unmeasured) variables that defined the hidden relationships within the measured, observable variables. The seven factors all had Eigenvalues greater than 1.0, based on using both the Kaiser criterion [10] and Cattell's scree plots [11]. The Eigenvalues from these seven factors measured the variance from all the variables that were accounted for by a given factor. As seen in Table 1 below the seven factors explain approximately 63 % of the original variance in the data.

The initial MLFA solution left the factors and factor scores (i.e. estimates of actual subject scores for individual observations for each factor) in an un-rotated seven-dimensional space, making the given solution difficult to interpret. As an aid for interpretation, factor scores were rotated and normalized using the varimax orthogonal transformation [12, 13]. Varimax rotations were so named because the process maximizes the sum of the variances (e.g. Vari-Max) of the squared loadings (i.e. the squared correlations between variables and factors). 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.

Table 1 Eigenvalues, named factors, and explained variance

Factor	Eigenvalue	Total variance (%)	Cumulative eigenvalue	Cumulative (%)
Eyes up-on road	11.8	28.1	11.8	28.1
Lat. & long. lane variation	3.9	9.3	15.7	37.4
Situation awareness	2.8	6.7	18.5	44.1
No glances to unassigned	2.5	5.9	21.0	50.0
Cross centerline	2.0	4.8	23.0	54.7
Touch centerline	1.3	3.2	24.3	57.9
Range, range rate variation	2.0	4.7	26.3	62.6

Factor loadings are correlations of an observed variable with the underlying factor and are conceptually similar to Pearson's r, the common correlation coefficient. When the loadings are squared, the resulting values describe the percentage of variance within a specific independent variable explained by a given factor. Based on their individual directionality and magnitude, factor loadings must be interpreted through a highly subjective activity where "names" are applied to each factor based on the magnitude and direction of the rotated loadings. 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 by 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.

2.2 *Multivariate Analysis of Variance (MANOVA)*

During the varimax process explained above, individual scores for each test subject on each variable and task were normalized, and then rotated resulting in a matrix of factor scores representing numerical values defining a person's relative spacing or standing on each of the seven latent factors. Test subjects' factor scores on each factor were analyzed using Multivariate Analysis of Variance (MANOVA). The MANOVA tested hypotheses that there were statistically significant differences between gender, age groups (Younger 19–39, Middle 40–59, Older 60–79), and Task Type (Visual-Manual, Auditory-Vocal, Just Drive, Combination). Initial results found no significant main effect for gender. So Gender was excluded and the MANOVA was rerun. The seven rotated factors were the independent variables. As can be seen in Table 2, there were highly statistically significant main effects for Task Type ($p < 0.0000001$) and Age Group ($p < 0.000001$). There were no significant interactions.

Test subjects' factor scores on each factor were subsequently averaged by age group and task type. Means were then plotted on radar plots that present a multi-dimensional driving workload profile for different groupings of subjects based on how type of task or age group affects workload. Figure 1 below presents a

Table 2 Multivariate tests of significance

	Test	Value	F	Effect df	Error df	p
Intercept	Wilks	0.521	116.04	7	883	0.0000001
TaskType	Wilks	0.065	191.38	21	2536	0.0000001
AgeGroup	Wilks	0.941	3.8752	14	1766	0.000001
TaskType vs. Age	Wilks	0.939	1.3291	42	4145	0.076328

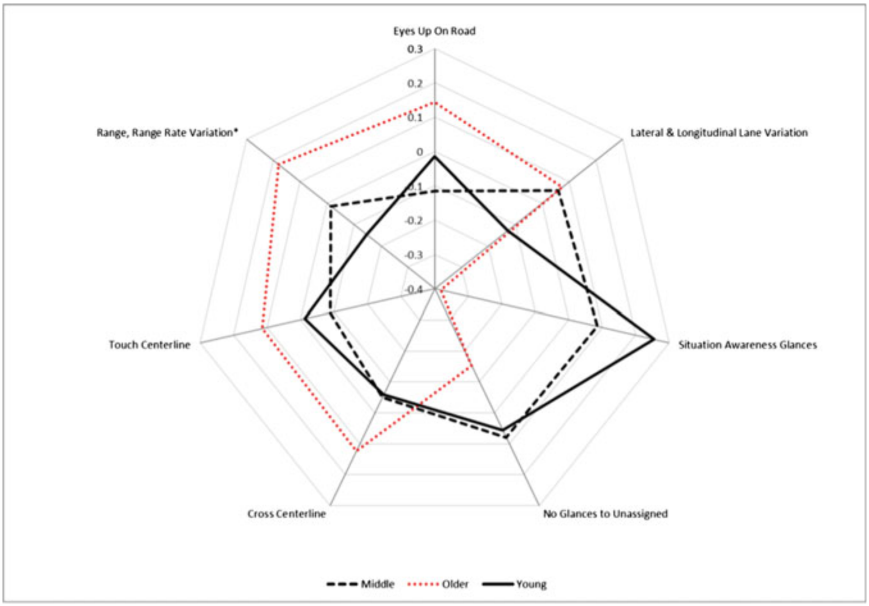


Fig. 1 Driver workload vs. task type

profile graphically comparing workload related to type of task. Note that Factor 1, the Eyes up-On Road factor, is in the 12 o'clock position while Factors 2 through 7 rotate in a clockwise direction.

2.3 ANOVA—New Visual-Manual Tasks

As mentioned above, the test track analyses included several additional visual-manual tasks that were felt to be too unsafe for drivers to perform while driving on a major highway. The original on-road tasks were also very short, lasting only 10 s or less. So the DWM team added six new tasks to test track analyses that were significantly longer in duration but required more eyes off the road time for test subjects. These new tasks included Route Tracing, Destination Entry, Read Easy, Read Hard, Map Easy, and Map Hard. An initial hypothesis was that these new tasks had significantly more work compared to the original set of visual manual tasks.

A second one-way Analysis of Variance (ANOVA) was used to analyze the 42 subjects' factor scores. However, this ANOVA only tested if the new visual-manual tasks were different, i.e. had more workload, compared to the original visual-manual tasks. Table 3 below presents the ANOVA results.

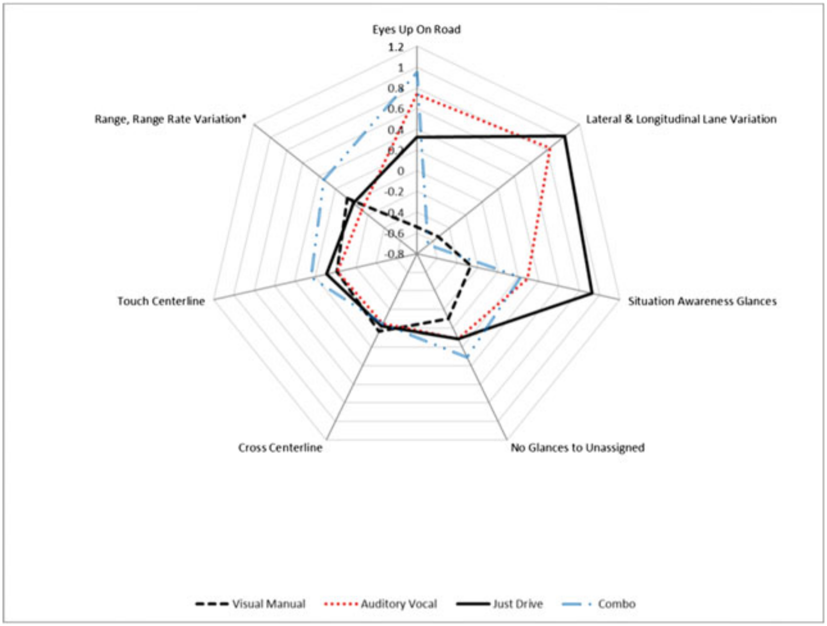


Fig. 2 Driver workload vs. age group

Table 3 ANOVA results comparing old vs. new visual-manual tasks

	Test	Value	F	Effect	Error	p
Intercept	Wilks	0.1133	562.9	7	504	<0.001
Old. vs. new	Wilks	0.6785	34.1	7	504	<0.001

As can be seen in the ANOVA, there were highly statistically significant differences between the two groups of visual-manual tasks. Figure 3 below shows the radar plot from this ANOVA. This is a relative cross-sectional comparison of the seven dimensions of work load differences between the two visual-manual task groups. When one examines these results and the associated radar plot, it becomes apparent that the older tasks did have less workload compared to the newer tasks, but the comparison is based on the interpreted factor names. For example, comparing the two groups on factor 1, Eyes Up-On Road, the older tasks had positive averaged factor scores, meaning that test subjects spent more time with eyes on the road when performing one of the original visual-manual tasks. Alternately, for factor 1, the new tasks have negative average scores, suggesting that test subjects were spending significantly less time with eyes on the road while performing one of the new visual-manual tasks. Similar comparisons can be made for the remaining six factors. However, note that factor 5, cross centerline, and factor 7, Range, Range Rate Variation, showed no significant differences between the two groups.

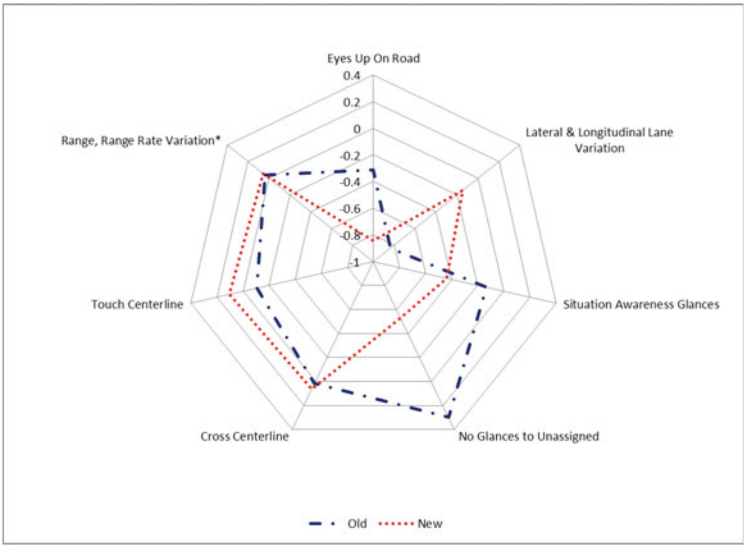


Fig. 3 Driver workload vs. age group

3 Discussion—Conclusions

These results from the MANOVA affirm the qualitative differences as shown in the proceeding figures. Given the original data, there are notable differences in driving workload as defined by the dependent grouping variables. The purpose of the MFLA using the DWM test track data was to examine the very large DWM data set and identify hidden structure that was indicative of driver workload. This MLFA, as well as the earlier reported analyses of on-road data, found that these DWM data contained extensive multicollinearity in the data set, hidden relationships, i.e. the latent structure that could be described by a minimum of seven factors while still being able to explain $\sim$63 % of the original variation in the data. The seven underlying factors began to reveal key effects of multitasking on driver workload caused by interference from secondary tasks. The actual driving task was shown to be multidimensional in nature, meaning that it was represented in the data by simultaneous effects on multiple factors. Driver workload, shown in the multivariate results, appears to be affected by different allocations of driver resources across input modalities as well as being affected by task type or age group. The MANOVA confirmed that within these data there are highly significant main effects due to age group, and task type.

It should be noted that the current results were based on the DWM data collected as test subjects performed secondary tasks while driving on an automotive test track. While ostensibly a driving task, drivers on the test track were able to drive in a tightly coupled three car platoon without having to devote attention to other vehicles that would have been driving near the instrumented test car driven on the

Table 4 On-road vs. test track factors and explained variance

	DWM on-road factors	Road explained variance	DWM test track factors	Track explained variance
1	Eyes up-aud. vocal work load	28.5	Eyes up-on road	28.1
2	Mirror & Sit. aware glances	11.9	Lateral & longitudinal lane variation	9.3
3	Glances to road	5.8	Situation awareness glances	6.7
4	Safety check glance variability	5.4	No glances to unassigned	5.9
5	% Task glance to road-not task related	5.9	Cross centerline	4.8
6	Crossing center stripe	4.7	Touch centerline	3.2
7	Variability-range and speed	3.8	Range, range rate variation	4.7

highway. Initially the author hypothesized that these differences in the driving task, e.g. road versus test track, would have some notable differences in workload. As in the first DWM on-road analyses, this analysis found that the MLFA could reduce the dimensionality of the original 42 independent variables by identifying seven new factors. Both analyses explained approximately 63 % of the original variance. However, when the rotated factor loadings were interpreted and named, qualitative differences in workload between the two data sets were observed.

Table 4 below compares the named seven factors from both analyses and includes the per cent explained variance. Figure 4 from the first DWM MLFA [7], also presents the radar plot from the on-road analysis. A visual, albeit qualitative, comparison of Fig. 4 compared to Fig. 1 will show the observed differences in workload.

Figure 5 below also presents the radar plot from the on-road analysis related to difference in age group. Again, one can visually compare Fig. 5 to Fig. 2 above to note that there are age related differences in driver workload but that there are also qualitative differences in workload based on the testing venue.

When qualitatively comparing test track versus on-road data, an additional difference in driver workload is that there are no gender differences and no two or three way interactions. The original on-road analyses did have these statistically significant differences. Based on these qualitative comparisons, it is apparent that both venues have demonstrable levels of workload. However, the individual factors are qualitatively different, and in addition, the test track results are statistically simpler, possibly showing that drivers were able to devote more attention to the driving task when there is no additional traffic on the road.

In applying this methodology on the test track data, once again there was an explicit recognition that it was exploratory in nature, and that the underlying dimensions it identified would need to be attributed with meaning and interpreted

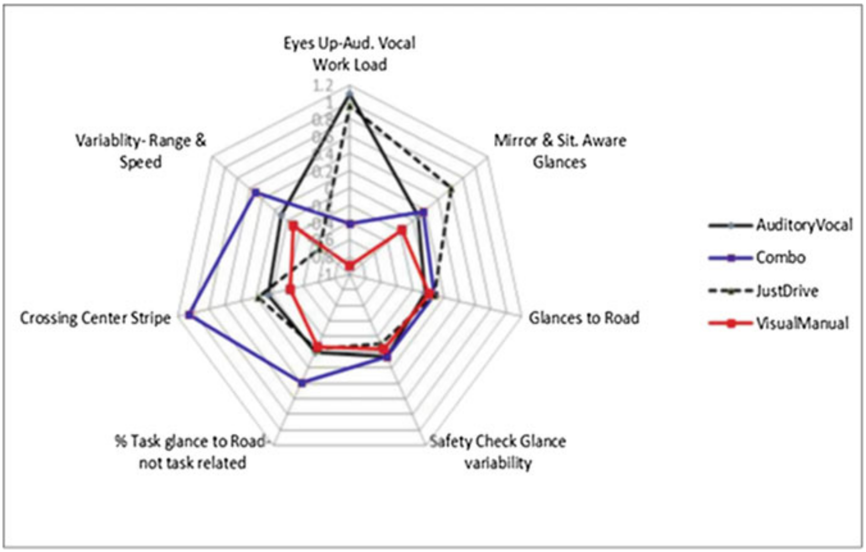


Fig. 4 On-road task type differences in workload

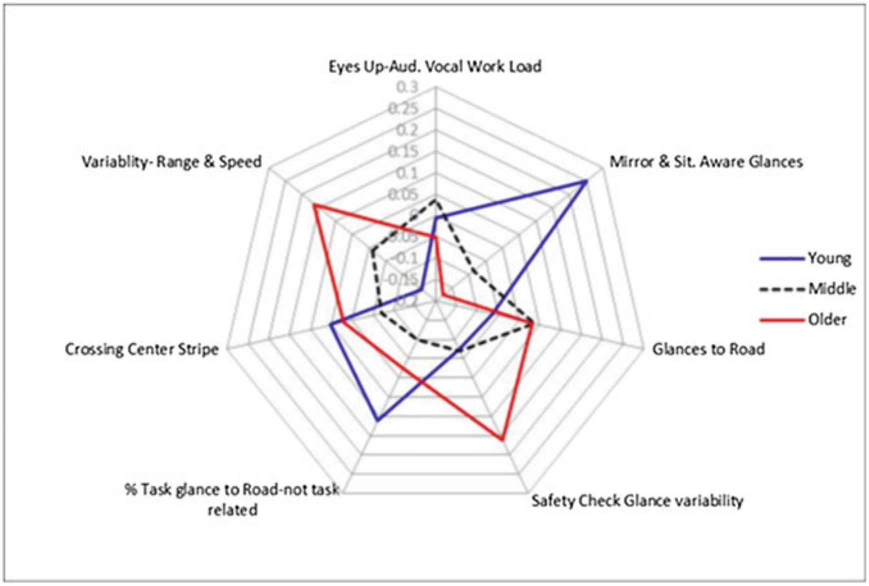


Fig. 5 On-road age group differences in workload

through subjective analysis. One must reiterate that this is exploratory work, and that the nature of the dimensions could change if the input to the analysis were different. Similarly, the interpretations of the underlying dimensions may be refined as a deeper understanding of the data set is acquired over time. The next step in this process will be to combine both test track and on-road data in one full MLFA of the DWM data using the same MFLA approach and then statistically compare the resulting factors scores using MANOVA. Assuming analysis provides similar results, one could then begin the laborious process of using Confirmatory Factor Analysis (CFA), to test whether measures of a construct are consistent with a researcher's understanding of the nature of that factor. CFA tests whether the data fit a hypothesized measurement model where the hypothesized model is based on theory and/or previous analytic research.

References

1. Carsten, O.M.J., Merat, N., Janssen, W.H., Johansson, E., Fowkes, M., Brookhuis, K.A., Merat, N., Carsten, O.M.J. (eds.).: HASTE Final Report, Version Number 3.1 (Nov 2005)
2. Neale, V.L., Dingus, T.A., Klauer, S.G., Sudweeks, J., Goodman, M.: An Overview of the 100-Car Naturalistic Study and Findings. National Highway Traffic Safety Administration, Paper 05-0400 (2005)
3. Breuer, J., Bengler, K., Heinrich, C., Reichelt, W.: Development of Advanced Driver Attention Metrics, (ADAM). Daimler Chrysler and BMW, Germany (2004)
4. Angell, L., Auflick, J.L., Hogsett, J., Austria, P.A., Kochhar, D., Biever, W., Diptiman, T., Kiger, S.: Driver Workload Metrics Task 2 Final Report, National Highway Traffic Safety Administration, Report Number DOT HS 810 635 (2006)
5. Wickens, C.D.: Processing resources in attention. In: Parasuraman, R., Davies, D.R. (eds.) Varieties of Attention, pp. 63–102. Academic Press Inc., New York (1984)
6. Auflick, J.L.: Resurrecting driver workload metrics: a multivariate approach. In: 6th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences, Procedia Manufacturing, vol. 3, pp. 3160–3167 (AHFE 2015)
7. Thurstone, L.L.: The vectors of the mind. Psychol. Rev. **41**, 1–32 (1934)
8. Lawley, D.N.: The estimation of factor loadings by the method of maximum likelihood. Proc. Roy. Soc. Edinb. (A) **60**, 64–82 (1940)
9. Kaiser, H.F.: The application of electronic computers to factor analysis. Educ. Psychol. Meas. **20**, 141–151 (1960)
10. Cattell, R.B.: The scree test for the number of factors. Multivar. Behav. Res. **1**, 245–276 (1966)
11. Kaiser, J.F.: The varimax criterion for analytical rotation in factor analysis. Psychometrika. **23** (3), 187–200 (September 1958)
12. Crawford, C.B., Ferguson, G.A.: A general rotation criterion and its use in orthogonal rotation. Psychometrika **35**, 321–332 (1970)
13. Spjotvoll, E., Stoline, M.R.: An extension of the T-method of multiple comparison to include cases with unequal sample sizes. J. Am. Stat. Assoc. **68**(344), 975–978 (1973)

Effect of Mental Workload and Aging on Driver Distraction Based on the Involuntary Eye Movement

Le Anh Son, Hiroto Hamada, Makoto Inagami, Tatsuya Suzuki and Hirofumi Aoki

Abstract We have shown that the driver distraction can be quantitatively estimated from the vestibulo-ocular reflex (VOR) that is a type of involuntary eye movement. However, the optokinetic reflex (OKR) was not considered in the previous models. Here, we developed a new model with both VOR and OKR. Using the new model, we investigated the effect of mental workload and aging on the involuntary eye movement as well as the driving performance. In this study, we evaluate driver distraction of younger group (age 20–59) and older group (age 60 and above) while driving with/without mental workload. Total 12 participants (6 in each group) who drive on a daily basis participated in the experiment to evaluate driver distraction. As expected, we succeed in applying VOR and OKR models to evaluate driver distraction while driving. Based on that, the effects of mental workload and aging on driver distraction were analyzed. The results indicate that the older group shows worse performance, especially under the distracted driving condition.

Keywords Driver distraction · Vestibulo-ocular reflex · Optokinetic reflex

L. Anh Son (✉) · T. Suzuki
Department of Mechanical Science and Engineering, Nagoya University,
Furo-Cho, Chikusa-Ku, Nagoya 464-8603, Japan
e-mail: leanhsonvn@gmail.com

H. Hamada
Vehicle Engineering Development Division, Toyota Motor Corporation,
Toyota, Japan

M. Inagami · H. Aoki
Institute of Innovation for Future Society, Nagoya University,
Furo-Cho, Chikusa-Ku, Nagoya 464-8601, Japan

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_30

1 Introduction

1.1 Aging Problems

Currently, the aging of population occurs in most countries in the world, especially in developed countries. For example, according to the statistics in Japan, the percentage of people with age 65 and over increases year by year. In 2014, the percentage of people with age over 65 years old is 26.0 %. It increases to 39.9 % in 2060 [1].

As suggested by previous researches, aging cause a complex issue that include transportation especially driver distraction [2–7]. For example, Thompson, et al. found out that older drivers made more driving safety error than middle-aged drivers while completing an audio serial addition task [8]. In order to support old drivers, we need to develop a method to evaluate driver distraction based on behavior.

1.2 Driver Distraction

According to previous researches, eye movement can be simulated by using head motion [9–12]. In case of focusing on one target, when the head turned to the left, the eyes move in the opposite direction to stabilize the visual image based on the input to the vestibular organ. The vestibular system, which is a sensory mechanism in the inner ear, provides the principal contribution to the sense of balance and spatial orientation. The system consists of two parts: the otolith and the semicircular canal. In this research, we applied a VOR model proposed by Merfeld and Zupan [13] which can deal with the interaction between otoliths and the semicircular canal.

Based on the simulation of eye movement, in 2008, Obinata's group has used the VOR model to evaluate mental workload [4] and driver distraction in terms of memory-decision workload [5]. They presented a new method for quantifying mental workloads by utilizing the VOR. Based on the model-based approach, they found that the driver's eye movement deviates from the simulation results by the VOR model under distraction, and distraction can be estimated quantitatively. They did not, however, consider changes in gaze direction of the driver.

Drivers need to do multi tasks, that is obtain information from the environment and control the vehicle. To obtain visual information during movement (not only in vehicle but also during walking, etc.), people need to stabilize the visual image of the environment. As determined by Schweigard & Mergner to reduce the effect of movement, the optokinetic model (OKR) was combined with VOR model [14–17]. In most previous researches on OKR-VOR interaction, passive and active head movements were used. By using active and passive head movements, Schweigard & Mergner created the relationship between visual pattern motions in space with subject's head movements. Based on that information, the negative feedback loop

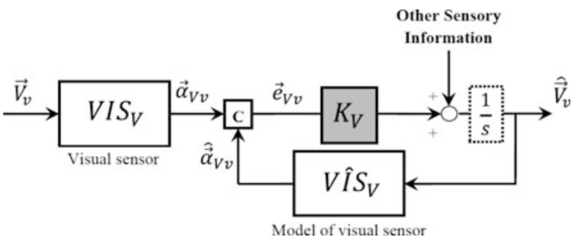


Fig. 1 Generic visual model path way proposed by Newman

was made to stabilize the image on the retina and the VOR only as a useful addition which compensates for the limited bandwidth of the OKR during high frequency/velocity head rotations.

In another way, Newman [16] and Clark et al. [15, 17] dealt with the interaction between VOR and OKR by including static and dynamic visual sensory information from four independent visual sensors (visual velocity, position, angular velocity, and gravity). Because the Merfeld and Zupan's model has only first-order lag characteristics for the eye muscle, we combined it with Robinson's model as a final common-path segment [11]. The processed of visual input is shown in Fig. 1.

2 Method

2.1 VOR—OKR Interaction

In order to consider the interaction between semicircular canal and otolith, in this research, we applied the VOR model proposed by Merfeld and Zupan. The model is shown in Fig. 2.

In this model, linear acceleration (α) and angular velocity (ω) of head-fixed coordinator frame are input. The orientation of gravity (g) with respect to the head is calculated using a quaternion integrator the angular velocity. The otolith is modeled to unify and respond to the gravito-inertial force ($f = g-a$). At the same time, by using high-pass filter of angular velocity, the signal of semicircular canals is calculated. After that, afferent signals from the canals and otoliths are compared in the central nervous system (Internal processing). This model combines four free-parameters (k_{ω}, $k_{f\omega}$, k_f, and k_a). The parameter values are determined by the feedback of this error.

According to Schweigard & Mergner, the visual pattern motion in space and time is matched with subject's head movements. It can be calculated by multiple head position with one parameter (k).

In a similar way with Newman et al. [9, 10, 15], one negative feedback loop was created which is important for image stabilization on the retina and the VOR only as a useful addition which compensates for the limited bandwidth of the OKR during high frequency/velocity head rotations. Detail of the model is shown in Fig. 3.

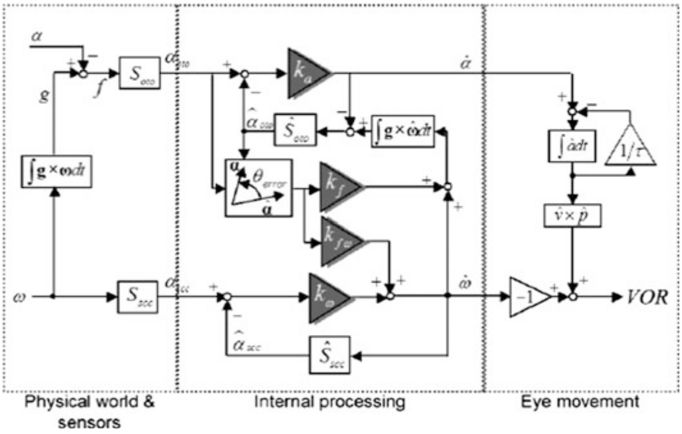


Fig. 2 The VOR model by MATLAB Simulink

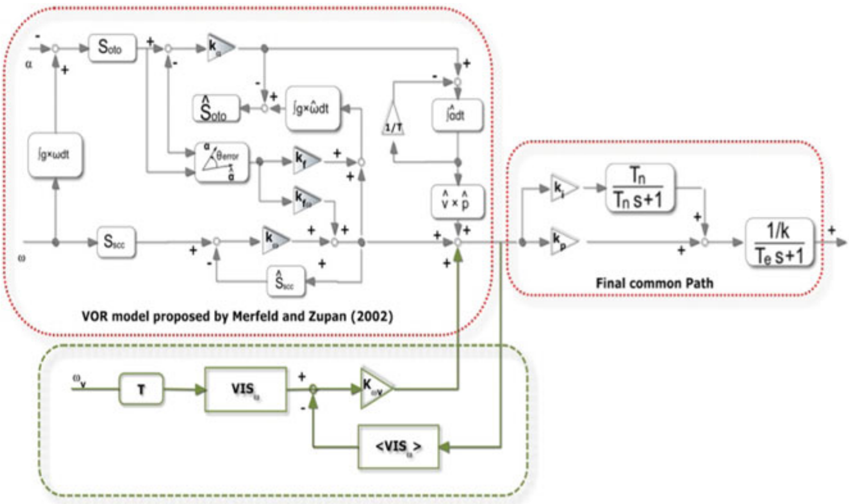


Fig. 3 VOR—OKR model

In case of driving, it is difficult to get target information from vision because the target changes time by time. Therefore, we applied technique from Schweigard and Mergner [14] by using head movement in active condition. The visual information (VS) was calculated based on head movement (HS) following the function: VS = k * HS. VS was used as input of our OKR model.

After that, a visual input is processed by the visual sensor (VIS) to generate a visual sensory estimate. This estimate is compared to an expected visual sensor evaluated from the internal model of visual sensor (<VIS>). Then the difference

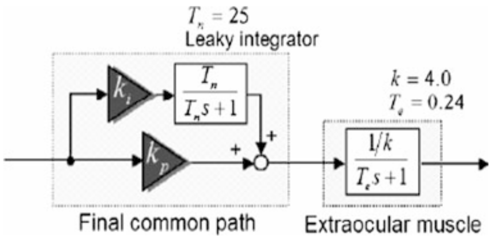


Fig. 4 Final common path model

between visual sensory estimate and the internal model of the visual sensor is weighted with a residual weighting parameter (K_V) and added to the rate of changing of estimated state.

After calculating the eye movement, the final common path proposed by Robinson [11] was applied. In this part, two parameters (k_i and k_p) were used, based on the different types of muscle fibers present in muscles of the eye (Fig. 4).

2.2 Experiment Setup

In the experiment, subjects were asked to drive by following a course on the seat of a driving simulator with six degrees of freedom. The driving simulator had a cylindrical 360° screen of 6 m in diameter. The simulator was controlled by CarSim (Mechanical Simulation Co.) which can simulate the dynamic behavior of a vehicle (Fig. 5). By controlling Carsim with MATLAB Simulink (MathWorks), the seat was moved at a prefixed frequency on the vertical and horizontal planes.

In this study, the participant followed the course with visual stimulus by mean of driving simulator. Eye movement was captured by using SmartEye pro (Smart Eye AB) with four cameras on the dashboard. To collect information on movements of the head, we used a Fastrak electromagnetic tracker (Polhemus Inc.).

Fig. 5 The experimental setup of the driving simulator

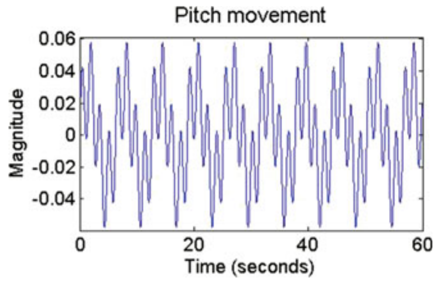


Fig. 6 Input for seat movement

An n-back task was used to create driver distraction. In the present n-back task, a series of one-digit numbers was verbally presented at an interval of 2 s. The subject was asked to judge whether or not each number was equal to preceding one and respond by pressing buttons on the steering wheel.

In order to simulate vibration of the vehicle, the seat was moved by the control of CarSim with MATLAB Simulink. The seat vibration consists of two components: vertical and horizontal. Figure 6 shows the profile of input for the seat vibration.

3 Results and Discussion

Subjects were divided into younger (less than 60 years old) and older (60 years old and above) groups, each of which consisted of 6 subjects.

3.1 VOR and OKR Interaction

To confirm a new model combined VOR and OKR models, we compared the eye movement of two cases: only VOR model and VOR/OKR model without mental workload. By comparing the observed eye movement and the simulation results from the models, the VOR/OKR model showed lesser mean square error than the VOR model only. For example, the mean square error of the observed eye movement and calculated one from the model reduced from 1.1E−03 to 4.3E−04 in Subject 18. We confirmed similar results in other subjects. Moreover, Fig. 7 shows that by combining OKR model, the simulation provides better matching.

As a result, the model consists of both VOR and OKR shows better performance compared with VOR model only.

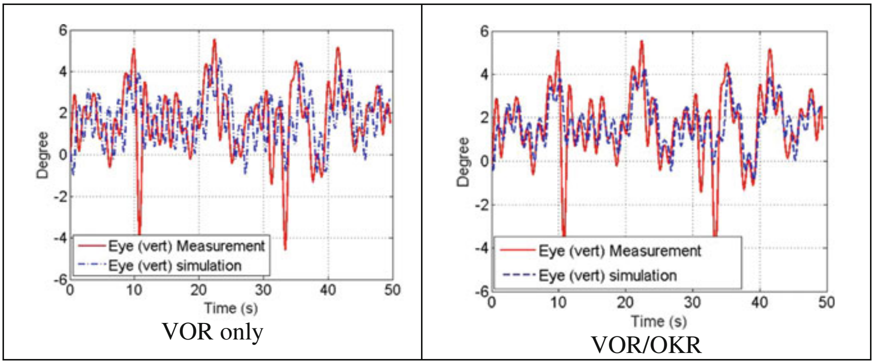


Fig. 7 Eye movement

3.2 *Effect of Driver Distraction on Eye Movement*

In this study, the effect of n-back task on driver distraction was evaluated by using the mean-square error between the measured and simulated eye movement in both vertical and horizontal direction.

All subjects participated in this experiment with two trials: drive without mental workload and drive with mental workload. The eye movement was simulated based on head movement using VOR/OKR model.

Younger Group. Figure 8 shows an example of the eye movement in vertical direction without and with mental workload for total 50 s. In this case, the mean square error for vertical direction without mental workload was 4.3E−04 and it increased to 1.9E−03 with mental workload.

As shown in Fig. 8, the time and frequency response became mismatched with mental workload. Consequently, the mean square error increased from the case without mental workload. We found similar results in other subjects in younger group (Table 1).

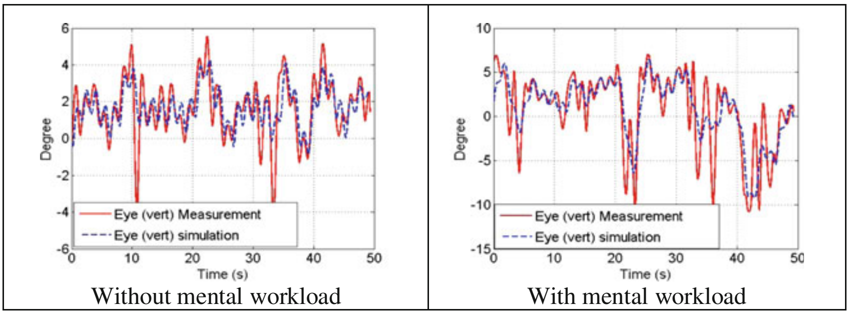


Fig. 8 Driving without/with mental workload (example of Subject 18)

Table 1 Mean square error of each younger subject

Subject	Without mental workload			With mental workload			Ratio of with/without
	Horizon	Vertical	Total	Horizon	Vertical	Total	
S13	4.1E−03	2.0E−03	6.1E−03	4.7E−03	4.7E−03	9.4E−03	1.54
S14	2.0E−03	5.3E−04	2.5E−03	2.7E−03	1.7E−03	4.4E−03	1.74
S17	3.2E−03	1.0E−03	4.2E−03	3.9E−03	1.3E−03	5.1E−03	1.23
S18	1.7E−03	4.3E−04	2.1E−03	2.8E−03	1.9E−03	4.7E−03	2.20
S21	1.0E−03	5.0E−04	1.5E−03	1.4E−03	8.3E−04	2.2E−03	1.46
S22	2.4E−03	7.8E−04	3.2E−03	2.6E−03	1.8E−03	4.4E−03	1.38

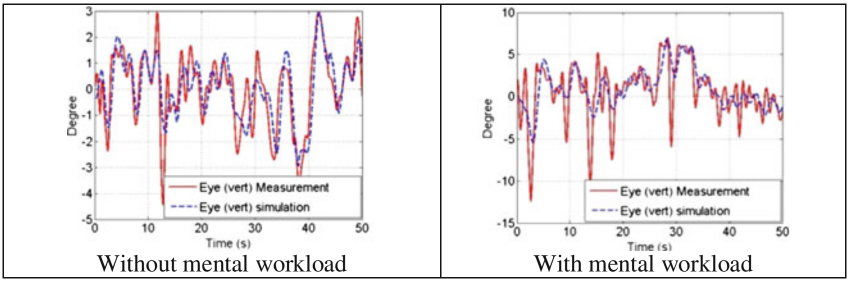


Fig. 9 Driving without/with mental workload (example of Subject 9)

Older Group. The results of older group show the same trend with young people. The mean square error was increased in the condition with mental workload. Figure 9 shows an example of subject 9.

In the case of driving without mental workload, the eye movement of subject 9 matched well in both time and frequency response. Inversely, when the subject 9 drove with mental workload, the simulation became mismatched and it increased the mean square error (Table 2).

Table 2 Mean square error of each older subject

Subject	Without mental workload			With mental workload			With/ without
	Horizon	Vertical	Total	Horizon	Vertical	Total	
S1	2.5E−03	4.0E−04	2.9E−03	3.0E−03	1.2E−03	4.2E−03	1.45
S2	3.5E−03	3.8E−04	3.9E−03	4.0E−03	7.4E−04	4.7E−03	1.22
S4	1.5E−03	1.2E−03	2.7E−03	4.4E−03	1.7E−03	6.1E−03	2.26
S9	1.7E−03	2.3E−04	1.9E−03	2.8E−03	1.4E−03	4.2E−03	2.17
S10	2.3E−03	2.0E−03	4.3E−03	2.4E−03	3.7E−03	6.1E−03	1.42
S12	5.4E−03	9.2E−04	6.3E−03	1.1E−02	2.4E−03	1.4E−02	2.17

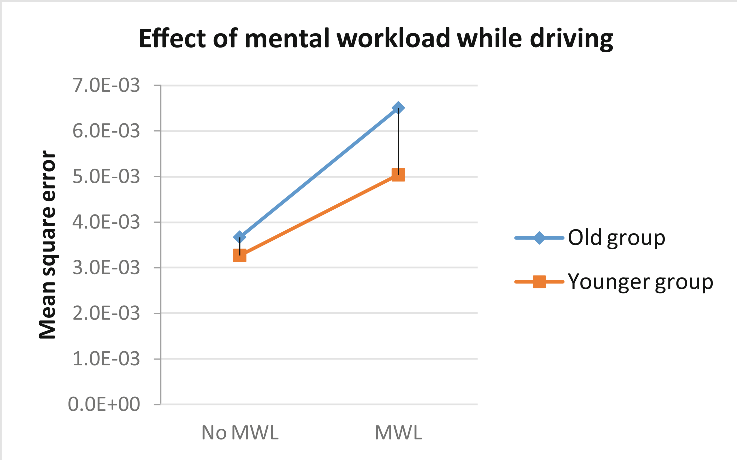


Fig. 10 Effect of metal workload while driving

In the same way with younger group, the mean square errors were higher in the condition with mental workload. The rates of increase in the error from 1.22 times to 2.17 times. We once again confirmed that the older drivers are more distracted by the secondary task.

Cross Effect. Figure 10 compares the increase of the mean square error between older and younger groups. The mean square error of older group increased more sharply than that of younger group (for older group, the average of mean square error from 3.7E−03 without mental workload to 6.5E−03 with mental workload. On the other hand, younger group got small larger with average from 3.3E−03 to 5.0E−03). This means that older people were more distracted when driving with the secondary task.

3.3 Effect of Driver Distraction on Driver Speed

In this experiment, we asked subjects to drive at a speed from 40 to 60 km/h. We analyzed the vehicle speed recorded during experiment.

By using standard deviation (SD) of speed, the results show that with mental workload, almost driver was more difficult to keep the sustainable speed. The SD of speed while driving with mental workload was bigger compare with without mental workload (Fig. 11) except subject 14. In addition, as similar with mean square error, the older people show worse performance compare with younger people with higher SD.

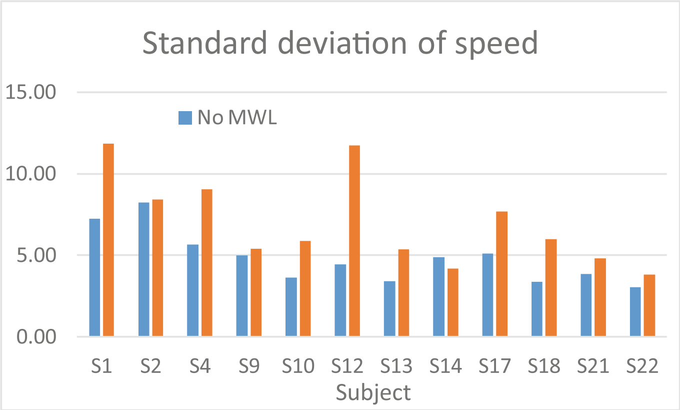


Fig. 11 Standard deviation of the vehicle speed

4 Conclusion

We succeed in applying VOR model and OKR model to evaluate driver distraction while driving. Based on that, the effects of mental workload and aging on driver distraction were analyzed. The results confirm that the older people shows worse performance than younger, especially under distracted driving. It means that they need more support from technology while driving to reduce the mental workload.

References

1. Cabinet Office of Japan, White Paper on Aging in Japan.: Available: http://www8.cao.go.jp/kourei/whitepaper/w-2015/html/zenbun/index.html (2015)
2. Arai, A., Arai, Y.: Self-assessed driving behaviors associated with age among middle-aged and older adults in Japan. Arch. Gerontol. Geriatr. **60**(1), 39–44 (2015)
3. Restrepo, H.E., Rozental, M.: The social impact of aging populations: some major issues. Soc. Sci. Med. **39**(9), 1323–1338 (1994)
4. Stinchcombe, A., Gagnon, S.: Aging and driving in a complex world: exploring age differences in attentional demand while driving. Transp. Res. Part F Traffic Psychol. Behav. **17**, 125–133 (2013)
5. Son, J., Park, M., Park, B.B.: The effect of age, gender and roadway environment on the acceptance and effectiveness of advanced driver assistance systems. Transp. Res. Part F Traffic Psychol. Behav. **31**, 12–24 (2015)
6. Donmez, B., Liu, Z.: Associations of distraction involvement and age with driver injury severities. J. Safety Res. **52**, 23–28 (2015)
7. Curry, A.E., Pfeiffer, M.R., Durbin, D.R., Elliott, M.R.: Young driver crash rates by licensing age, driving experience, and license phase. Accid. Anal. Prev. **80**, 243–250 (2015)
8. Thompson, K.R., Johnson, A.M., Emerson, J.L., Dawson, J.D., Boer, E.R., Rizzo, M.: Distracted driving in elderly and middle-aged drivers. Accid. Anal. Prev. **45**, 711–717 (2012)

9. Merfeld, D.M., Good, L., Hospital, S.: Modeling the vestibulo-ocular reflex of the squirrel monkey during eccentric rotation and roll tilt. Exp. Brain Res. **106**(1), 123–134 (1995)
10. Zupan, L.H., Merfeld, D.M.: Neural processing of gravito-inertial cues in humans. IV. Influence of visual rotational cues during roll optokinetic stimuli. J. Neurophysiol. **89**, 390–400 (2003)
11. Robinson, D.A.: The use of control systems analysis in the neurophysiology of eye movements. Annu. Rev. Neurosci. **4**, 463–503 (1981)
12. Merfeld, D.M., Zupan, L.H., Vestibular, J., Eye, M., Infirmary, E.: Neural processing of gravitoinertial cues in humans. III. Modeling tilt and translation responses. J. Neurophysiol. **87** (2), 819–833 (2002)
13. Merfeld, D.M., Zupan, L.H.: Neural processing of gravitoinertial cues in humans. III. Modeling tilt and translation responses. J. Neurophysiol. **87**, 819–833 (2002)
14. Schweigart, G., Mergner, T., Evdokimidis, I., Morand, S., Becker, W.: Gaze stabilization by optokinetic reflex (OKR) and vestibule-ocular reflex (VOR) during active head rotation in man. Vision. Res. **37**(12), 1643–1652 (1997)
15. Clark, T.K., Newman, M.C., Oman, C.M., Merfeld, D.M., Young, L.R.: Human perceptual overestimation of whole body roll tilt in hypergravity. J. Neurophysiol. **113**(7), 2062–2077 (2015)
16. Newman, M.C.: A multisensory observer model for human spatial orientation perception, Master dissertation (2009)
17. Clark, T.K., Newman, M.C., Oman, C.M., Merfeld, D.M., Young, L.R.: Modeling human perception of orientation in altered gravity. Front. Syst. Neurosci. **9**, 68 (2015)

A Study on the Positioning of a Mounted Mobile Phone to Reduce Distraction While Driving Among Young Adults

Angelique Mae Alconera, Lakan Garcia, Jeremy Christine Mercado and Alyssa Jean Portus

Abstract This study examined three locations for mounting a mobile phone inside a vehicle to minimize distraction while driving. The locations considered were the three most preferred locations based on a survey among 76 young adult drivers. Nineteen right-handed drivers aged 19–25 participated in a series of visual, auditory, and tactile distraction tests while driving using a simulator. Total distraction time, response accuracy, and the effect to driving skill were recorded for each test. At $\propto = 0.10$, the mean time that the drivers' hands were off the steering wheel was found to be different among the locations (P-value = 0.096). Further analyses suggest that the samples taken from the top-left of the steering wheel do not differ (P-value = 0.6614) with the most favorable location, the top-right of the steering wheel. This implies that driving distraction is minimal when the mobile phone is mounted in either of these two locations.

Keywords Ergonomics · Philippine driving · Mobile phone · Driving distractions

1 Introduction

Driving is a complex activity that requires focus and attention. Drivers monitor several things simultaneously including regulating speed, maintaining a safe following distance, looking out for incoming vehicles, pedestrians and traffic signs. Drivers act on all these while making sure that safety rules and regulations are

A.M. Alconera (✉) · L. Garcia · J.C. Mercado · A.J. Portus
University of the Philippines—Diliman, Quezon City 1101, Philippines
e-mail: giquealconera@gmail.com

L. Garcia
e-mail: lakandgarcia@gmail.com

J.C. Mercado
e-mail: christine.mercado2012@gmail.com

A.J. Portus
e-mail: aaportus@up.edu.ph

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_31

observed. As much as driving already seems like an arduous task for many, drivers in the Philippines in particular should take more precaution. Driving conditions in the country are more demanding as drivers and pedestrians lack the discipline necessary for a peaceful and well-coordinated thoroughfare.

Despite the unfavorable driving conditions, according to the President's 6th State of the Nation Address, car sales in the Philippines grew by 27 % from 2014 to 2015. Albeit being an indicator of economic boom, this may be a cause of worry given the notorious lack of discipline of many Filipino drivers. The Association for Safe International Road Travel even deems bus and jeepney drivers as downright reckless.

1.1 Mobile Phones as an Inevitable Part of the Filipino Lifestyle

Despite the increased awareness of the dangers of using mobile phones while driving, the activity still remains prevalent among drivers. This can be linked to the fact that mobile phones are now considered a necessity, especially by individuals from the Generation Z. In a study conducted by Siddiqui [1], 89 % of respondents aged 20–25 years old said that they consider cell phones a necessity rather than a luxurious item. On the other hand, 43 % of the respondents claimed that they use their mobile phones almost always, sometimes unconsciously, and 63 % of the respondents feel that they feel incomplete without their mobile phones.

Aside from mobile phones being one of the most accessible media of communication, it also has features and applications that are deemed essential for the lifestyle of some. As can be observed, many drivers use various mobile phone applications while driving specially to navigate through the streets. Evidently, although the usage of mobile phones while driving poses serious safety risk, they are still a necessity and the usage of such while driving is inevitable.

As the need for using a mobile phone while driving becomes unavoidable, many drivers have opted to mount their mobile phones on various locations in the vehicle such as their dashboards, radio consoles, near the rear-view mirror or windshields. Currently, there is still no standard location wherein the phone should be placed. Thus, this study would like to answer the question, "Where is the position of mounting the mobile phone in a vehicle such that distraction is minimized?"

1.2 Scope and Limitation

This study is limited only to drivers who use a regular-sized mounted mobile phone. The findings in this study do not necessarily apply to mounted tablets or laptop computers that may also be used for navigation. Furthermore, this study is limited to only one model of mobile phone and one sedan type vehicle. In addition, this study focused only on drivers aged 19–25. Young and Regan [2] found in their research that older drivers often choose not to divert attention while driving. Lastly, only Filipino drivers who are right-hand dominant and have at least one year of driving experience were considered in this study. The aspects looked into were only the auditory, visual and tactile distractions of driving. These were the senses chosen as these are the ones commonly used while using a navigation application.

2 Review of Related Literature

In an experiment conducted by Strayer and Johnston [3] wherein participants were made to perform a simulated-driving task (pursuit-tracking) under different scenarios, drivers who were engaged in cellphone conversations missed twice as many simulated traffic signals regardless if the phone is handheld or not. Tracking error increased when drivers used cellphones to perform attention-demanding word generation tasks. Both tasks divert the attention of the driver into an engaging cognitive context which disrupts driving performance.

In a study on the effects of cell phone use on peripheral vision [4], it was determined that cell phone conversations tend to artificially constrict the visual field of the driver. This suggests that cell phone use will decrease the perceptual visual field, making the driver less aware of one's environment, consequently increasing the chances of accidents.

The same detrimental effect is observed even when the driver is simply interacting with a voice-based messaging system. In particular, the ability of a driver to do a complete scan of an intersection reduced by 11 % when they were conversing with a voice-based console [5]. It can be seen through these that even though the secondary task of the driver mainly draws its resource pool from the auditory sense, such an activity also had an effect on the visual perception of the driver.

Another activity that a driver usually engages in is manually typing or inputting while driving. In a study by Cheung [6] regarding the effects on driving performance and text messaging behavior when tasks are done simultaneously, he found a serious degradation of vehicle control and increase in lateral and lane deviation compared to when the driver is simply focused on driving. Furthermore, the given task, which was to enter an address, took significantly longer to complete. Regan et al. [7] also performed different experiments regarding this matter. It was found out that manually dialing a phone number has the most detrimental effect on the performance of the driver.

These studies may mean that the time where at least one of the hands is not on the steering wheel may lead to making the driver more distracted. This, as well as the auditory and visual effects of dual-tasking while driving, is taken into consideration in the method of gathering the data for this study.

In sum, this literature review shows the adverse effects of dual-tasking through cellphone use while driving and the potential danger that it brings. These dangers come from the diversion of attention from performing demanding tasks simultaneously. The study aims to address this issue through the concept of multiple resource theory and statistical models.

3 Methodology

Eleven male participants and eight female participants participated in the study. All participants have at least one year of driving experience, are right hand dominant, and are 19–25 years of age.

3.1 Setup

Toyota's Vios model was used in the study as it is the most common vehicle purchased in the Philippines. In addition, the study conducted by Mohamed and Yusuff [8], proves that the Toyota Vios is the best among its close competitors in terms of spaciousness, steering adjustability, driver reach ability to surrounding components, driver view, vibration, and noise among others.

The dimensions of the standard car were replicated in an indoor setting. In particular, the distance from the steering wheel to the window, distance from the car seat to the radio component, and the distance from the steering wheel to the gear stick were taken and were used in the setup. The distance from the car seat to the steering wheel, window, and radio component were not computed as the car seat can be adjusted to suit the comfort of the driver. The distance between the steering wheel and the car seat were decided upon by the participant. To drive through the prepared simulator, the participants made use of a modified steering wheel.

The mobile phone used in the setup is an ASUS Zenfone 2 model. This was mounted on a stand and was placed in different pre-set locations. To reduce variability among tests, the same model of mobile phone and phone mode, volume and brightness were maintained in all tests. The participant adjusted the angle of inclination of the mobile phone, as is the situation in real life. The experimental setup is shown in Fig. 1.

Fig. 1 The experimental setup was made of a makeshift steering wheel, a mobile phone mounted using a tripod, a laptop and projector which allowed the user to view and interact with the driving simulator

3.2 Location of Mount

Nine different locations were initially chosen for the study. These locations were selected as they are either in the same horizontal plane as the usual location where drivers hold the steering wheel, the same horizontal plane as the normal line of sight of a driver or the same vertical plane as the console of the car. The figure below shows the locations where the mounted mobile phone is to be located (Fig. 2).

In order to determine the three (3) most preferred location of the mounted phone, a survey was conducted to 76 individuals who are 19–25 years of age and who have at least one year of driving experience. The participants of the survey were asked to rank the locations in terms of preference, where 1 is the most preferred and 9 the least preferred. Table 1 below shows that locations 6, 4, and 8 were the most preferred. These three locations were used in the driving simulation in the setup shown above.

Location 6 is the area to the top-right of the steering wheel, herein referred to as the Top location. Location 4 is the area to the top-left of the steering wheel, herein referred to as the Left location. Lastly, location 8 is the area to the right of the steering wheel, which is referred to as the Right location.

Fig. 2 Proposed locations of mounted mobile phone

Table 1 Average rankings of locations according to preference of drivers

Location	Average	Rank
1	6.11	8
2	6.55	9
3	6.05	7
4	3.82	2
5	4.46	4
6	3.51	1
7	5.75	6
8	4.07	3
9	4.68	5

3.3 Tests

In order to determine the most ergonomic location of the mounted mobile phone, participants were subject to dual-task activities and the effects of the distractions were observed and recorded. The distractions fall into the following categories: auditory, visual, and tactile.

The participants played the 3D Speed Driver game. While doing the tests, the participants were instructed to avoid all barriers along the way. The number of barriers hit was recorded to account for the general driving skill of each participant. Before the actual tests, participants were allowed to play the game until they confirm that they have become familiar and comfortable with the controls. This is done so that the results can be directly attributed to the distractions, not by the driving task.

3.4 Visual Distraction Test

In visual distraction happens when drivers do not have their full attention at the road. While using navigation applications, drivers often glance at their mobile phones for various reasons like monitoring routes, confirming road blocks, and noting estimated arrival times. These examples of visual distraction were accounted for. This was conducted by letting the participant verbally respond to a simple question displayed as a text message. The questions were "Is there a clock in front?", "Is there a barricade in front", and "Are there wheels in front?".

If the item in question was visible on the lane where the participant is currently in, the participant was expected to answer with a 'Yes'. Otherwise, the answer was to be a 'No'. Once the question popped up, the driver was expected to immediately respond to the stimuli. The questions specified above appeared at different intervals. The total time in the game that the participants' eyes looked at the mobile phone, read the question, and went back to the road were recorded. The responses of the participants were also recorded, regardless of whether they were accurate or not.

3.5 Auditory Distraction Test

Navigation applications commonly use artificial voice commands to give drivers directions or notify them of road hazards. Drivers divert their attention from the road to comprehend and react to these commands. In this part of the test, the participants were tasked to respond to turning commands that were given via a mobile phone speaker. For simplicity, the drivers were instructed to immediately follow the command regardless of any immediate roadblocks.

Similar to the previous test, commands were given to the participant three times. These commands were simply to "Turn Right" or "Turn Left". The reaction time of each participant towards each command was recorded.

3.6 Tactile Distraction Test

Distraction due to physical stimuli also arises from the use of navigation applications. This happens when the driver does tasks that require taking at least one hand off the steering wheel to interact with the mobile phone. An example of this activity is when the driver zooms in the map while navigating. This distraction was simulated by instructing the participants to input patterns on the mobile phone. The participants were asked to input the patterns at random intervals.

The amount of time that the participant's fingers and eyes were on the screen of the mobile phone was recorded. In addition, the number of times that the participant made corrections while inputting the patterns was recorded.

3.7 Measurement Error

All of the tests were video recorded. Reaction times were obtained using a stopwatch. This method may be prone to error as the recorded times have very short durations; usually less than a second. To reduce the error, data points were taken three times and the averages of these recordings were used for analysis.

4 Results and Discussion

Kolmogorov-Smirnov Test for Normality was used to determine if the factors taken in each distraction test come from a normal distribution. Using Minitab Statistical Software, the P-values taken from each test were determined and are shown in Table 2. With $\alpha = 0.10$, only the samples in factor F were proven to come from a Normal Distribution.

As this factor satisfies the assumptions for a parametric analysis, ANOVA at $\alpha = 0.10$ was used to determine if there is any significant difference among the locations. From this, a *P*-value of 0.187 is derived. This implies that there is no significant difference between the three locations.

The rest of factors do not have a normal distribution, hence were analyzed using Nonparametric Statistical tests. The results of Kruskal-Wallis Test at $\alpha = 0.10$ are shown in Table 3. From this, it can be seen that there is a significant difference among locations only in factor G (*P*-value = 0.096). Consequently, no further analysis was done for other factors.

Table 2 Results of Kolmogorov-Smirnov test for normality

Test	Index	Factor	P-value	Conclusion
Auditory	A	Reaction time	0.028	Nonnormal
Auditory	B	Roadblocks hit	<0.01	Nonnormal
Visual	C	Amount of time eyes focused on the phone	<0.01	Nonnormal
Visual	D	Roadblocks hit	<0.01	Nonnormal
Visual	E	Missed questions	<0.01	Nonnormal
Tactile	F	Amount of time eyes focused on the phone	>0.15	Normal
Tactile	G	Amount of time hands are on the phone	<0.01	Nonnormal
Tactile	H	Average repetitions needed to input pattern	<0.01	Nonnormal
Tactile	I	Roadblocks hit	<0.01	Nonnormal

Table 3 Results of Kruskal-Wallis tests

Test	Index	Factor	*P*-value	Conclusion
Auditory	A	Reaction time	0.267	Nonnormal
Auditory	B	Roadblocks hit	0.482	Nonnormal
Visual	C	Amount of time eyes focused on the phone	0.460	Nonnormal
Visual	D	Roadblocks hit	0.937	Nonnormal
Visual	E	Missed questions	0.288	Nonnormal
Tactile	G	Amount of time hands are on the phone	0.096	Nonnormal
Tactile	H	Average repetitions needed to input pattern	0.813	Nonnormal
Tactile	I	Roadblocks hit	0.322	Nonnormal

Nonparametric tests are generally limited compared to tests for samples that have a normal distribution. Because of its limitation, instead of comparing means, only the medians of the samples for this factor were compared. To do this, one location is first considered to be the "best" location then the other two locations were compared to it.

Clearly, it is more desirable that the amount of time wherein the hand of the driver is on the mobile phone is minimized. Among the three locations, the least median amount of time is spent when the mobile phone is mounted in the top location (median of 1.7617 s), followed by the left location (median of 1.9074 s) and by the right location (median of 1.9686 s). From this, the top location is considered to be the "best" location in terms of tactile distraction.

Mann-Whitney Tests at $\alpha = 0.10$ shown in Table 4 indicate that there is no significant difference in the medians of all the data samples taken at the left location

Table 4 Results of Mann-Whitney tests

Locations to compare	*P*-value	Conclusion
Left vs. top	0.6614	No significant difference between location
Right vs. top	0.0410	There is significant difference between locations

and the top location. This means that the levels of tactile distraction that the participants experienced are statistically equal and are minimal when the mobile phone is placed at these locations.

Overall, the result of the tests and statistical analyses that follow show that at $\alpha = 0.10$, only the amount of time that the hands of the driver are on the mobile phone proved to be significantly different among the three different locations that were tested. Among these, the left and top locations appear to be the locations where the drivers experience the least amount of distraction while driving.

These results can be used as baseline data for further studies for determining the optimal location of the mobile phone. It is recommended that more samples be taken and that more locations be considered while conducting further studies. In addition, it is suggested that an $\alpha = 0.05$ be used instead of α or 0.10 to draw more accurate and precise results and conclusions.

5 Conclusion

The amount of time that the driver's hands are on the mobile phone has a significant effect on driver distraction. The location with the least amount of time was the top location, which was then used as a basis of comparison. Furthermore, no significant difference was found between the left and top locations at a 0.1 level of significance. This implies that the most favorable position of mounting a mobile phone in a vehicle that will reduce distraction is either to the top-left or the top-right of the steering wheel.

References

1. Siddiqui, S.I., Jabeen, S., Mumtaz, M.: Whether Cell Phone Is a Necessity or a Luxurious Item? Middle-East J. Sci. Res. **19**(1), 61–65 (2014)
2. Young, K., Regan, M.: Driver distraction: a review of the literature. In: Faulks, I.J., Regan, M., Stevenson, M., Brown, J., Porter, A., Irwin, J.D. (eds.) Distracted Driving, pp. 379–405. Australasian College of Road Safety, Sydney (2007)
3. Strayer, D.L., Johnston, W.A.: Driven to distraction: dual-task studies of simulated driving and conversing on a cellular phone. Retrieved from http://www.weber.edu/wsuimages/psychology/Research/ResearchArticles/Psychological%20Science-2001-Strayer-462-6.pdf (2001)
4. Maples, W., DeRosier, W., Hoenes, R., Bendure, R., Moore, S.: The effects of cell phone use on peripheral vision. Optometry, 36–42 (2009)
5. Biondia, F., Turrill, J., Coleman, J., Cooper, J., trayer, D. (n.d.). Cognitive distraction impairs drivers' anticipatory glances: an on-road study. In: Proceedings of the Eighth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 22–29
6. Cheung, D.: Text messaging while driving: its effects on driving performance and text messaging behaviour. Faculty of Life Sciences, London (2010)

7. Driver distraction as a breakdown of multi-level control. In: Regan, M.A., Lee, J.D., Young, K. (eds.) Driver Distraction: Theory, Effects and Mitigation, pp. 41–56. CRC Press, Boca Raton, FL
8. Mohamed, Z., Yussuf, R.M.: Automotive ergonomics: passenger cars interior. In: International Conference on Ergonomics. Kuala Lumpur (2007)

Characterization of Driver Perception Reaction Time at the Onset of a Yellow Indication

Ihab El-Shawarby, Hesham Rakha, Ahmed Amer and Catherine McGhee

Abstract The research presented in this paper characterizes driver perception-reaction times (PRTs) in a controlled field environment at the onset of a yellow-indication transition in high-speed signalized intersection approaches. The study characterized the impact of driver gender, driver age, roadway grade, mean approach speed, platooning scenarios (leading, following, or alone), and time-to-intersection (TTI) on the driver PRT. This characterization is critical for the efficient and safe design of traffic signal clearance timings. The study demonstrates that the driver PRT is higher for female and older drivers (60 + age group) as compared to male and younger drivers. The PRT is larger when vehicles travel along an upgrade section. Driver PRTs are typically higher if they are following a vehicle that runs a yellow light. Furthermore, driver PRTs decrease when they are followed by another vehicle. Finally, driver PRTs increase as the TTI at the onset of the yellow interval increases.

Keywords Perception reaction time · Driver behavior · Traffic safety · Traffic signals

I. El-Shawarby · H. Rakha (✉) · A. Amer
Virginia Tech Transportation Institute, Blacksburg, USA
e-mail: hrakha@vt.edu

I. El-Shawarby
Faculty of Engineering, Ain-Shams University, Cairo, Egypt

H. Rakha
Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, USA

C. McGhee
Virginia Center for Transportation Innovation and Research, Charlottesville, USA

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_32

1 Introduction

The driver perception-reaction time (PRT) is of significant importance in highway design. For example, it is used to estimate the stopping sight distance in the computation of horizontal and vertical profiles in highway design [1] and also used in the calculation of the yellow interval duration in traffic signal design [2]. At the onset of a yellow-indication transition on high-speed signalized intersections, a driver decides to either stop safely or to proceed through the intersection before the end of the yellow interval. Accordingly, the proper design of traffic signals requires the computation of a yellow interval that entails an estimate of the driver's PRT. The state-of-practice in different dilemma zone alleviation strategies typically recommends a 1.0 s PRT, which is assumed to equal or exceed the 85th percentile brake PRT [3]. However, a number of studies have demonstrated that brake PRTs are much longer than 1.0 s and that the 85th percentile PRT is more in the range of 1.5–1.9 s [4]. These studies have also demonstrated that the PRT on high-speed intersection approaches (greater than 64 km/h or 40 mile/h) are lower, with 85th percentile PRTs in the range of 1.1–1.3 s.

A review of the fundamental concepts associated with reaction times will give a better understanding of the factors related to PRT. A comprehensive survey by Green [5] summarized PRT results from most studies published until 2000. The study concludes that when a driver responds to the onset of a yellow light at a traffic signal, the total reaction time can be split into a mental processing (perception) time which is the time required for the driver to decide on a response and a movement (reaction) time which is the time used for lifting the foot from the accelerator and touching the brake pedal. Because the mental processing time is an internal quantity that cannot be measured directly and objectively without a physical response, it is usually measured jointly with movement time. Furthermore, Green demonstrated that older female drivers respond more slowly than younger male driver.

A study by Caird et al. [6] which used a driving simulator to analyze the behavior of 77 drivers approaching signalized intersections at speeds of 70 km/h concluded that the PRT did not differ by age and was affected by time-to-intersection (TTI). The PRT grand mean was 0.96 s, ranging from 0.86 s for drivers closest to the intersection stop line to 1.03 s for drivers farthest from it. In another study [7], conducted by the present investigators, the authors used data gathered from a controlled field test on 60 test participants to characterize driver brake PRTs at the onset of a yellow indication at a high-speed signalized intersection approach. Participants were instructed to drive the car at 72 km/h (45 mph) on a closed test course with no other vehicles. The study concluded that the 1.0 s, 85th percentile PRT that is recommended in traffic signal design procedures is valid and consistent with field observations and demonstrated that either a lognormal or beta distribution is sufficient for modeling brake PRT. Another study by Gates et al. [8] recorded vehicle behavior at six signalized intersections in the Madison, Wisconsin area comprising 463 first-to-stop and 538 last-to-go records. The study found that addition of potential predictor variables related to the activity of other vehicles nearby had statistical significant correlation with stop–go activity. The analysis of the brake-response time for first-to-stop vehicles showed that the 15th, 50th, and 85th percentile brake-response times were 0.7, 1.0, and 1.6 s, respectively.

Previous studies have either been conducted in a driving simulator, or were gathered in a controlled field environment without considering the interaction with other vehicles, or gathered from the field by randomly recording driver behavior. This study extends previous research efforts by characterizing driver PRT behavior in a controlled field environment considering different approach speeds and the influence of other vehicles on driver response at the onset of a yellow indication on high-speed signalized intersection approaches.

The objectives of this study is to characterize driver PRT behavior to investigate the need to design driver and situation-specific yellow times as opposed to using the current approach of one size fits all yellow times. This characterization is critical in the efficient and safe design of traffic signal clearance timings within the emerging Connected Vehicle initiative to provide driver-specific in-vehicle warnings and communicate this information over wireless networks to other vehicles and to various roadside infrastructures.

2 Experimental Design

The field experiment involved in this paper analysis was conducted at the Virginia Department of Transportation's (VDOT) Virginia Smart Road facility, located at the Virginia Tech Transportation Institute (VTTI). It is a 3.5 km (2.2 mile) two-lane road with one four-way signalized intersection and with limited traffic access. The horizontal layout of the test section is fairly straight, and the vertical layout has a substantial grade of 3 % [9]. Because participants turned around at the end of each run, half of the trials run by each participant were on a 3 % upgrade and the other half were on a 3 % downgrade.

Three vehicles were used in the study, one was driven by test participants (accompanied by the in-vehicle experimenter) and the other two vehicles were driven by two research assistants. One of them was either leading or following the test vehicle, whereas the other vehicle was crossing the intersection from the conflicting approach when the traffic light was green. A real-time data acquisition system (DAS) was installed inside the trunk of a 2004 Chevrolet Impala. The vehicle was also equipped with a differential GPS unit, a longitudinal accelerometer, sensors for accelerator position and brake application, and a computer to run the different experimental scenarios. The data recording equipment had a communications link to the intersection signal control box that synchronized the vehicle data stream with changes in the traffic signal controller. Phase changes were controlled from the instrumented car using the GPS unit to determine the distance from the intersection and a wireless communications link to trigger the phase changes. Twenty-four licensed drivers were recruited in three equal age groups (under

40-years-old, 40 to 59-years-old, and 60-years-old or older); equal numbers of males and females were assigned to each group. The experiment involved test-track driving, for six sessions, once per day, where each participant was assigned to six different test conditions. The different test conditions were based on two instructed vehicle speeds of 72.4 km/h (45 mile/h) and 88.5 km/h (55 mile/h) and three platoon conditions (leading, following, and no other vehicle).

Participants drove loops on the Smart Road, crossing the four-way signalized intersection where the data were collected, 24 times for a total of 48 trials, where a trial consists of one approach to the intersection. Among the 48 trials, there were 24 trials in which each yellow trigger time to stop-line occurred four times. On the remaining 24 trials the signal indication remained green. This scheme would result in yellow/red signals being presented on 50 percent of the 48 trials; conversely, 50 % of intersection approaches would be green indication. To examine whether willingness to stop varies with speed, the onset of yellow was based on the time-to-stop line (between 2.0 and 4.6 s) at the instructed speed rather than on distance from the stop line.

A 4-s yellow indication at the 72.4 km/h (45 mile/h) instructed speed and a 4.5-s yellow indication at the 88.5 km/h (55 mile/h) instructed speed were triggered for a total of 24 times (four repetitions at six distances). The yellow indications were triggered when the front of the test vehicle was 40.2, 54.3, 62.5, 70.4, 76.5, and 82.6 m (132, 178, 205, 231, 251, and 271 ft) from the intersection for the 72.4 km/h (45 mile/h) instructed speed and 56.7, 76.2, 86, 93.6, 101, and 113 m (186, 250, 282, 307, 331, and 371 ft) for the 88.5 km/h (55 mile/h) instructed speed to ensure that the entire dilemma zone was within the range.

3 PRT Measurements and Analysis

Video and vehicle performance data were assembled digitally from the test vehicle DAS. Video frame, driver's information (subject number, age, and gender), platoon, trial number, and condition were reported in the data file. The data that were gathered included but were not limited to: current state and duration of the traffic signal, vehicle heading, speed, acceleration, distance to intersection, brake application (on/off).

Although the drivers were instructed to drive at either 72.4 km/h (45 mile/h) or 88.5 km/h (55 mile/h), the instantaneous approach speeds at the yellow interval onset varied considerably. In the case of the 72.4 km/h (45 mile/h) instructed speed, speeds varied from 64.5 to 88.7 km/h (40.1–55.1 mile/h), with a mean of 74.4 km/h (46.2 mile/h), a median of 74.4 km/h (46.2 mile/h), and a standard deviation of 2.3 km/h (1.4 mile/h), as illustrated in Fig. 1a. Alternatively, in the case of the 88.5 km/h (55 mile/h) instructed speed, the test vehicle speed varied from 73.5 to 97.5 km/h (45.7–60.6 mile/h), with a mean speed of 89.3 km/h (55.5 mile/h), a median of 89.5 km/h (55.6 mile/h), and a standard deviation of 2.7 km/h (1.7 mile/h), as illustrated in Fig. 1b. Consequently, the average approach

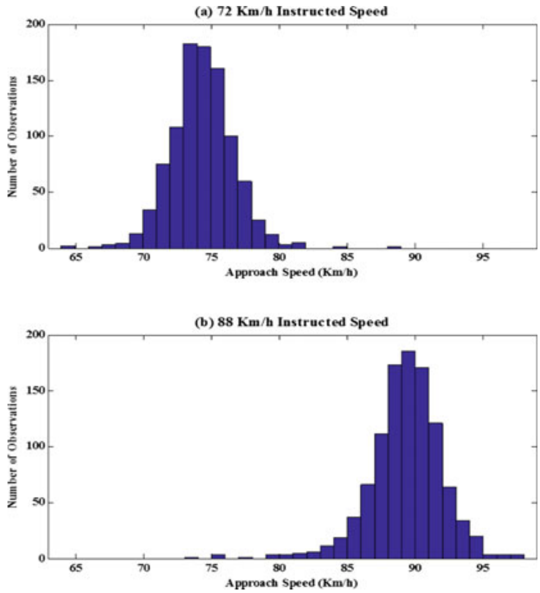


Fig. 1 Histogram of perception–reaction times

speed tended to be slightly higher than the instructed speed in both cases. The histograms of the two approach speeds appear very close to a normal distribution.

A total of 2016 data records were available for analysis, of which 971 data records were for those who were instructed to drive at 72.4 km/h (45 mile/h) and 1045 observations were for a speed of 88.5 km/h. The TTI for the 72.5 km/h instructed speed ranged from a minimum of 1.93 s to a maximum of 4.69 s with a mean equal to 3.59 s, a median of 3.66 s, and a standard deviation of 0.45 s. The remaining 1045 data records for an instructed speed of 88.5 km/h (55 mile/h) included TTIs ranging between 2.31 and 5.33 s with a mean equal to 3.98 s, a median of 3.99 s, and a standard deviation of 0.53 s. In other words, drivers decelerated at farther distances when they traveled at higher speeds.

A study of driver PRTs at the onset of the yellow indication was performed considering all stopping events. As was noted earlier, it was possible to determine the approach speed to the intersection, the TTI, and the brake application of the stopping vehicles from the data files. Using this information, PRT characteristics and profiles were examined. PRT was defined as the time elapsed between the onset of the yellow indication and the instant the driver started to press the brake pedal. The PRT ranged from a minimum of 0.22 s to a maximum of 1.52 s with a mean of 0.73 s, a median of 0.72 s, and a standard deviation of 0.18 s for the participants who were instructed to drive at 72.4 km/h (45 mile/h). Alternatively, the PRT ranged between 0.18 and 1.53 s with a mean equal to 0.74 s, a median of 0.72 s, and a standard deviation of 0.18 s for those who were instructed to drive at 88.5 km/h (55 mile/h). The histogram for the observed PRTs of the 2016 stopping

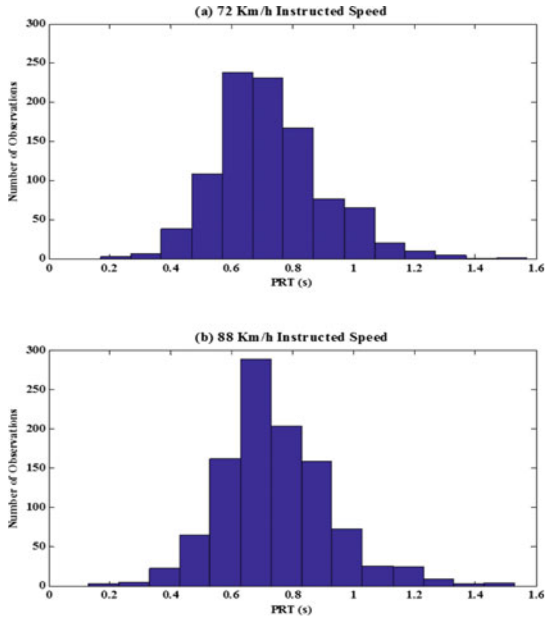


Fig. 2 Histogram of perception–reaction times

events for the two approach speeds is shown in Fig. 2. These figures demonstrate that the driver PRTs were very similar for both instructed speeds.

The observed 15th, 50th, and 85th percentile PRTs were 0.57, 0.72, and 0.92 s, respectively, in the case of the 72.4 km/h instructed speed; and were 0.58, 0.72, and 0.92 s, respectively, in the case of the 88.5 km/h instructed speed. The 85th percentile is consistent with earlier studies which showed that the 85th percentile PRT in high-speed intersection approaches (greater than 64 km/h or 40 mile/h) are in the range of 1.1–1.3 s. The lower value of PRTs in the current study can be attributed to the fact that the PRT was defined from the instant the signal indication changed to yellow until the driver touched the brake pedal, and not when the brake light was activated as in most studies; thus PRT did not include the time lag from the instant the driver presses the brake pedal until the brake lights activate. Consequently, the results appear to be consistent with other naturalistic field study findings. The data were sorted based on the driver's TTI, at the yellow-indication onset, into equal sized bins (equal number of observations) and the average TTI and PRT for each bin was computed (for illustration purposes only) given the large number of observations.

A general linear model procedure (GLM) was conducted using the SAS software to investigate the effects of the TTI, grade (uphill and downhill), age group (under 40-years-old, between 40 and 59-years-old, and 60 years of age or older), gender (male and female), and platoon (leading, following, and no other vehicle), for the two instructed speed levels (72.4 and 88.5 km/h) on PRT. The results showed that the TTI had a significant effect ($P < 0.0001$) on PRT for both the 72.4 km/h (45 mile/h) instructed speed and the 88.5 km/h instructed speed. The mean PRT

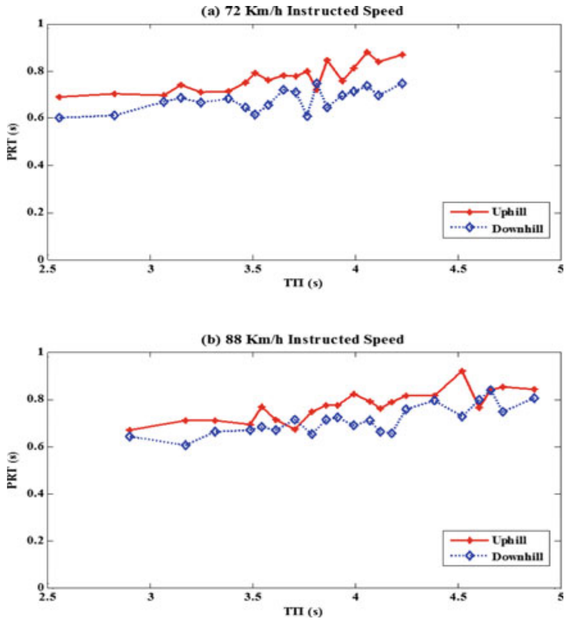


Fig. 3 Effect of roadway grade on perception–reaction times

estimates at each TTI for the two instructed speed levels and approaches (uphill and downhill) were used to illustrate various trends and effects, as demonstrated in Fig. 3. The results showed that the PRT on either approach (i.e., on upgrade or downgrade) exhibit similar trends, with slightly higher PRTs in the case of the uphill approach. This difference, which is significant, demonstrates that drivers traveling uphill might be pushing harder on the accelerator and take longer to release their push and move their feet to press the brake pedal while drivers traveling downhill might not be pressing the accelerator in order to maintain some desired speed and are more alert because they realize that the deceleration level needed to stop the vehicle is greater. For TTIs in the range of 1.93–4.69 s, for the 72.4 km/h (45 mile/h) instructed speed, the PRT ranged from 0.22 to 1.67 s for the uphill approach, and from 0.23 to 1.27 s for the downhill approach. Similarly, for the 88.5 km/h instructed speed, the PRT ranged from 0.22 to 1.53 s and from 0.18 to 1.43 s for the uphill and downhill approaches, respectively. These results occurred for TTIs ranging from 2.31 to 5.33 s, as shown in Table 1. Significant differences in PRTs were observed for the uphill ($M = 0.78$ s) and downhill ($M = 0.68$ s) conditions ($F\ (1,972) = 65.4$, $P < 0.0001$) for the 72.4 km/h (45 mile/h) instructed speed, and also for the uphill ($M = 0.77$ s) and downhill ($M = 0.71$ s) conditions ($F\ (1,1045) = 30.1$, $P < 0.0001$) for the 88.5 km/h (55 mile/h) instructed speed.

Female drivers appeared to have slightly longer PRTs when compared to male drivers for both the 72.4 km/h (45 mile/h) and 88.5 km/h (55 mile/h) instructed speeds, as shown in Fig. 4. For the 72.4 km/h (45 mile/h) instructed speed, the

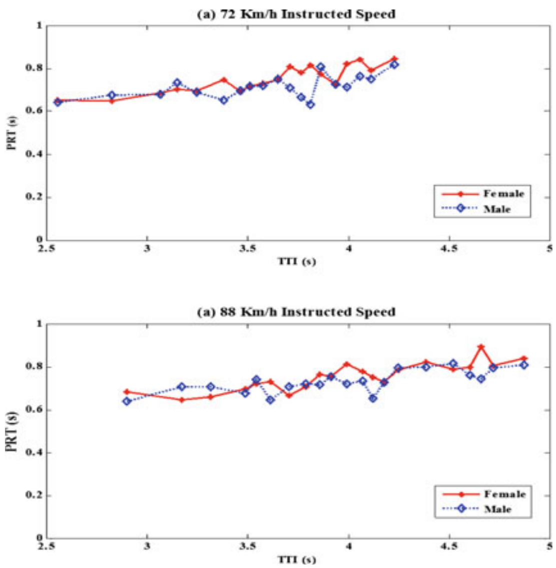
Table 1 Descriptive statistical results of PRT for grade, age, gender and platoon

Variable		N	PRT (s)						
			Min	Max	Mean	15 %	50 %	85 %	Std. Div.
(a) *72 km/h instructed speeds*									
Grade	Uphill	532	0.22	1.67	0.78	0.62	0.74	0.96	0.18
	Downhill	440	0.23	1.27	0.68	0.53	0.67	0.83	0.16
Gender	Female	462	0.23	1.42	0.75	0.58	0.73	0.93	0.18
	Male	510	0.22	1.67	0.71	0.57	0.68	0.88	0.17
Age	Older	317	0.42	1.42	0.79	0.63	0.78	0.93	0.15
	Middle	345	0.37	1.31	0.71	0.57	0.67	0.88	0.17
	Younger	310	0.22	1.67	0.70	0.53	0.67	0.88	0.20
Platoon	Following	309	0.22	1.47	0.74	0.58	0.72	0.92	0.18
	Leading	325	0.23	1.42	0.73	0.57	0.71	0.92	0.18
	Single	338	0.37	1.67	0.73	0.57	0.72	0.91	0.17
Overall		972	0.22	1.52	0.73	0.57	0.72	0.92	0.18
(b) *88 km/h instructed speeds*									
Grade	Uphill	544	0.22	1.53	0.77	0.62	0.73	0.93	0.18
	Downhill	501	0.18	1.43	0.71	0.53	0.68	0.87	0.18
Gender	Female	489	0.22	1.53	0.76	0.61	0.73	0.92	0.18
	Male	556	0.18	1.48	0.73	0.57	0.68	0.92	0.18
Age	Older	360	0.51	1.48	0.81	0.66	0.78	0.97	0.16
	Middle	366	0.22	1.53	0.72	0.57	0.68	0.88	0.18
	Younger	319	0.18	1.32	0.69	0.57	0.68	0.85	0.17
Platoon	Following	336	0.22	1.48	0.78	0.62	0.77	0.93	0.18
	Leading	351	0.18	1.27	0.71	0.57	0.68	0.88	0.17
	Single	358	0.32	1.53	0.74	0.58	0.72	0.90	0.18
Overall		1045	0.18	1.53	0.74	0.58	0.72	0.92	0.18

mean PRT was found to be 0.75 s for the female drivers and 0.71 s for the male drivers (as shown in Table 1; for the 88.5 km/h (55 mile/h) instructed speed, the mean PRT was found to be 0.76 and 0.73 s for the female and male drivers, respectively. Although the difference in PRT between male and female drivers seems small, the F-statistic generated from the GLM demonstrated that these differences are statistically significant for both approach speeds ($F(1,972) = 22.3$, $P < 0.0001$ for the 72.4 km/h instructed speed and $F(1,1045) = 10.9$, $P = 0.001$ for the 88.5 km/h instructed speed).

The data were utilized to analyze differences in driver PRTs associated with driver age. The PRT for drivers 60 years of age and older (M = 0.79 s for the 72.4 km/h instructed speed and 0.81 s for the 88.5 km/h instructed speed) were found to be significantly higher than those for the 40–59 age group (M = 0.71 s for the 72.4 km/h and 0.72 s for the 88.5 km/h) and those for the under 40-years-old age group (M = 0.70 s for the 72.4 km/h and 0.69 s for the 88.5 km/h), as shown in Table 1 and Fig. 5. The F-statistic generated from the GLM for the three age

Fig. 4 Effect of gender on perception–reaction times



groups demonstrated that significant differences, with P-values less than 0.0001, exist between the driver age groups for the 72.4 km/h (45 mile/h) instructed speed (F (2,972) = 33.9) and also for the 88.5 km/h (55 mile/h) instructed speed (F (2,1045) = 66.6).

A comparison of driver PRTs for the three different platooning scenarios (following, leading, or alone) was performed in order to characterize the effect of surrounding traffic on driver behavior. This behavior is important in designing traffic signal timing plans within the IntelliDrive initiative. The results showed that the PRT was not impacted by the platooning scenario in the case of the 72.4 km/h instructed speed (M = 0.74, 0.73, and 0.73 s for the leading, following, and alone scenarios, respectively), as shown in Table 1 and Fig. 6 (F (2,972) = 1.4, P = 0.26). Alternatively, in the case of the 88.5 km/h (55 mile/h) instructed speed, the mean PRT for the following, leading or single vehicle scenarios were 0.78, 0.71, and 0.74 s, respectively. Figure 6 shows that in the following platoon scenario, where the test vehicle was following another vehicle that proceeded through the intersection without slowing or stopping, the mean PRT was higher compared to the other two scenarios. A potential explanation for the higher PRT values could be that because the lead vehicle ran through the intersection, the subject driver was also inclined to proceed. This initial inclination to run increased the deliberation time for the drivers that eventually elected to stop. In the case of the leading platoon condition, a shorter mean PRT was observed (compared to the other two scenarios) as the driver may have been forced to decide faster in order to provide the following vehicle with sufficient braking time to avoid a rear-end collision. The F-statistic generated from the GLM demonstrated that significant differences, with P-values less than 0.0001 (F (2,1045) = 16.9), exist between the different scenarios.

Fig. 5 Effect of age on perception–reaction times

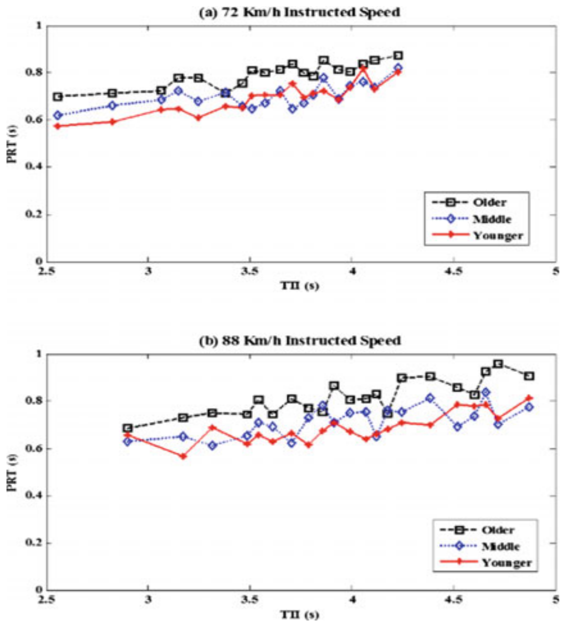
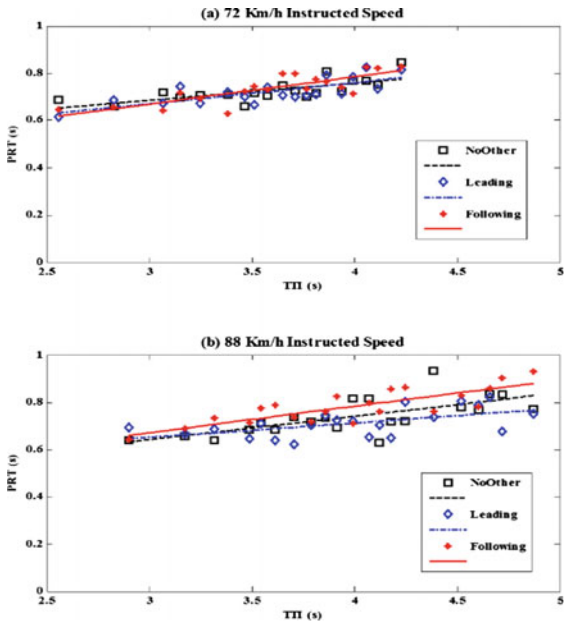


Fig. 6 Effect of surrounding traffic on perception–reaction times



The mean PRT for the 4-s (72.4 km/h instructed speed) and 4.5-s (88.5 km/h instructed speed) yellow interval was plotted against the TTI at the yellow interval change time, as shown in Fig. 7. The results show that the PRT tends to increase as

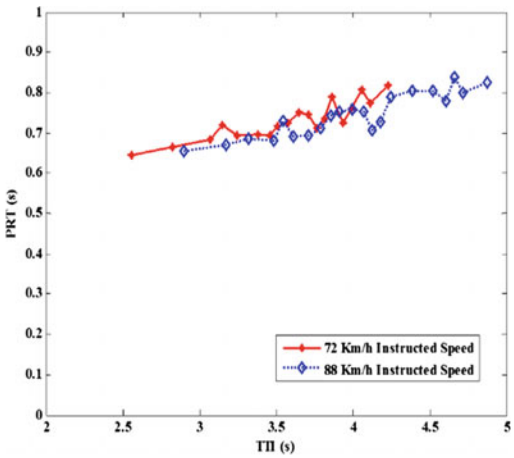


Fig. 7 Effect of vehicle speed on perception–reaction times

the TTI increases. For the range of TTIs from 4.69 to 1.93 s for the 72.4 km/h (45 mile/h) instructed speed, the PRT decreased from 1.53 to 0.22 s with a mean of 0.73 s, while for the range of TTIs from 5.33 to 2.31 s for the 88.5 km/h (55 mile/h) instructed speed, the PRT decreased from 1.53 to 0.18 s with a mean of 0.74 s. The figure demonstrates that the mean PRT curve for the 72.4 km/h (45 mile/h) instructed speed was higher than the 88.5 km/h (55 mile/h) instructed speed over the entire TTI range. This shift correlates to a lower PRT for those drivers traveling at higher speeds on the onset of the yellow interval. The results, however, indicate a similar PRT trend at different speeds. The results generated using the regression procedure (REG) of the SAS software showed that the approach speed had a significant effect, with P-values equal to 0.0006, on the average PRT.

4 Conclusions and Recommendations

The paper presented the results of a controlled field study that attempted to characterize driver PRTs at the onset of a yellow indication. Specifically, the study characterized the impact of driver gender, driver age, roadway grade, mean approach speed, platooning scenario (leading, following, or alone), and TTI on driver PRT. This characterization is critical in the efficient and safe design of traffic signal clearance timings and can be used within the Connected Vehicle initiative to provide the capability for vehicles to identify threats and hazards on the roadway and communicate this information over wireless networks to other vehicles and to infrastructure. Consequently, within the Connected Vehicle initiative the vehicle can gather information on the driver, the subject vehicle, and surrounding traffic conditions in addition to receiving Signal Phasing and Timing (SPaT) data to compute and display customizable in-vehicle warnings to the driver.

Initially the PRT values obtained from the study were compared to other field study results. The findings of this study are consistent with earlier studies, producing 85th percentile PRT values that are consistent with previously reported results. These findings demonstrate that the driver behavior observed in the controlled field experiment appear to be consistent with naturalistic and non-obtrusive field-observed driver behavior. Subsequently, the study demonstrated that the driver PRT is higher for female and older drivers (60 + age group) compared to male drivers and younger drivers of both genders. The PRT is larger for vehicles traveling along an upgrade given that the driver is typically accelerating when the yellow indication is initiated. Driver PRTs are typically higher if they are following a vehicle that runs a yellow light. Furthermore, driver PRTs decrease when they are followed by another vehicle. Finally, driver PRTs increase as the TTI at the onset of the yellow interval increases.

Acknowledgments The authors acknowledge the support of Yu Gao, Sangjun Park, Aly Tawfik, Sashikanth Gurram, Stephanie Shupe, and Mola-taf Al-Aromah for running the tests, and to the Center for Technology Development and the Smart Road Operations Group at VTTI; and thank Dr. Abdel-Salam G. Abdel-Salam for providing valuable comments in the statistical analysis. This work was supported in part by grants from the Virginia Transportation Research Council (VTRC), the Mid-Atlantic Universities Transportation Center (MAUTC), and SAFETEA-LU funding.

References

1. Dimitropoulos, I., Kanellaidis, G.: Highway geometric design: the issue of driving behavior variability. In: Proceedings of the International Symposium on Highway Geometric Design Practices, vol. 41, pp. 1–7, Boston, MA (1995)
2. ITE.: Determining vehicle signal change and clearance intervals. In: Technical Council Task Force 4TF-1. Institute of Transportation Engineers, Washington, DC (1994)
3. Milazzo, J., et al.: The effect of dilemma zones on red light running enforcement tolerances. Transportation Research Board, 81st Annual Meeting, Washington, DC (2002)
4. Taoka, G.T.: Brake reaction times of unalerted drivers. ITE J. **59**(3), 19–21 (1989)
5. Green, M.: “How long does it take to stop?” Methodological analysis of driver perception-brake times. Transp. Hum. Factors **2**(3), 195–216 (2000)
6. Caird, J.K., et al.: The effect of amber light onset time on older and younger drivers’ perception response time (PRT) and intersection behavior. Transportation Research Board, 84th Annual Meeting, Washington, DC (2005)
7. Setti, J.R., Rakha, H., El-Shawarby, I.: Analysis of brake perception-reaction times on high-speed signalized intersection approaches. Transportation Research Board, 86th Annual Meeting, Washington, DC (2007)
8. Gates, T.J., et al.: Analysis of dilemma zone driver behavior at signalized intersections. Transp. Res. Rec. J. Transp. Res. Board, 29–39 (No. 2030, Transportation Research Board of the National Academies, Washington, DC) (2007)
9. Rakha, H., et al.: Vehicle dynamics model for predicting maximum truck acceleration levels. J. Transp. Eng. **127**(5), 418–425 (2001)

Part V
Maritime—Human Performance and Safety Assessment in the Maritime Domain

Evacuations of Passenger Ships in Inclined Positions—Influence of Uphill Walking and External Stressors on Decision-Making for Digital Escape Route Signage

Sonja Th. Kwee-Meier, Alexander Mertens and Christopher M. Schlick

Abstract Numbers of ship passengers are increasing and modern cruise ships fit up to 8000 passengers. Nevertheless, technical failures and hazards on passenger ships can never be completely excluded, often requiring a fast and efficient evacuation of the ship in an inclined position. This paper aims at an integrated analysis how physical, mental and emotional stress affect decision-making for escape route signage in terms of decision times. 26 participants processed decision-making tasks with contradicting escape route signage while walking on a treadmill at 0°, 7°, and 14° with and without stressors, i.e. time limit and acoustic background noise. An inverse relationship between mental, emotional, and physical stress and decision times was found, that is, steeper uphill grades and higher stress levels were associated with shorter decision times.

Keywords Maritime safety · Evacuation · Decision-making Information processing · Escape route signage · Hazard · Disaster

S.Th.Kwee-Meier (✉) · A. Mertens · C.M. Schlick
Institute of Industrial Engineering and Ergonomics, RWTH Aachen University, Bergdriesch 27, 52062 Aachen, Germany
e-mail: s.meier@iaw.rwth-aachen.de

A. Mertens
e-mail: a.mertens@iaw.rwth-aachen.de

C.M. Schlick
e-mail: c.schlick@iaw.rwth-aachen.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_33

1 Introduction

Numbers of passengers on cruise ships are continuously increasing having evoked capacities of up to 8000 passengers on a single ship. Nonetheless, technical failures and hazards on passenger ships can never be fully ruled out. Inclined positions imply high physical needs for passengers on their escape route. Such hazards have already resulted in fatalities, such as on the Costa Concordia in 2012.

The situation in an evacuation is emotionally and physically very stressful due to implied conditions of an emergency with a possibly steep inclined ship position. Moreover, passengers need to process information, such as escape route signage, and decide for correct directions that might have changed from the presently deployed static escape route signage according to the specific circumstances, e.g. fire. Influences of environmental conditions, for instance, corridor width [1], knowledge of previously valid escape routes [2, 3], affiliative and social factors [4, 5] compete with signage information. In a prior study, we found that trust and decision-making for digital escape route signage can be positively affected by integrating temporal update information [6]. However, the results also revealed tendencies for rash decisions with increasing mental and emotional stress, for instance, due to time limits, which was already found to act as stressors on decision-making [7] and to induce anxiety [8].

Based on the Processing Efficiency Theory (PET) [9], Eysenck et al. [10] provided evidence in the Attentional Control Theory (ACT) that anxiety decreases information processing allowing different conclusions on decision times in anxiety states. Subsequently, more time is needed to process information. Anxiety is often related with urgency, which might cause the perceived need to make hasty decisions. Keinan et al. [11] found stress to enhance the tendency to decide before all decision alternatives have been scanned. Porcelli and Delgado's [12] results indicate that stress exacerbates the behavioral bias and enforces automated reactions. Stankovic et al. [13] investigated decision-making in the context of aviation emergencies in an eye-tracking study. The fixation numbers and dwell times decreased under stress. However, there was no effect of stress on the decision times. In conclusion, evacuation conditions with their implied stress are likely to reduce information processing efficiency prolonging processing times on the one hand. On the other hand, haste due to the emergency situation potentially counteracts against longer decision times. Thus, the influence of stress on decision times for escape routes remains uncertain with present literature to the authors' knowledge. Moreover, physical demands increase uncertainty of decision times under evacuation conditions on ships in inclined position. However, convincingly communicating escape route directions in times shorter than the decision times is crucial in evacuations.

Consequently, our research question that we focus on in this paper is: How do physical demands of escape routes on ships in inclined position, i.e. uphill grades, and mental and emotional stress influence decision-making in terms of decision times?

Therefore, we conducted an experimental study to investigate the effect of mental, emotional, and physical stress similar to an evacuation process on a passenger ship in inclined position. 26 participants walked on a treadmill at 0°, 7°, and 14° with and without applied mental and emotional stressors, i.e. time limit and acoustic background noise. The task was to decide for escape route directions according to contradicting signage. Perceived mental effort and state anxiety were evaluated. Decisions were recorded and examined with focus on how decision-making in terms of decision times is affected by the simulated evacuation conditions.

2 Methods

2.1 Participants

The 26 participants (13 m, 13f) had a mean age of 24.31 years ($SD = 2.57$). Ethical approval for the study was obtained by the ethics committee of the University Hospital of the RWTH Aachen University (EK 280/15). Exclusion criteria were pregnancy, cardiac pacemakers, severe illnesses, especially coronary heart diseases, and mobility, visual and hearing impairments. No participant violated the minimum boundary for visual acuity of ≥.8 or suffered from a red-green color-deficiency.

2.2 Experimental Conditions

Participants processed the same decision-making tasks in four conditions on a treadmill in the experimental setup depicted in Fig. 1. Four specifications of escape route signs were tested in randomized pairs against each other resulting in 36 decisions per condition with pauses of 2, 3, or 4 s in between. The minimum content was the standardized escape route sign according to DIN EN ISO

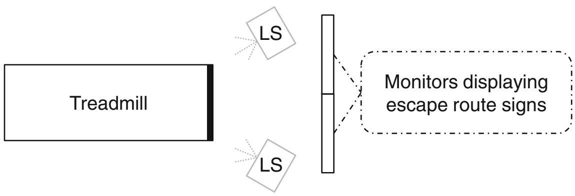


Fig. 1 Experimental setup with treadmill for level and uphill walking, loudspeakers (LS) pointing at the participant for acoustic stressors, and two monitors displaying contradicting pairs of escape route signs as decision-making task

Table 1 The four experimental conditions varying in grade and applied stressors

	Condition A	Condition B	Condition C	Condition D
Grade (°)	0	0	7	14
Time limit of for decisions (ms)	–	2000	2000	2000
Acoustic stressors (dB(A))	–	60–70	60–70	60–70

7010:2012 [14], optionally accompanied by either temporal update information, a flashing frame, or both elements.

The four experimental conditions varied in grade and applied stressors (see Table 1). Condition A was defined as level walking without stressors. In conditions B, C and D, mental and emotional strain was applied by acoustic stressors, i.e. continuous mumbling at 60–70 dB(A) played in a loop, and a time limit of 2000 ms for decisions with a maximum sign display of 1500 ms. Additionally, physical demands were increased in conditions C and D by uphill walking at 7° and 14° instead of level walking in conditions A and B. The walking speed was 2.7 km/h over all conditions. The condition order was permuted.

2.3 Experimental Protocol

Participants were comprehensively informed about the study, especially about the physical demands, signed informed consent, and answered questionnaires on demographic data, physical mobility, exercising, personality traits, and attitudes, for instance, technical affinity. Tests for sight and reaction times were conducted. Each participant walked on the treadmill for 5 min for familiarization to inhibit potential between-subject effects. The speed was increased from 0 to 1 km/h in the first minute, up to 2 km/h in the second minute, up to 2.7 km/h in the third minute, and kept constant during the last 2 min. Participants were secured on the treadmill with a safety harness in the training and the experimental conditions. There were breaks of 10 min between the training and in between the experimental conditions. At the beginning of each experimental condition, a standardized instruction was read to the participants. It covered a short description about the decision-making task under the simulated emergency conditions and how to decide for the preferred direction, i.e. left or right, according to the two contradicting escape route signs on the monitors by Wii remotes. They were informed whether they had to expect a time limit for the display of the signage and encouraged to make their decisions as fast as possible in every condition and latest when the displayed signs fade out in the limited time conditions B–D. Each condition started with two example escape route signs to demonstrate the principle and the potential time limit. Participants answered several questionnaires between the conditions and rated mental effort (RSME) [15], and state anxiety (MRF) [16, 17].

2.4 Apparatus

The treadmill was an h/p cosmos mercury treadmill with a walking surface of 1.5 by 0.5 m. The visual stimuli, i.e. the escape route signs, were displayed on two 22″ TFT monitors at 1.8 m height, adjusted upwards with treadmill grades in conditions C and D. The acoustic stressors for the background noise were played via two loudspeakers. The decisions were entered by two Wii remotes for right and left hand using the button on the backside for index finger inputs (B button).

2.5 Statistical Analysis

Statistical Analysis of Mental Effort and State Anxiety. The alpha level was set to $\alpha = .05$; tests were one-tailed as state anxiety and mental effort were assumed to increase with applied stressors and steeper grades. Repeated measures ANOVAs were conducted with the independent variable condition as within-subject factor. The assumptions for analysis of variance were not strictly met. Hence, Greenhouse-Geisser corrected measures of ANOVA were used for interpretation. The pairwise differences were analyzed by dependent-samples t-tests and effect sizes were evaluated by Cohen’s d [18] adjusted for dependent samples by Dunlap et al. [19] integrating the correlation in Formula (1) with n for the cases. Effects sizes were evaluated by Cohen’s criteria for $|d|$ with .2 = small, .5 = medium, and .8 = large effect size [18].

$$d = t[2(1 - r)/n]^{1/2}. \quad (1)$$

2.5.1 Decision Times

The alpha level was as well set to $\alpha = .05$, tests were two-tailed as literature is ambiguous about the influence of stress on decision times. Decision times are closely related to classical reaction times. Subsequently, the decision time data were positively skewed in the study, especially in condition A without time limit (see Table 4). The decision numbers were equal over the conditions and the maximum of missing values did not exceed 1.7 % [20]. Hence, the analyses of decision times were conducted using non-parametric measures and tests. Friedman’s ANOVA was used for examining the influence of the effect of the condition on the decision times. Pairwise test were conducted by Wilcoxon signed-rank tests. Effect sizes were calculated by Formula (2) with N for the number of observations over the compared conditions, not the number of cases [21]. They were evaluated by Cohen's criteria for $|r|$ with .1 = small, .3 = medium, and .5 = large effect size [18].

$$r = z/\sqrt{N}. \tag{2}$$

3 Results

3.1 Mental Effort

Mental effort for the decision-making tasks increased with applied stressors and steeper uphill grades (see Fig. 2). The influence of the factor condition on mental effort was significant with a large effect size, Greenhouse-Geisser corrected $F(2.17, 54.13) = 31.85, p < .001, \eta_p^2 = .56$. All pairwise comparisons between conditions were as well significant, indicated by dependent-samples t-tests revealing small to large effect sizes (see Table 2).

3.2 State Anxiety

State anxiety was evaluated by the mental readiness form with its three dimensions (MRF) [16, 17], indicating increasing levels of state anxiety with applied stressors and steeper uphill grades (see Fig. 3). The factor condition had a significant influence on state anxiety, Greenhouse-Geisser corrected $F(1.81, 45.28) = 15.50$, $p < .001, \eta_p^2 = .49$.

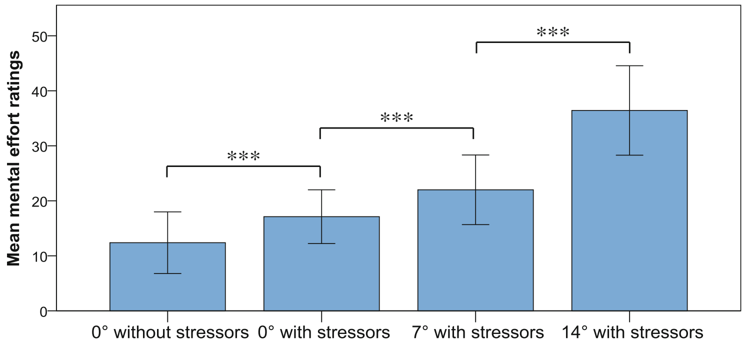


Fig. 2 Increasing ratings of mental effort (RSME) over the conditions varying in grades and applied stressors, means with 95 %, *p < .05, ** p < .01, *** p < .001, one-tailed, for effect sizes see Table 2

Table 2 Mental effort comparisons between conditions A (0° without stressors), B (0° with stressors), C (7° with stressors), and D (14° with stressors)

		Condition B–A	Condition C–B	Condition D–C
RSME	*t*(*df*)	2.1*** (25)	2.46*** (25)	5.41*** (25)
	d	.36	.33	.77

*p < .05, ** p < .01, *** p < .001, one-tailed

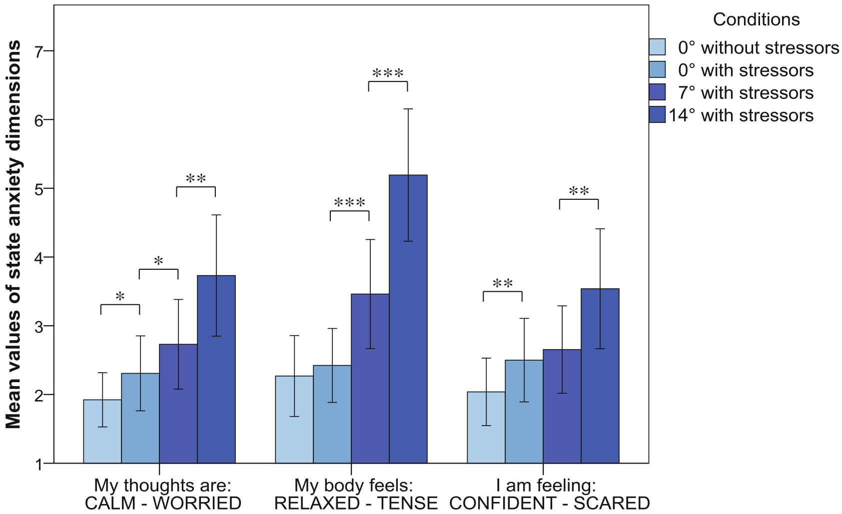


Fig. 3 Increasing state anxiety values according to the MRF dimensions over the four conditions, means with 95 %-CI, *p < .05, ** p < .01, ***p < .001, one-tailed, for effect sizes see Table 3

Dependent-samples t-tests showed that the second dimension of the MRF was only sensitive to changes in the grade (see Fig. 1) indicating that it might have been mainly influenced by physical demands in the study. Therefore, a new dependent variable MRFpsych is suggested, which is calculated as mean of the first and third dimension of the MRF to depict exclusively emotional strain, showing significant small to medium effects for each pairwise comparison of conditions (see Table 3).

3.3 Decision Times

Decision times were significantly different over the conditions, $\chi^2(3) = 178.81$, $p < .001$. Table 4 depicts the decreasing mean and median decision times with applied stressors and steeper uphill grades. However, valid decisions slightly decreased over the conditions A–D (see Table 4). The distribution of decision times in condition A without stressors, including time limit, is positively skewed and

Table 3 State anxiety comparisons between conditions A (0° without stressors), B (0° with stressors), C (7° with stressors), and D (14° with stressors)

		Condition B–A	Condition C–B	Condition D–C
MRF1: calm	*t*(*df*)	2.30* (25)	2.19* (25)	3.291** (25)
Worried	*d*	.30	.29	.50
MRF2: relaxed	*t*(*df*)	.750 (25)	3.51*** (25)	5.31*** (25)
Tense	*d*	–	.58	.77
MRF3: confident	*t*(*df*)	2.90** (25)	.78 (25)	3.11** (25)
Scared	*d*	.32	–	.44
MRFpsych	*t*(*df*)	2.90** (25)	1.83 (25)*	3.70** (25)
(MRF1 + 3)	*d*	.31	.20	.48

* $p < .05$, ** $p < .01$, *** $p < .001$, one-tailed

strongly leptokurtic implying that most decisions were made quickly, for instance, 82.2 % within 1500 ms. However, the distribution is heavy-tailed on the side of high decision times with a maximum decision time of 12.289 s although participants had been instructed to decide as fast as possible. Therefore, there is a strong variation in the decision times in condition A without stressors and a noticeable difference between the mean and the median (see Table 4). Thus, we utilized the median and non-parametric methods for further analyses (Fig. 4).

The decision times significantly decreased in each condition with small to medium effect sizes (see Fig. 3), resulting in medium effects between level walking without stressors and uphill walking at 7° and 14° with stressors, $r = -.30$ and $r = -.32$ respectively (see Table 5).

The effect size of the comparison of decision times between level walking with and without stressors, i.e. time limit and acoustic background noise at 60–70 dB (A), is small to medium. Participants made more than 80 % of the decision within the first 1500 ms, to which the presentation time of the escape route signs was limited in the other conditions, already in the level walking condition without stressors, including time limit. The percentage of decisions within the first 1500 ms

Table 4 Descriptive statistics on the decision time distributions in conditions A (0° without stressors), B (0° with stressors), C (7° with stressors), and D (14° with stressors)

	Condition A	Condition B	Condition C	Condition D
Decisions	936	928	926	921
Missing decisions	0	8	10	15
Mean	1142.6	845.5	824.7	789.5
Median	867.5	820.0	792.0	770.0
Within 1500 ms (%)	82.2	97.3	96.2	97.5
Within 2000 ms (%)	90.3	99.1	98.9	98.4
Skewness	5.5	0.7	1.2	1.0
Kurtosis	47.0	0.6	2.5	1.8

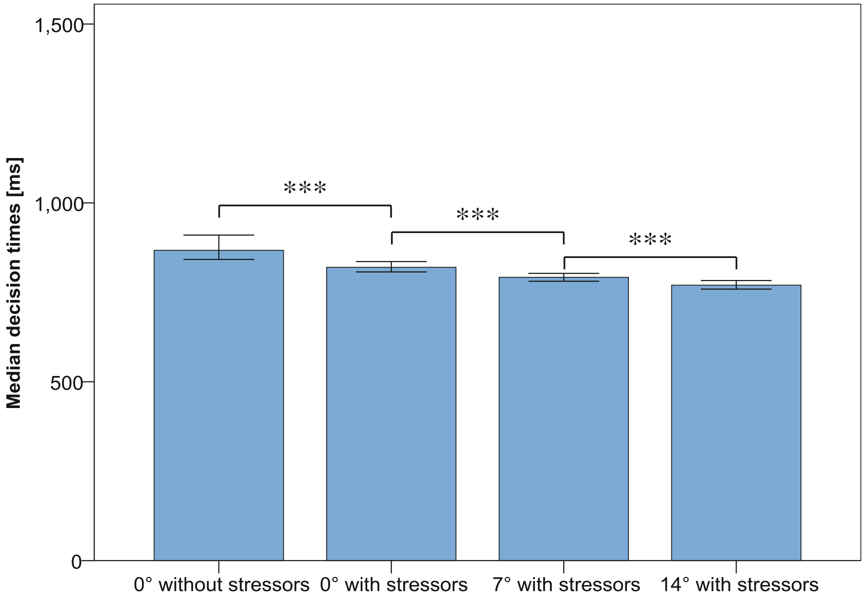


Fig. 4 Decreasing decision times for direction decisions according to escape route signs with increasing grades and applied stressors, medians with 95 %-CI, *$p < .05$, ** $p < .01$, ***$p < .001$. For effect sizes see Table 5 (The confidence intervals do not depict exactly 95 % but most likely a slightly larger interval as the median is a non-parametric measure and works with ranks)

Table 5 Comparisons of decision times between the condition A (0° without stressors), B (0° with stressors), C (7° with stressors), and D (14° with stressors)

		Cond. A	Cond. B	Cond. C	Cond. D
Condition A	*N*	–			
0° without stressors	*Z*	–			
	r	–			
Condition B	*N*	929			
0° with stressors	*Z*	−9.54***	–		
	r	−.22	–		
Condition C	*N*	926	919		
7° with stressors	*Z*	−11.60***	−3.34***	–	
	r	−.30	−.08	–	
Condition D	*N*	921	914	912	
14° with stressors	*Z*	−13.87***	−6.24***	−3.57***	–
	r	−.32	−.15	−.08	–

* $p < .05$, ** $p < .01$, *** $p < .001$, two-tailed

further increased to 97 % with applied stressors. In other words, with 1 % missing decisions, participants made their decisions only in 2 % of the cases after the announced soon end of the possible time to react (escape route signage fades out) although knowing that this option existed from the instruction at the beginning of each condition (Latest decision when escape route signs fade out). This largely unused time until exhaustion of the time limit suggests that decreased decision times are irrationally evoked by the increases in mental and emotional strain (see Sects. 3.1 and 3.2).

The effects with increasing uphill grades are small but noteworthy for two reasons. First, in condition B with stressors but only with level walking, the median decision time for escape route signs is 820 ms. In comparison, this is only 14 % longer than the time needed to read a traffic sign (720 ms) with 1 line with approximately 10–13 characters [22]. Second, these further decreases in decision times with steeper uphill grades show that increases in mental and emotional strain (see Sects. 3.1 and 3.2) not only arise because of higher physical demands but significantly affect decision times in the inverse direction despite higher physical demands and decreased information processing efficiency.

4 Discussion

The presented study investigated the influence of increased physical uphill walking demands of ships in inclined position referring to evacuation conditions, i.e. steep uphill grades and stress, on decision-making in terms of decision times. Conditions A and B were level walking and C and D uphill walking at 7° and 14°, respectively. In condition B, C, and D, stressors, i.e. time limit and acoustic background noise, were applied.

The stressors evoked significantly higher mental effort for decision-making and higher state anxiety, whereas decision times decreased with applied stressors from condition A–B. The decision times further significantly decreased with (steeper) uphill walking. Effects of grade were also found on mental effort for decision-making and state anxiety contradicting purely physiological demand reasoning for illogically decreasing decision times with increasing physical demands, but rather because of induced increasing mental and emotional strain by (steep) uphill grades.

Increasing state anxiety levels imply higher information processing times due to decreased efficiency according to the Attentional Control Theory (ACT) [10]. However, the inverse relationship was found. While mental effort and state anxiety increased in every condition from A to D with applied stressors and (steeper) uphill grades, reaction times decreased. This inverse relationship can be explained by the perceived stress, which was found to evoke several effects on decisions. For instance, it leads to decreased attention to the all possible choices [11, 13] and enhances the tendency to fall back to automated behavioral patterns under stress [12].

The effect of decreasing decision times under these controlled conditions was significant between all conditions and is expected to be even stronger under real evacuation conditions. However, the conditions with the stressors, i.e. time limit and background noise, and (steep) uphill grades already significantly affected mental effort and also state anxiety.

From a methodological point of view, the positively skewed and extremely leptokurtic decision times of condition A with the heavy tail for high decision times in comparison to the distributions of the other conditions are of interest for stress and evacuation related research. Participants were instructed to decide as fast as possible in every condition, including condition A, but the decision time results show that more than 2 s were taken for a decision in 10 % of the cases. In light of the distribution measures of the conditions B–D with stressors (see Table 4), it becomes clear that this instruction is not enough to evoke emergency similar decision-making behavior but external stressors and ideally challenging and realistic physical demands like (steep) uphill walking in conditions C and D are needed.

The participants' age was quite homogenous and young. Therefore, the generalizability of results is limited to this young age group in the first instance. Against the background of the relatively high ages on cruise ships and passenger ships in general with 72 % same or older than 50 years [23], we will conduct this experiment as well with people who are 60 years old and older.

5 Conclusion

The simulated evacuation conditions particularly regarding ships in inclined positions led to significantly shorter decision times for directions according to digital escape route signage. The applied stressors, i.e. time limit and acoustic background noise, resulted in higher mental and emotional strain, as did (steeper) uphill grades. In conclusion, an inverse relationship between mental, emotional, and physical stress and decision times was found. Against the background of the generally short decision times, the further decreasing decision times imply a high risk of rash decisions for directions in evacuations according to automated behavioral patterns. Well-designed and attention attracting signage [24] could help preventing people from rash decisions, that is, following cues in wrong directions into potentially dangerous routes, for instance, because of social influences by accompanying passengers [4, 5].

Acknowledgments The study was conducted within the project SIREVA funded by the German Federal Ministry of Education and Research (BMBF) in the context of the national program "Research for Civil Security" and the call "Maritime Security" (FKZ: 13N12954).

References

1. Vilar, E., Rebelo, F., Noriega, P., Duarte, E., Mayhorn, C.B.: Effects of competing environmental variables and signage on route-choices in simulated everyday and emergency wayfinding situations. Ergonomics **57**(4), 511–524 (2014)
2. Sime, J.D.: Affiliative behaviour during escape to building exits. J. Environ. Psychol. **3**(1), 21–41 (1983)
3. Johnson, C.W.: Lessons from the evacuation of the World Trade Center, Sept 11th 2001 for the future development of computer simulations. Cogn. Technol. Work **7**(4), 214–240 (2005)
4. Mawson, A.R.: Understanding mass panic and other collective responses to threat and disaster. Psychiatry: Interpers. Biol. process. **68**(2), 95–113 (2005)
5. Kinateder, M., Müller, M., Jost, M., Mühlberger, A., Pauli, P.: Social influence in a virtual tunnel fire—influence of conflicting information on evacuation behavior. Appl. Ergon. **45**(6), 1649–1659 (2014)
6. Meier, S.T., Bützler, J., Schlick, C.M.: The influence of information presented on digital escape route signage on decision-making under mentally and emotionally strenuous conditions. In: Proceedings 19th Triennial Congress of the IEA, pp. 1–8 (2015)
7. Maule, A.J., Mackie, P.: A componential investigation of the effects of deadlines on individual decision making. Contemp. Issues Decis. Mak. 449–461 (1990)
8. Maule, A.J., Hockey, G.R.J., Bdzola, L.: Effects of time-pressure on decision-making under uncertainty: changes in affective state and information processing strategy. Acta Psychol. **104** (3), 283–301 (2000)
9. Eysenck, M.W., Calvo, M.G.: Anxiety and performance: the processing efficiency theory. Cogn. Emot. **6**(6), 409–434 (1992)
10. Eysenck, M.W., Derakshan, N., Santos, R., Calvo, M.G.: Anxiety and cognitive performance: attentional control theory. Emotion **7**(2), 336 (2007)
11. Keinan, G., Friedland, N., Ben-Porath, Y.: Decision making under stress: scanning of alternatives under physical threat. Acta Psychol. **64**(3), 219–228 (1987)
12. Porcelli, A.J., Delgado, M.R.: Acute stress modulates risk taking in financial decision making. Psychol. Sci. **20**(3), 278–283 (2009)
13. Stankovic, A., Aitken, M.R.F., Clark, L.: An eye-tracking study of information sampling and decision-making under stress: implications for alarms in aviation emergencies. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 58(1), pp. 125–129 (2014)
14. DIN EN ISO 7010:2011: Graphical symbols—safety colours and safety signs—registered safety signs (ISO 7010:2011). German version EN ISO 7010:2012 (2012)
15. Zijlstra, F.R.H.: Efficiency in work behavior: a design approach for modern tools. Delft University Press, Delft (1993)
16. Krane, V.: The mental readiness form as a measure of competitive state anxiety. Sport Psychol. **8**, 189-189 (1994)
17. Murphy, S., Greenspan, M., Jowdy, D., Tammen, V.: Development of a brief rating instrument of competitive anxiety: comparison with the CSAI-2. Paper for the Association for the Advancement of Applied Sport Psychology, October 1989, Seattle, USA (1989)
18. Cohen, J.: Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates Inc., New Jersey (1988)
19. Dunlap, W.P., Cortina, J.M., Vaslow, J.B., Burke, M.J.: Meta-analysis of experiments with matched groups or repeated measures designs. Psychol. Methods **1**(2), 170–177 (1996)
20. Miller, J.: A warning about median reaction time. J. Exp. Psychol. Hum. Percept. Perform. **14** (3), 539 (1988)

21. Rosenthal, R.: Meta-analytic procedures for social research, 2nd edn. Sage, Newbury Park (1991)
22. Long, J., Hall, R.: Readability of information road signs: A preliminary study. Road Transp. Res. **9**(4), 11–20 (2000)
23. IMO, International Maritime Organization.: Guidelines for evacuation analysis for new and existing passenger ships. MSC Circular n. MSC/Circ. 1238, 30 Oct 2007
24. Nilsson, D.: Exit choice in fire emergencies-Influencing choice of exit with flashing lights. Lund University, Lund (2009)

FRAM in FSA—Introducing a Function-Based Approach to the Formal Safety Assessment Framework

Gesa Praetorius, Armando Graziano, Jens-Uwe Schröder-Hinrichs and Michael Baldauf

Abstract Formal Safety Assessment (FSA) is a structured methodology in maritime safety rule making processes. FSA takes organizational, technical and human-related factors into concern. While the method allows for the use of expert input during the identification of hazards and risk control options, the FSA guidelines give preference to assessment methods grounded in quantitative risk assessment. No specific guidance is given on how expert input should be obtained. This article therefore presents the findings of a pilot study with the objective to introduce the Functional Resonance Analysis Method (FRAM) as a method to enrich FSA studies through structured expert input. Two focus groups (n = 6) were conducted to compare hazards and risk control options identified in one scenario with the help of fault tree analysis and FRAM. The results of the study show that FRAM has the potential to enrich hazard identification as a complementary tool.

Keywords Risk assessment · Functional resonance analysis method (FRAM) · Hazard identification · Formal safety assessment

G. Praetorius (✉) · A. Graziano · J.-U. Schröder-Hinrichs · M. Baldauf
Maritime Risk and System Safety (MaRiSa) Research Group,
World Maritime University, P.O. Box 500, 201 24 Malmö, Sweden
e-mail: gp@wmu.se

A. Graziano
e-mail: agr@wmu.se

J.-U. Schröder-Hinrichs
e-mail: jus@wmu.se

M. Baldauf
e-mail: mbf@wmu.se

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_34

1 Introduction

Formal Safety Assessment (FSA) is a key feature in the International Maritime Organization (IMO) rule-making process. It is a systematic and structured framework to be used in studies to evaluate maritime risks with the objective to provide suitable options for risk control to be covered by IMO regulations [1]. The FSA framework consists of a five step approach including hazard identification, risk analysis, identification of risk control options, cost-benefit assessments and recommendations to decision makers [2]. FSA encourages a systematic assessment including organizational, technical and human-related factors. While FSA allows for the use of expert input during the identification of hazards and risk control options, many of the suggested methodologies for eliciting, formalizing and quantifying expert input are rooted in quantitative risk assessment.

Quantitative risk assessment is often of limited value when the system context of an operation is to be considered. The FSA guidelines, however, lack methods that may be able to capture this context in a more suitable way. This article therefore presents the findings of a pilot study with the objective to enrich the FSA framework through applying the Functional Resonance Analysis Method (FRAM) as a method to obtain structured expert input in the hazard identification and risk control option phase of the methodology. Two focus groups with master students (n = 6) in the World Maritime University (WMU) M.Sc. program Maritime Safety and Environmental Administration (MSEA) were conducted to test whether the application of different methods (fault tree analysis, FRAM) would result in the identification of different hazards and risk control options for a given scenario. The case chosen for the comparison was a collision avoidance scenario involving two vessels on collision course in the open sea in good visibility conditions. This rather simple scenario was chosen in order to limit the number of variables to be considered for the hazard identification. Even though the methodology is normally a five step approach, only the steps one (hazard identification) and three (identification of risk control options) were carried out in the groups. Both groups used two different approaches—a fault tree based approach and FRAM—in the context of FSA and discussed the results of their considerations after the exercise in the light of limitations and suitability of the different tools used.

2 Background

Before the results of the study are discussed, some theoretical background for the research presented in this article should be given.

2.1 Formal Safety Assessment (FSA)

Originally introduced in the response to the Alpha Piper disaster in 1988, FSA is currently being used in the IMO to support the rule-making process [3] with the objective to propose proactive, systematic, transparent and cost-effective maritime legislation [4]. First FSA interim guidelines were published in 1997 and IMO member states were invited to submit FSA studies for evaluation and further discussion. Over the years, many FSA studies have been submitted to IMO meetings on a regular basis and have addressed vital safety concerns on different ship types over the years. Very comprehensive FSA studies were, e.g., the FSA studies resulting from the EU financed project SAFEDOR reviewing general issues of concern on specific ship types, such as LNG Carriers (MSC 83/INF.3), container vessels (MSC 83/INF.8), crude oil tankers (MEPC 58/INF.2) etc. These are just a few examples to highlight that there is a lively debate about the application of FSA in the work of IMO.

The FSA guidelines have been revised and extended several times on the basis of the evaluations of the various FSA studies submitted to IMO. The Human Element, as human factors are referred to in the IMO, was recognized fairly soon. In 2006, it was also noted that for many issues to be assessed not enough data may always be available. Reliance on historical data may also prevent that FSA studies can be proactive. This is why an option for the inclusion of expert judgment was accepted in cases where no or insufficient data may be available for a thorough analysis.

However, the FSA guidelines remain deeply rooted in probabilistic approaches to risk assessment. This is reflected throughout the 2013 version of the FSA guidelines. Paragraph 3.2.3, as an example, highlights that a "proactive approach is reached through the probabilistic modelling of failures and development of accident scenarios". It can be debated if such an objective is realistic. However, this is not a question for this paper. Appendix 3 of the guidelines contain recommendations for methods to be used during hazard identification. Most of them are quantitative risk assessment tools, such as Fault Tree Analysis, Event Tree Analysis, Failure Mode and Effect Analysis (FMEA) etc. The same applies to methods suggested for the consideration of human factors. The main tool to be used for that purpose is Human Reliability Analysis (HRA), as suggested in Appendix 1 of the 2013 FSA Guidelines. A core component of the considerations is the quantification of human error probabilities with techniques like the Technique for Human Error Rate Prediction (THERP) and Human Error Assessment and Reduction Technique (HEART).

At the same time, the FSA guidelines acknowledge that many of the safety aspects evaluated in a FSA study require an evaluation of the entire system to which that aspect belongs. The generic model mentioned in the FSA guidelines (IMO [1]; Fig. 3 on page 17) specifically shows a diagram in which the Technical/Engineering System is embedded in a Personnel Subsystem which is part of an Organizational/Management Infrastructure in a wider Environment. It is therefore not quite clear, why the guidelines do not include tools that are more suitable to include the organizational context of safety performance in safety assessments [5, 6].

There is one option in the guidelines that potentially open up for such considerations. Especially in relation to expert input one tool is suggested during the hazard identification. This tool is an influence diagram. Even though this tool is also embedded in probabilistic thinking during HRA, it is more flexible and less linear than fault trees and event trees.

On the basis of the discussion above, the idea was developed within the MaRiSa group at WMU to compare a linear and a systemic tool and evaluate the results in the context of hazard identification during FSA studies. The two tools to be compared are Fault Tree Analysis and the Functional Resonance Analysis Method (FRAM).

2.2 Functional Resonance Analysis Method (FRAM)

FRAM is a method to analyze and model complex sociotechnical systems, in which functions are distributed over human operators, organizations and technology. The method focuses on the concept of performance variability and ways in which systems manage and monitor potential and actual variability. FRAM is based on four basic principles; the principle of equivalence of successes and failures, principle of approximate adjustments, principle of emergence and the principle of functional resonance e.g. [7, 8]. The principle of equivalence of successes and failures expresses that the only difference in between these two is the judgment of the outcome. While an action is deemed a success if it has the desired outcome, the same action can be identified as a failure when negative and unforeseen consequences occur. How these consequences arise is accounted for by the principle of approximate adjustments.

Sociotechnical systems are complex systems acting in an uncertain and dynamic environment. Functions are distributed over people, technology and organization that adjust their performance to be able to meet the demands the system is facing in the current situation. As this adjustment is based on the availability of resources (e.g. time, manpower) it will always be approximate. Consequently, everyday performance is, and needs to be variable to help the system to successfully adapt its functioning to the current operational conditions. The last principle, the principle of functional resonance, highlights the potential of the variability in multiple functions to resonate, and therefore reinforce and even amplify itself, so that the outcome of a function might carry an unusually high amount of variability, which the system is not able to manage given the current condition. As a result, accidents might occur.

FRAM consists of four steps, which are used to model the system based on functions and to identify sources of performance variability as well as measures to manage, dampen or monitor it. In Step 1 all necessary system functions are defined. The aim is to afford a consistent description as a basis of the analysis. All functions are described in form of their six aspects (Input, Output, Time, Control, Precondition, Resources/Executing conditions). These aspects describe the basic characteristic of an activity and help to understand relations among functional units within a system.

The functions that are the focus of the analysis are called foreground functions. Functions that are required by the foreground functions, but which do not themselves contribute to the variability being investigated, are called background functions [9]. Background functions represent the context and while they do not vary during the time frame specific for the analysis, they shape the performance and affect how events progress [10].

Step 2 helps to identify the variability of the functions in the FRAM model. The functions performance can vary in various ways. While functions involving humans tend to vary a lot, technical functions usually show a stable performance over time. Organizational functions do not show the same extent of variability as human functions, but show a delayed effect on these. There are three types of variability that can be characterized in a function: endogenous (internal), exogenous (external), and upstream-downstream coupling variability. Most interesting for an event analysis is the upstream-downstream coupling variability as it can become the basis for functional resonance. Upstream functions are carried out before downstream functions in the instantiation of the model, which means that variability in the earlier will impact on the performance of the latter [9].

In Step 3 of the analysis, an instantiation is created to see how performance variability can propagate through the system. It can help to understand how performance variability within some functions can amplify or dampen the variability of other functions, as the instantiation provides a way of simulating the functions performance within a specific operational condition to identify vulnerabilities and strengths of the system at work. The final step, Step 4, is used to suggest ways in which performance variability can be monitored, managed or eliminated. However, the analyst should keep in mind that varying performance also is an indicator for the flexibility to adapt performance to specific conditions, which means that eliminating variability can make the system rather inflexible and brittle [11].

Within the research presented here, step 1 to 3 have been carried out to build a model for "Collision Avoidance", which is described in more detail in Sect. 3.1.

2.3 Fault Tree Analysis (FTA)

FTA is a standard tool in safety and reliability engineering. It has been widely used in a number of industries since the early 1960s and is covered by various industry standards and norms (e.g. IEC 61025).

FTA is a top down approach in risk assessment to identify what factors in what combination may cause a top level event. Factors on a lower level come together in "and" or "or" relationships in order to cause an event on a higher level. The main objective of this method is to calculate the overall probability of the top level event in the tree.

In order to allow for a comparison of a fault tree with FRAM, the fault tree method had been slightly changed. Instead of identifying how an accident is caused, a tree was developed showing how an accident can be avoided. By doing so, a success tree had been developed. This allows for a discussion how vulnerable certain parts of this tree are and how their function can be strengthened. This approach is closer in line with the FRAM approach and therefore is more suitable for direct comparison.

3 Methodology

The following section describes the methodological approach of this study. Firstly, two models for Collision Avoidance, one based on the Fault Tree Analysis (FTA) and another based on FRAM were developed. Secondly, a hazard identification exercise was conducted with the help of two focus groups to explore the two models' ability to elicit expert input and formalize it in a way that it can contribute to a FSA.

3.1 Modeling Collision Avoidance

Prior to the focus group exercise, two models for collision avoidance were developed. The scenario modeled is based on two vessels in the open sea on opposite courses. The scope of discussion was open as it included options that may lead to a collision as well as circumstances under which the collision can be avoided.

The FRAM for Collision Avoidance. The FRAM model was constructed based on a task analysis carried out as part of the EU financed research project CyClaDes, a project focusing on crew-centered design. The task analysis was adapted in WMU's MaRiSa group, by a team member with a navigational background. This ensured that the functions identified and described in the model reflected a sufficient level of granularity. It also ensured appropriate couplings between the functions covered by the model.

All in all, a total of 25 functions, 6 background (e.g. (to) man the vessel, (to) leave the port etc.) and 19 foreground (e.g. (to) plot position, (to) detect target, etc.) were identified in step 1. In step 2 and step 3 the functions' variability (internal, external and in the couplings) was identified based on an accident analysis using TRACEr-MAR [12] which was also part of the aforementioned project CyClaDes [13]. This was done to indicate what functions were the most likely to vary and may cause a spread of variability through the system. Based on this analysis, five of the functions considered crucial for the safe conduct of the vessel in a collision avoidance scenario were identified ((to) monitor AIS, (to) perform lookout, (to) detect target, (to) set ARPA-alarm and (to) communicate intentions). These functions were subsequently highlighted in the model for the purpose of the exercise to enhance the comprehension of the model.

The Success Tree—An Adapted FTA for Collision Avoidance. As a FRAM-model depicts the functional units of a system and the couplings among those, it was necessary to take into consideration that the Fault-tree would need to be adapted to be able to explain the source of failure, but also offer the ability to elicit preconditions for successful avoidance.

3.2 Data Collection

Participants. Six master students pursuing a M.Sc. Degree in Maritime Safety and Environmental Administration (MSEA) were recruited for this pilot study. The participants were chosen based on their previous work experience, all had worked onboard ocean-going vessel and served as navigation officers prior to joining the university. All participants were male and the age ranged from 28 to 49 years coming from six different countries outside Europe.

Focus Groups. Two focus group interviews were conducted to be able to gain insights in whether, and if so how, FRAM can contribute to the identification of hazards and potential risk control options.

A focus group interview is a moderated group interview in which several participants are asked a certain number of questions, which are then discussed (often) openly in the group. The group discussion features one or two moderators who guide and steer the informants during their discussions [14]. Focus groups are generally good for understanding the fundamental issues and perceptions and the attitudes, thoughts and feelings of the informants. They are also highly useful for brainstorming. However, the moderator has few opportunities to control the outcome of the interview, as it is difficult to steer interactions among the participants [15].

3.3 Hazard Identification Exercise

The focus groups for the hazard identification exercise took place simultaneously in meeting rooms at the university's premises. Each group consisted of three participants and was led by a research team of two, who were responsible for introducing the exercise, present both models and moderate the discussions, as well as to document the identified hazards and risk control options. In addition, audio recordings and pictures of notes on the whiteboard were made. All notes by the participants were collected as well.

The group exercise took approximately 100 min and both groups started by introducing the participants to the aim of the study, the overall structure of the focus group interview and explained the right to withdraw informed consent at any time.

It was explained that the study focuses on two phases of the FSA; "Hazard Identification" and "Risk Control Options". The participants were instructed that their task would be to identify hazards and their potential consequences followed by the task to identify risk control options. To counterbalance any effects of order in which the participants were exposed to the model, one group started the exercise by using FRAM, while the other group started their hazard identification based on the FTA-model.

The results of the exercise were captured on a whiteboard to enable discussions about identified hazards and their control options. When a consensus about hazards, consequences and risk control options had been reached, the three participants were exposed to the second model and they were asked whether they would like to make any changes to the documented results on the whiteboard in terms of additional hazards or risk control options.

In a final step of the exercise, the participants were asked to comment on the use of the two models for the task they had been assigned. The focus was on capturing the advantages and disadvantages of each of the two models with regards to the task.

3.4 Identification of Hazards and Risk Control Options

Table 1 shows the various hazards that the two groups identified based on the model they were presented first. Group 1 started the hazard identification based on the FRAM model, while group 2 was first exposed to the success tree.

The hazards and risk control options identified by group 1, represented a wide range of aspects stretching over technology, organization and the human operator. The three participants focused in general little on identifying error(s), and more on how various functions were interconnected due to the couplings. Hazards often addressed aspects of the overall system design ranging from the design of technical equipment, to issues of manning and training. Furthermore, in general the hazards identified related to variability in the output of a function and its likely effect if that variability would be carried further through the coupling to another function. One example was variability in the output of the function "to monitor conning" (list item 3), which might arise due to the changes in manning of vessels. Less crew also

Table 1 Overview of identified hazard

Group 1	Group 2
1. Too high traffic density when monitoring the AIS 2. Target data is not complete 3. Monitoring the conning display is varying due to less manning on the bridge 4. Wrong reading of data within the integrated bridge system 5. Information overload in the electronic chart display 6. Erroneous alarms/signals in the decision support 7. Wrong data input 8. Manoeuver without communicating to the other vessel 9. Communicate with the wrong vessel 10. Choose inappropriate action (course alteration too small, misunderstanding of port and starboard)	1. Malfunction of the radar system 2. Deck officer without the necessary knowledge 3. Fatigue or use of prohibited substances 4. Bad weather conditions 5. Inappropriate actions by other ships 6. Small fishing vessels

Table 2 Overview of identified risk control options

Group 1	Group 2
1. Lookout to support visual target detection	1. Maintenance
2. Use of the ECDIS	2. Redundancy Systems
3. Additional watchkeeper	3. Increase look out
4. Calibration of alarms in decision support (the log is kept without any calibration)	4. Good selection process and training on board
5. Correlate to other sensors	5. Continuous supervision from senior officers
6. Back-up ECDIS on the bridge	6. Appropriate manning level
7. Alarm switch	7. Compliance with the regulations about rest and work hours
8. Paper chart as back-up	8. Radar setting and capacities
9. Officer-of-the-Watch (OOW) is habituated 30 min prior to watch to compare (Change to SOPs)	9. Increase look out
10. Guidance by company to communicate (SOPs) at any time possible	10. Take early actions
11. To alter course remarkably to give a clear signal	11. Avoid close quarters situations
12. Check AIS	12. Radar settings
13. Broadcast intentions to vessel(s) nearby	13. Increasing look out
14. Education and regular assessment of the OOW	
15. Additional training for OOWs	
16. Better integration of technology	

means that the same tasks need to be conducted with less manning, affecting the workload of other crewmembers and the bridge-team. A suggested risk control option was to ensure that additional resources (lookout or additional watchkeeper) would be present.

A similar observation about differences in between the groups can be made based on Table 2. While there is an overall concurrence in the identified RCOs with regards to manning (additional watchkeeper and increase lookout), there are differences in addressing underlying technical and organizational aspects beyond the immediate causes. In this pilot study only the group that first worked with the FRAM model identified several measures that addressed the company's responsibilities to provide the right preconditions for a successful operation in terms of adapted Standard Operational Procedures (SOPs), as well as regular training and assessment of competencies.

However, some comments ("We are doing this from our heads" (participant 2, group 1)) suggest that it was not always easy for the participants to relate to the model they were presented with. As an expert in a domain, it is not always easy to follow the formality of the model rather than falling back on one's own knowledge from actual work experience.

4 Applicability of Methods Within the FSA

Table 3 presents an overview of the positive and negative aspects identified by the informants concerning the two models used in the exercise.

Table 3 Overview of the participants' feedback on the use of the models

FRAM		FTA	
Positive	Negative	Positive	Negative
• The six aspects decompose the task with additional relevant information and help the user to understand differences interdependencies and influences with other tasks • More detailed equipment use in the representation to identify RCOs • Looks at multiple-source errors/lacks • Inductive and deductive approach in one • Easier to identify RCOs • More options to help decision making (more concrete model) • Good for macro-approach as it is covering the whole system • Shows human multi-tasking • Shows couplings among functions	• Complicated approach that does not generate concrete requirements for system design (hard to translate into design) • Seems to be difficult to quantify • Requires more time/learning • Not intuitive when seen the first time • Can be confusing	• Simple/easy to apply to smaller applications/isolated problems • Very clear for given case (fault-tree develops only into one direction) • Easy to recognize urgent hazards • Easily quantifiable • Useful for micro-approach • Easier to collect and portray adjustments made through RCOs	• Not covering all possible hazards and scenarios • Inter-connections between qualification, human element and equipment are not properly reflected • Difficult to identify RCOs • Limited due to its deductive nature • It does not take into account additional aspects (e.g. Time, Resources, etc.)

In general, the participants were more positive towards the use of the FRAM model than towards the success tree. They felt that the FRAM provided a more detailed description of the onboard work and therefore was better able to capture the overall complexity of the scenario in comparison to the success tree.

Furthermore, by depicting the interdependencies among functions, the FRAM was also more supportive in the identification of RCOs and their likely effects as the participants could use the visualization to track the couplings among the various functions.

One of the participants also highlighted that he thought of the FRAM model as a macro-level for analysis as all functions, background and foreground, were visible making it possible to reason beyond looking for a single point failure. However, while the participants were overall positive towards the use of FRAM, they also showed concern for the lack of quantification of variability and its spread.

The FSA consists of multiple steps and requires a cost-benefit analysis to evaluate the effectiveness of a RCO. Therefore, using only a FRAM might cause difficulties when RCOs need to be compared in terms of cost and effect in later phases of the assessment. Furthermore, it was stressed that FRAM is rather time-consuming and not as straight forward to understand as models that are based on tree-structures, such as the success tree. Therefore, it is of great importance to consider the choice of method for hazard identification in relation to the subject under assessment as well as in relation to the further steps of the assessment framework. For issues involving mainly technical systems or isolated problems, an approach based on a FTA with a model depicting a clear and easily assessed structure, might be more suitable.

5 Discussion

This study has taken a first approach to consider enriching two of the steps of the FSA methodology. The results show that different types of models applied in the hazard identification and identification of RCOs also elicit different types of feedback. Overall, the participants were positive towards the use of less traditional methods, such as FRAM, to complement commonly used approaches.

Furthermore, among the aspects addressed and criticized by Psaraftis [16] with regards to the FSA, two have been addressed in this pilot study focusing on hazard identification and RCOs. Firstly, the overreliance on expert opinion or, differently, the lack of a rigorous elicitation method to quantify experts' judgment, emerged as one of the major deficiencies in recently submitted FSA studies. Secondly, it was observed that many of the hazards identified in step 1 of recent FSAs, were ultimately not addressed in the following steps, reducing their thoroughness. Consequently, this suggested the testing of a complementary approach which could enrich hazard identification and still facilitating experts.

The results of the study show that FRAM can be considered as a complement to traditional risk assessment approaches, such as FTA, but probably not as a standalone method suitable for the FSA. Both focus groups were positive towards the function-based model and highlighted its added value as being able to portray the complexity of everyday maritime operations as well as interdependencies between tasks, which normally go unnoticed.

Some of the participants highlighted how the six aspects embedded in the design and the multiple possible interdependencies between functions, manage to portray a level of detail (in terms of tasks) which results in a comprehensive picture very close to the reality of ship operations. From the discussion, the six characteristics were considered a significant advantage of FRAM compared to the FTA, since "*you are able to determine which one influence which. You can see it clearly*" (Participant 2, group 2). Furthermore, the hazards identified differed depending on which visualization had been presented first, with the FRAM showing an emphasis on risk control options related mostly to organizational aspects, such as manning, education and fit between (design of) technology, operator and work environment. The FTA mostly generated direct advice on how to avoid errors and discussions circled around how to cut or add to the branches of the fault tree. The participants further stated that the FRAM offered a macro perspective of the system at work, while also enabling the participants to discuss and reason about the effects of risk control options. However, the lack of quantification and the complexity of the visualization was stressed as a disadvantage and the participants felt that the FTA would be a better option for the identification of hazard and risk control operation for less complex scenarios or a micro perspective on an isolated problem.

6 Concluding Remarks

While the use of failure-based approaches has a long history in risk assessment, this research has tried to show that there is a potential to study the introduction of new approaches into traditional frameworks. Within the maritime domain the FSA is the foremost tool to support decisions regarding safety-related measures. While the focus has so far mostly been on creating scenarios where the estimated risk is as low reasonably possible, the results show that there is a general lack to take everyday complexity and the interdependence of work environment, organizational setup and operator performance into consideration. The study seems to support concerns that the scenario-based approach of FSA is too limited for a thorough hazard identification and RCO evaluation.

This study presents a first attempt of introducing a complex modelling method into the FSA framework, and while several disadvantages were highlighted, it has been very salient that complementary approaches to identify hazards and risk control options are needed. FRAM presents one potential candidate for such a complementary approach.

References

1. IMO: Revised guidelines for Formal Safety Assessment (FSA) for use in the IMO rule-making process. IMO document MSC-MEPC.2/Circ.12. International Maritime Organization (2013)
2. Soares, C.G., Teixeira, A.P.: Risk assessment in maritime transportation. Reliab. Eng. Syst. Saf. **74**, 299–309 (2001)
3. Sames, P.C.: Introduction to risk-based approaches in the maritime industry. In: Papanikolaou, A.D. (ed.) Risk-Based Ship Design. Methods, Tools and Application, pp. 1–15. Springer, Berlin (2009)
4. Kontovas, C.A., Psaraftis, H.N.: Formal safety assessment: a critical review. Mar. Technol **46**, 45–59 (2009)
5. Schröder-Hinrichs, J.-U.: Human and organizational factors in the maritime world—are we keeping up to speed? WMU J. Marittime Aff. **9**(1), 1–3 (2010)
6. Schröder-Hinrichs, J.-U., Baldauf, M., Ghirxi, K.T.: Accident investigation reporting deficiencies related to organizational factors in machinery space fires and explosions. Accid. Anal. Prev. **47**(10), 1187–1196 (2011)
7. The Four Basic Principles of the FRAM. http://functionalresonance.com/basic-principles.html
8. Herrera, I.A., Woltjer, R.: Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis. Reliab. Eng. Syst. Saf. **95**, 1269–1275 (2010)
9. Hollnagel, E.: FRAM: The Functional Resonance Analysis Method—Modelling Complex Socio-Technical Systems. Ashgate Publishing Company, Farnham (2012)
10. Hollnagel, E., Hounsgaard, J., Colligan, L.: FRAM—The Functional Resonance Analysis Method—A Handbook for the Practical Use of the Method. Centre for Quality, Middelfart (2014)
11. Woods, D.: Essential characteristics of resilience. In: Hollnagel, E., Woods, D., Leveson, N. (eds.) Resilience Engineering: Precepts and Concepts, pp. 21–34. Ashgate Publishing Group, Farnham (2006)
12. Schröder-Hinrichs, J.-U., Graziano, A., Kataria, A., Praetorius, G.: TRACEr-MAR—guidebook for the technique for the retrospective & predictive analysis of cognitive errors adapted to the MARitime domain. MaRiSa Occasional Paper no. 2. WMU (2016)
13. Kataria, A., Praetorius, G., Schröder-Hinrichs, J.U., Graziano, A., Teixeira, A.P., Soares, C.G., Golyshev, P., Kähler, N., Ventikos, N., Lykos, G.V., Sotiralis, P., Hüffmeier, J., Fälth, J., Ljungkvist, M.: Deliverable 1.1 results of accident and incident analysis (2014)
14. Kuniavsky, M.: Observing the User Experience. A Practitioner's Guide to User Research. Elsevier Science, Amsterdam (2003)
15. Bryman, A., Bell, E.: Business Research Methods. Oxford University Press, Oxford (2015)
16. Psaraftis, H.N.: Formal safety assessment: an updated review. J. Mar. Sci. Technol. **17**, 390–402 (2012)

Designing New Interfaces for Submarines: From Cognitive Work Analysis to Ecological Interface Design

Daniel Fay, Neville A. Stanton and Aaron Roberts

Abstract Current submarine control rooms show a high degree of technical evolution, although future additions may cognitively overload operators. Additional sensors, capabilities and technology may prove difficult to effectively use for even the most trained operators. To mitigate this, the Command Team Experimental Test-Bed project aims to assess current methods of work, and how they can be improved. Cognitive Work Analysis will be used to assess current interfaces, with results informing new Ecological Interface Designs. It is anticipated that these new interfaces will reduce operator workload. This paper details how completed analyses are directly informing interfaces, ensuring that they meet required needs.

Keywords Human factors · Ecological interface design · Cognitive work analysis

1 Introduction

The Command Team Experiment Test-Bed (ComTET) is a project with the objective of identifying and developing new ways of working for submarine command teams [1]. A need for new ways of working will be necessitated by increased data, improved technologies, and a drive for crew reductions; these changes stand to improve a submarines command room, although may lead to cognitive overload for operators. There are many factors that could mitigate these effects on future crew, one of which is the design of an operators' station interface. Given that operators use their stations to interact with the submarines data, sensors,

D. Fay (✉) · N.A. Stanton · A. Roberts
Transportation Research Group, Faculty of Engineering and the Environment,
University of Southampton, Burgess Road, Southampton SO16 7QF, UK
e-mail: D.T.Fay@southampton.ac.uk

N.A. Stanton
e-mail: N.Stanton@southampton.ac.uk

A. Roberts
e-mail: Apr1c13@southampton.ac.uk

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_35

technology, and crew, it is imperative that stations facilitate, rather than hinder work. The focus of this research was to perform an evidence-based assessment of current interfaces.

Submarine command teams use passive sonar sensors to gather data. In early generations of submarines, sonar was listened to using headphones, and a crank that altered which bearing; these were passed to Target Motion Analysis (TMA) operators, who would plot the signals location on a paper chart. Command rooms have evolved since in response to technology and innovation [2]. Initial improvements have created digital versions of paper tools for operators to use, seeking to improve operator processes. Submarines show a high degree of technological evolution [3].

Recently Submarine Combat system upgrade projects have moved towards using Consumer Off-the Shelf (COTS) systems, as opposed to Military Specification (MILSPEC) ones [4–6], as well as modular open systems [7]. These projects have clearly shown to be beneficial in a variety of aspects, however it is unclear as to whether the associated user interfaces departed from their predecessors enough to make full use of new capabilities. There is little literature with regards to the design of currently operational interfaces, although there is literature regarding creating new interfaces to remedy problems regarding their usability, cognitive requirements, or functionality [8–10]. These sources all indicate that whilst a submarines command room is more advanced than ever, there is a clear need to present the information in highly usable interfaces.

2 Background Theory

2.1 Stations to Be Redesigned

To demonstrate the effectiveness of interface redesigns, it was decided to choose stations whereby the gains would be most evident. Through a series of discussions held with submarine and human factors SMEs it was decided that the Sonar and TMA would be candidates for this, due to their complex nature as well as importance within the command team. This importance can be seen in the sum of communications between the two sets of stations during scenarios, an example of which is published by Stanton [3].

Sound Navigation and Ranging (Sonar) is a system for locating and ranging objects via means of sound propagation and listening. This is often achieved passively, by listening to ambient noise. Sound detected by sensors can be plotted for analysis by TMA. Sonar is especially important for submarines as it allows them to gain a picture of their surroundings, as underwater visuals cannot be used. By listening to sounds underwater other objects can be detected and classified. This will then allow the control room to determine which actions to take.

TMA is the process of analysing positional data from contacts derived from sensors to produce a location, and predicted movements. This is called a solution, consisting of the speed, course, heading, and bearing of a vessel. Solutions are fed into a tactical map view, which can inform movement, attacks, or other changes. Such analysis is important, as it converts data into knowledge about contacts, allowing actions that can ensure mission success and platform safety.

2.2 Cognitive Work Analysis (CWA)

CWA [11, 12] is a framework developed for modelling complex socio-technical systems (systems constructed of social (humans, or other actors) and technological agents), to understand layers of constraints and how work could proceed within a system given their presence. Sanderson [13] further defines CWA as an approach to analyse, evaluate and design complex systems that have a high degree of human-computer interaction. Throughout all stages, there is an emphasis on how work could be conducted, as opposed to how it is conducted [14]. CWA has been applied, for a number of purposes [15], including interface design and evaluation [12, 16].

There are five different phases to CWA, which can have differing names and do not have to all be completed for an effective analysis. These are, as defined by Vicente [12]: Work Domain Analysis (WDA), Control Task Analysis (ConTA), Strategies Analysis (StrA), Social Organisation and Cooperation Analysis (SOCA) and Worker Competencies Analysis (WCA). As StrA and WCA were not used they will not be described. They were not used as they not often used to design interfaces [15, 17]; WDA is used for individual interface design, with ConTA and SOCA being used for interfaces that will form part of a system.

WDA considers objects within a system, what work can be performed using them and for what purposes [12, 18]. This is achieved using an Abstraction Hierarchy (AH), a diagram that models the system using five different levels of abstraction, ranging from physical objects to the systems reason(s) for existing, linked using means-ends connections, see Table 1. These connections may be verified using the 'Why-What-How' triad for each item in each layer of the AH, allowing users to either go from top to bottom to discover how a function is achieved, or bottom to top to assess how why objects exist and what they do. Once a hierarchy is complete, it is possible for practitioners to understand what an object in the system does by examining at links that go upwards, and understand how this is done by following downwards links. Figure 1 presents two examples of items within an AH, linked to form a 'Why-What-How' triad, with how these connections should be interpreted displayed below.

CTA addresses context in relation to activity. It can be modelled using a Contextual Activity Template (CAT) [19] table. Context is provided as situations, which can either be temporal or spatial, or a combination of both. Such analysis helps to identify different types of constraints and whether they are hard (requiring

Table 1 The levels of an abstraction hierarchy, with definitions of each level

Level	Items on this level should:
Functional purpose	Capture all reasons for design or procurement of a system. Each function provided by a system should be an item
Values and priority measures	Be a measure of how a system achieves its functional purposes
Purpose related functions	Link abstract ideals at the hierarchy's top levels to figurative items present in the bottom two levels, by grouping affordances into actions whose completion would count towards evaluating a Value and Priority Measure
Object related processes	Show possible affordances provided by the physical objects below. These affordances should be generic, so as not to constrain subsequent levels
Physical objects	Consists of both internal and external physical objects of the system, including user interface objects

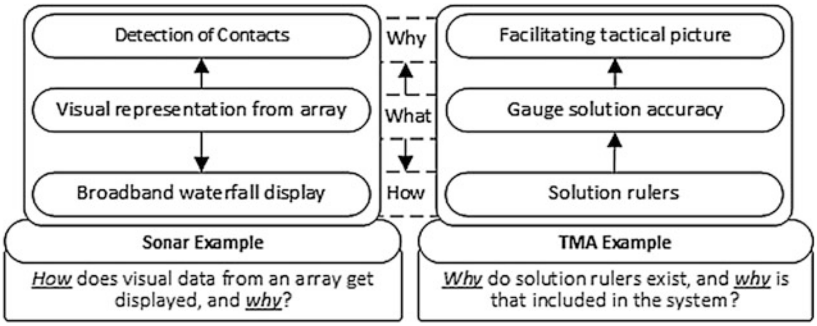


Fig. 1 Example items in an abstraction hierarchy for both sonar and TMA, with physical objects at the bottom. Note that *arrows* are only used for clarity, they are not included in the actual diagram

changes to environment or system) or soft (requiring changes in attitudes or behaviour) [14]. The vertical axis is typically populated either from the object-related processes or the purpose-related functions level from the AH. The horizontal axis is populated from possible scenarios. Finally, a graphical key is added to each row, showing which situations a can task be carried out and when it is [19]: An empty box indicates an action can never be carried out, a dashed outline indicates that an activity could, but is not, and a circle with solid lines through connecting scenarios denotes that an activity can and is typically performed. An example excerpt is show in Fig. 2a, showing available ORPs for a sonar operator on the vertical axis, and situations of work along the horizontal axis. As can be seen, the high level of training for operators, combined with a high level of technological sophistication, allows tasks to be completed across a variety of contexts. At this stage, actors are not included in the analysis, with the CTA's main purpose being to

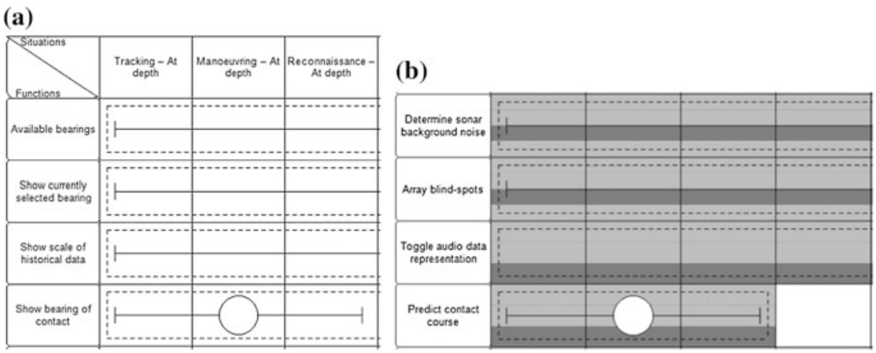


Fig. 2 **a** Example CAT (*left*). **b** Example SOCA CAT (*right*)

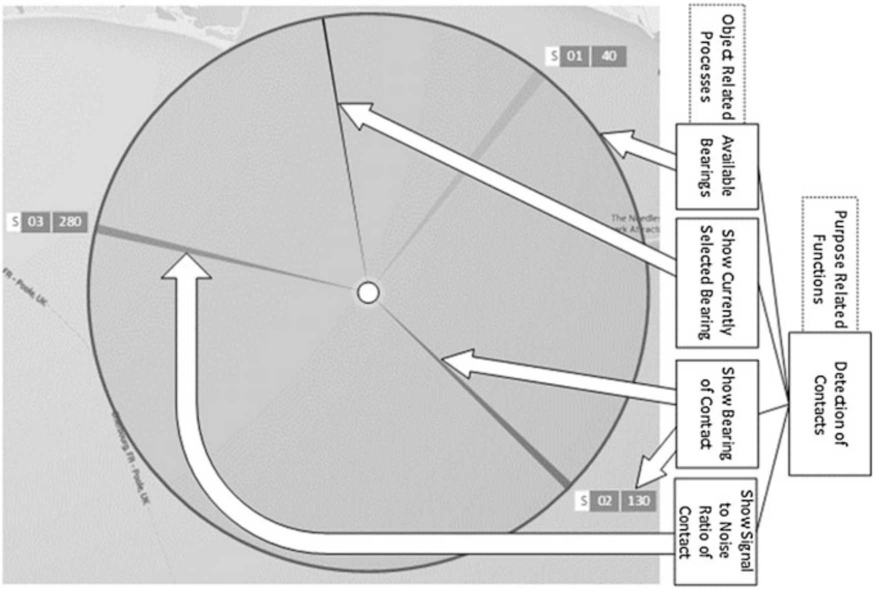


Fig. 3 A simple interface design, created from a sample sonar AH

show constraints for each activity with regards to whether or not it can be undertaken within a given scenario.

SOCA includes the actors of the system [20] as colours overlaid onto diagrams from other CWA stages. This is achieved by looking at constraints imposed by social and organisational structures and specific actor roles or definitions [15]. Actors are assigned a colour, which is then used to shade activities that they are capable of doing; it is important to note that shading denotes a capability to perform, rather than actual completion. With multiple actors, shading is split into discreet, equally proportioned blocks. For this analysis a SOCA CAT was performed, see Fig. 2b for an excerpt, allowing for analysis which tasks were performed, when they were performed, and who by. Figure 2b builds upon Fig. 2a by adding shading, representing sonar operators, to indicate that they will be the actors performing the actions laid out in the CAT.

2.3 Ecological Interface Design

Ecological Interface Design (EID) is a theoretical framework for designing human-machine interfaces of a complex nature [21], and is based upon the Skills, Rules, and Knowledge Taxonomy (SRK Taxonomy) [22] and the AH. Vicente [21] and Rasmussen [22] identify two primary objectives of EID: To ensure that required mental processing does not exceed task demands for a user, and to support all three levels of cognitive control described from in the SRK taxonomy. These inform user actions, and as such must be easy to interpret in routine, as well as unfamiliar, situations. For both types of situation, interpretation of an interface should not detract from a user's cognitive capabilities, that could better utilised executing their desired functionality. Enabling a user to understand the current environment at a glance enables them to use their full cognitive capabilities for assessing and choosing an appropriate course of action. Using CWA and EID to create interfaces has been extensively performed, as evidenced by McIlroy and Stanton [16], whose paper analysed usage of different CWA stages in relation to EID designs. In their paper they analyse the usage of different CWA stages used when producing an EID, alongside whether the SRK is mentioned, for over seventy entries. Their analysis indicated EID usage in a wide variety of fields, including aviation, military, and Information Technology. Additionally, they found a mix of CWA stages were used, with WDA being the most common, appearing in all but two entries.

Rasmussen's SRK Taxonomy posits that each taxa is a discrete distinction of human behaviour originating from fundamentally different representations of environmental constraints [12, 22]; however it is noted that they are not necessarily mutually exclusive. Each taxa is followed by 'Based Behaviour' when describing associated actions taken at that level, and can be abbreviated as: SBB, RBB, KBB. The taxa, and their associated environmental information perception contributories, are shown in Table 2. Both signs and symbols can be deemed as affordances, defined as what an environment offers an individual [23], and will contribute to how an environment or interface is manipulated. The SRK Taxonomy is important as it guides design principals, so as that all information in an interface is displayed in a manner that takes advantage of human perception and psychomotor abilities [24]. By capitalising on these innate abilities for most of the interface design, higher cognitive capability and capacity can be preserved to respond to events which require higher cognitive function to process and respond to. Vicente [21] defines a design principal for each taxon to guide creation of an EID as follows:

Table 2 The three taxis of Rasmussen SRK taxonomy, with associated perceptual contributories

Taxa*	Perception^
Skill—rote learned responses from an individual's perceptual motor system, in response to signals	Signals—environmental information, created via navigational as well as manipulatory means, sensed as time-spatial signals
Example—focusing* on a new sonar trace^ that has appeared on screen	
Rule—priorly learned sequential, or concurrent, actions to complete a desired or given end goal, encompassing manipulation of necessary	Signs—environmental information that indicates activation or modification of predefined actions or manipulations is possible. These indications refer to situations or behaviour by convention or prior experience, and do not represent concepts or functional properties of the environment
Example—setting a tracker* on the new sonar trace^	
Knowledge—actions generated in response to a new unfamiliar situation with no control rules available from previous encounters. These actions will be towards a goal based upon an analysis of the environment, in tandem with operators overall aims, taking both signs and symbols into account	Symbols—provide information pertaining to functional properties, and can be used for both reasoning as well as comprehension by virtue of providing this information
Example—deciding on a course of action* if the system cannot classify the trace^	

- SBB: An operator should be able to directly act upon the interface, with displayed objects being isomorphic to the environmental object it represents
- RBB: A consistent one to one relationship between work domain constraints and signs presented within the interface should be provided
- KBB: The work domain should be represented as an abstraction hierarchy, serving as an external mental model to support knowledge based problem solving

3 Method

The described method was carried out for both Sonar and TMA stations within ComTET. Construction of each diagram only proceeded after multiple consultations with submarine as well as human factors SMEs, visits to the Royal Navy's Talisman trainer, and over sixty hours of experimentation in ComTET. This was to ensure that all aspects of the system were fully understood, and as such could be effectively analysed.

3.1 Creating the CWAs

Abstraction Hierarchies were created in accordance with the requirements for each level laid out in Table 1, with the requirements being met as follows, with examples listed underneath each point:

- Functional Purpose: Given that all stations fulfil the same functional purpose in most cases, a set of purposes was devised and reused. Changes or additions were added to certain diagrams, so as not to constrain them and to provide a more accurate representation. A submarines main tenets (Remain safe, remain undetected, and complete the mission) were not included at this level, as they are better suited to being measures of performance, and they do not accurately provide a reason for the system existing in the control room.
  - A sonar station exists for Detection, Classification, Localisation, and Tracking (DCLT).
  - A TMA station exists to assist tactical picture generation.
- Value and Priority Measures: As with the functional purposes, all items were similar across diagrams, although greater variation was employed to represent different consoles functionality. Usage of keywords 'maximise' and 'minimise' created clear, quantifiable measures, including the submarines three main tenets.
  - A sonar station should maximise contact DCLT.
  - A TMA station should maximise solution accuracy, plus provide a representation of uncertainty.
- Purpose Related Functions: Functionality groups were identified from the object related processes. These functionality groups were aligned to processes.
  - A sonar station should provide functionality for DCLT of contacts.
  - A TMA station should provide functionality to predict contact actions.
- Object Related Processes: Each physical object was assessed for affordances, whether alone or in combination with another object. Inherent affordances, such as 'Push button' were not included so as not to detract from system related affordances.
  - A sonar stations bearing tapes display available bearings to an operator.
  - A TMA stations cuts allow an operator to identify detections of a contact.
- Physical Objects: All objects in the operators' screen were broken down and added, using a left to right, top to bottom approach. The breaking down of each component was performed by determining if the object was a group of objects, each with separate affordances.
  - A sonar stations waterfall can be broken down into the waterfall display, a bearing tape, a time tape, tracker indicators, etc.
  - A TMA stations solution rule can be broken down into a directional arrow, cut markers, a drag handle, etc.

To create the Contextual Activity Templates for both diagrams, the vertical axis, actions, was populated from the AHs Object-Related Processes, and the horizontal axis, situations, was populated using a common and an analysis specific set of situations. The common situations were derived by combining each scenario run by ComTET with an at depth, or periscope depth modifier. ComTET uses three scenarios, tracking, manoeuvring, and reconnaissance, as a low and high workload version, however as the processes are of comparable nature in each, a decision was made to ignore workload when populating the scenarios. Whilst the analysis specific situations would yield sufficient results by themselves, they would not have allowed for a comparison between different analyses; these comparisons will allow commonalities to be discovered, providing an exploitation vector for potential role mergers.

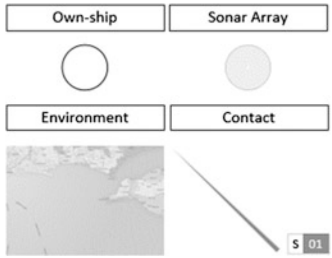
A SOCA CAT was constructed from the finished CAT diagram, with the inclusion of 'system' actor. This actors purpose was to represent all supporting actions that the system takes, in addition to logic checks that support users; for example, not assigning a tracker unless the Signal to Noise Ratio (SNR) is over a certain threshold on a sonar station. Whilst there was no explicit automation, logic built into the system creates safe-guards, which could be likened to system actions. An example of both a SOCA and SOCA CAT for sonar can be seen in Fig. 2.

Upon completion of the analyses, a thorough validation process was carried out to ensure accuracy and completeness. This process involved consulting with various SMEs, across differing relevant domains, and analysing ComTET scenarios. Results from this extensive validation led to several corrections being made; however, these corrections were not due to errors, rather ensuring that creational rules had been followed across all diagrams, and changes to represent a consensus on system functionality. With these 'finalised' CWAs, experimenters visited HMS Drake to perform an analysis of the Talisman trainer, in conjunction with currently serving submariners. Interviews were held and scenarios were monitored across several days, allowing experimenters to construct station CWAs that could be compared to those performed on Sonar and TMA stations of DW. Compiled analyses of this visit were then compared to their DW counterparts; there were differences in which role performed which function as well as differences in methods to completing tasks, but overall the comparison validated both the DW CWAs and ComTETs simulation capabilities. It should be noted that ComTETs capabilities have been explicitly validated before, this was an additional measure.

3.2 From CWA to EID

To inform an EID from both CWAs, the AH was the most used diagram. Both the SOCA and SOCA CAT were used to ensure that operators could carry out required functions in required contexts, in addition to identifying where potential automation could be added. Table 3 shows the steps taken, and example output, to translate each AH to an EID, which could then be refined to ensure optimal usability.

Table 3 A generalizable process for translating an AH to an EID, with a scaled down sonar example

Taxa (s)	Step	In example
SBB	List all objects within the expected operational environment, including those identified by Purpose Related Functions	• Own-ship • Environment • Contact • Sonar array
SBB	Create a Physical Object to represent each object, using skeuomorphism where appropriate	 *Not to scale*
SBB	Identify relevant properties that each object will have, using certain Object Related Processes nodes to ensure completeness	• Own-ship—course • Environment—land/sea boundaries • Contact—bearing, aural Strength • Sonar array—range, bearings covered
RBB	Identify constraints, if any, for each of these properties	• Courses and bearings can only be 0°–360° • Aural strength can vary greatly depending on the background noise • The range and bearings covered for a sonar array is contingent on which one is being used
SBB RBB KBB	Create Physical Objects to represent these properties and their constraints, or alter the Physical Object of the object it is attached to. Consider creating different states for when constraints have been broken, or properties are different to expected values	• Course and bearing constraints are already enforced by being displayed on a circle • To represent aural strength in relation to background noise, the 'Contact' Physical Objects' transparency will change • For this example, an array capable of surveying all 360° is employed, with the range being displayed by a thick border
–	For each Object Related Process:	
SBB RBB KBB	Ensure the outcome or information, detailed in the connected Object Related Process, can be determined using existing Physical Objects. If not, add Physical Objects to accomplish this	• The selected bearing is not currently shown, so a Physical Object similar to a contact detection is added to represent the current array location • The bearing of a contact is shown by its location, but the exact bearing is not; a text field will be added to the marker to display this

(continued)

Table 3 (continued)

Taxa (s)	Step	In example
SBB RBB	If it always occurs, try to make the associated Physical Objects as non-intrusive as possible	• As 'Detection of Contacts' always occurs the user should have to explicitly interact with them before the interface changes, to prevent issues when moving a cursor across the screen
RBB KBB	If it needs to be triggered, add a Physical Object, or affordance, to do so. Altering the states of related Physical Objects dependent on this trigger is strongly advised	• 'Detection of Contacts' does not need to be triggered, so this step does not apply
RBB KBB	If it can be remedial or diagnostic, list all causes of requirements. For each cause, consider Physical Objects to represent this	• 'Detection of Contacts' is not remedial of diagnostic, so this step does not apply
SBB RBB KBB	Arrange all Physical Objects on screen, ensuring the structure of the environment is represented	See Fig. 3

The target supported SRK taxa's are displayed as a guide, but all taxa's should be considered where appropriate

The process allowed for systematic capturing of all required environmental objects that would need to be displayed, what constraints there were, how to display them, and how to manage activation of functions to interact with the environment. This systematic elicitation provides support for behaviour at each level of the SRK Taxonomy, creating signs, signals, and symbols to facilitate appropriate levels of cognitive control; by integrating a model of the environment with an AH for interface design, it is possible to identify how best to do so. Additionally, it is believed that this process is generalizable to other scenarios where a practitioner is creating an EID from a CWA.

4 Results and Future Work

By following the process outlined it has been possible for experimenters to systematically create EID prototypes from CWAs, with a focus on AHs. It is believed that such a systematic elicitation of design direction can explicitly build upon a CWA to create designs, as opposed to informing them and starting a design from scratch. It is acknowledged that this method is not fool proof, and some subjective liberties will have to be taken, although not to the extent that a practitioner should feel without guidance from their analysis.

For future work, these interfaces will be extensively reviewed by human factors and submarine SMEs to ascertain their suitability; this process will be in-depth, to ensure a high standard of interface. Based on these reviews, it is anticipated that each interface will be iteratively improved before being testing their efficacy. Experimental testing is planned to follow.

5 Conclusion

This paper has presented how the ComTET team is moving towards testing new interfaces employing an EID approach, based on CWA analyses, to assess their impact in the submarine command room. A generalizable method for translating an AH from a CWA was presented, ensuring that analysis work informed each design in a direct manner, taking into consideration the environment of the system.

It is anticipated that experimental results will show a reduction in operator mental workload, coupled with improved task performance. Such a result would make a strong, evidence based case for future exploration of submarine command team improvements via interface redesigns. However, the authors recognise that there are many factors affecting team performance, interface design is just one piece of the puzzle. To that end, as the ComTET project progresses, many different ideas will be tested so as to address as many facets of such a complex challenge.

References

1. 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)
2. Dominguez, C., Long, W.G., Miller, T.E., Wiggins, S.L.: Design directions for support of submarine commanding officer decision making, In: Proceedings of 2006 Undersea HSI Symposium: Research, Acquisition and the Warrior, pp. 6–8 (2006)
3. Stanton, N.A.: Representing distributed cognition in complex systems: how a submarine returns to periscope depth. Ergonomics **57**, 403–418 (2014)
4. Scott, R.: Open for business: a new model for submarine sonar. Acquisition Community Connection (2006)
5. Hammond, P.: Open architecture vs open source in defence systems. Presented at the ACCU, Oxford, UK (2007)
6. Gosling, P.: Evolutionary Trends In UK Sonar, *RUSI Defence Systems*, Vol. 11(2) 31 October (2008)
7. Marcantonio, R.T., Sanford, E.G., Tillman, D.S., Levine, A.J.: Addressing the design challenges of open system architecture systems on U.S. navy ships—building out of the box (2007)
8. Dominguez, C., Long, W.G., Miller, T.E., Wiggins, S.L.: Design directions for support of submarine commanding officer decision making (2006)
9. Duryea, D.M., Lindstrom, C.E., Sayegh, R., Submarine imaging systems: developing improved capabilities and technologies, pp. 69430X–69438X (2008)

10. Hunter, A., Hazen, M., Randall, T.: VICTORIA class submarine human-in-the-loop experimentation plan. In: Presented at the 19th International Command & Control Research & Technology Symposium (ICCRTS), Alexandria, VA (2014)
11. Rasmussen, J., Pejtersen, A.M., Goodstein, L.P.: Cognitive Systems Engineering (1994)
12. Vicente, K.J.: Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. CRC Press, Boca Raton (1999)
13. Sanderson, P.: Cognitive work analysis and the analysis, design, and evaluation of human-computer interactive systems, In: Computer Human Interaction Conference, 1998. Proceedings. 1998 Australasian, pp. 220–227 (1998)
14. 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, 2nd edn. Ashgate, Farnham (2013)
15. Jenkins, D.P., Stanton, N.A., Salmon, P.M., Walker, G.H.: Cognitive Work Analysis: Coping With Complexity. Ashgate, Farnham (2009)
16. McIlroy, R.C., Stanton, N.A.: Ecological interface design two decades on: whatever happened to the SRK taxonomy? IEEE (2014)
17. Burns, C.M., Hajdukiewicz, J.: Ecological Interface Design. CRC Press, Boca Raton (2004)
18. Rasmussen, J.: The role of hierarchical knowledge representation in decisionmaking and system management, Systems, Man and Cybernetics, IEEE Transactions on, vol. SMC-15, pp. 234–243 (1985)
19. 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**, 371–394 (2006)
20. Stanton, N.A., McIlroy, R.C., Harvey, C., Blainey, S., Hickford, A., Preston, J.M., et al.: Following the cognitive work analysis train of thought: exploring the constraints of modal shift to rail transport. Ergonomics **56**, 522–540 (2012)
21. Vicente, K.J., Rasmussen, J.: Ecological interface design: theoretical foundations. IEEE Trans. Syst. Man Cybern. **22**, 589–606 (1992)
22. Rasmussen, J.: Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Trans. Syst. Man Cybern. 257–266 (1983)
23. Gibson, J.J.: The Theory of Affordances. Hilldale, USA (1977)
24. Vicente, K.: Ecological interface design: supporting operator adaptation, continuous learning, distributed, collaborative work. In: Proceedings of the Human Centered Processes Conference, pp. 93–97 (1999)

The Command Team Experimental Test-Bed Phase Two: Assessing Cognitive Load and Situation Awareness in a Submarine Control Room

Aaron Roberts, Neville A. Stanton and Daniel Fay

Abstract The Command Team Experimental Test-Bed (ComTET) is a body of work examining the functionality of submarine command teams with an emphasis upon future ways of working. 10 teams of 8 participants (80 participants in total) received extensive training at one of the operator stations in the submarine control room simulator (e.g. sonar or periscope operator). The teams then completed 3 different scenarios under higher and lower work demand conditions. The Work Load (WL) and Situation Awareness (SA) of the command team was assessed using a variety of standardized subjective rating scales (e.g. NASA TLX), physiological measures (e.g. ECG), in play cognitive capacity assessments (e.g. duel task paradigm) and SA assessments. The communication(s) between all team members were recorded allowing the use of the Event Analysis of Systemic Teamwork (EAST) method to examine performance. Preliminary results indicate that the WL of operators in a submarine command team varied as a result of scenario type and scenario demand. The initial results are discussed alongside future analysis plans.

Keywords Human factors · Submarine · Command team · Work load

1 Introduction

The submarine control room has evolved across many decades of operations and so represents a highly advanced system, but this does not mean that it cannot be improved [1]. Submarines of the future will be required to handle greater volumes

A. Roberts (✉) · N.A. Stanton · D. Fay
Transportation Research Group, Faculty of Engineering and the Environment, University of Southampton, Boldrewood Campus, Burgess Road, Southampton SO16 7QF, UK
e-mail: apr1c13@soton.ac.uk

N.A. Stanton
e-mail: N.Stanton@soton.ac.uk

D. Fay
e-mail: D.T.Fay@soton.ac.uk

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_36

of data (e.g. extra sensors), with potentially reduced crew sizes, a drive prevalent in many domains [2–4]. Previous research has identified the types of decisions that submarine command teams are required to make and it is clear that such decisions are informed by the integration of information from disparate sources [5, 6]. A key challenge for future submarine command teams is to effectively manage increasing volumes of data of greater complexity, whilst ensuring that Work-Load (WL) and Situation Awareness (SA) is maintained at a level which facilitates optimal performance [1, 3].

A submarine control room is an excellent example of a socio-technical system, involving complex interactions between technological agents and human operators [1, 7]. Previous work investigating WL and SA in submarine control rooms has typically focused upon such concepts as being held in the mind of the operators, often whilst performing particular operational sub-tasks such as track management [8, 9]. However, placing WL and SA solely in the mind of the individual operatives does not facilitate a complete understanding of the sociotechnical system as a whole but rather the cognitive capacities of the operator, which are accepted to be limited [10].

Distributed cognition and SA is defined as the requirement of multiple individuals to communicate and coordinate tasks effectively to complete common higher objectives, often with a reliance on technology [11, 12]. To understand the functionality of a submarine command team in its entirety it is important to understand the operations completed and relationships between all agents within the sociotechnical system, not just the minds of the humans acting as operators. The work of Stanton [1] provided valuable insight into one of the key operations routinely completed by operational submarines from a socio-technical perspective, returning to periscope depth from below 60 meters. However, such work did not objectively examine the SA and WL of the submariners. It is also important to consider a greater range of operations routinely completed by submarines.

The first phase of the Command Team Experimental Test-bed (ComTET) program was the building of a submarine control room simulator with sufficient fidelity to undertake such a program of research [2]. The aim of the current work (the second phase of ComTET) is to conduct a set of rigorous experiments with high statistical power in the submarine simulator. This will allow the examination of the WL and SA of submariners from a socio-technical perspective.

2 Method

2.1 Participants

10 teams of 8 individuals (80 participants in total) were recruited opportunistically using posters (e.g. placed on University of Southampton campus and submarine Museum) and by directly contacting local groups with a maritime or military

interest (e.g. University Military undergraduates, Royal National Lifeboat Institute). A total of total of 71 males and 9 females participated with an age range of 18–52 from a variety of backgrounds primarily including undergraduate students and graduate recruits for Ministry of Defence supported companies. The study protocol received ethical approval from the University of Southampton Research Ethics Committee (Protocol No: 10099) and MoDREC (Protocol No: 551/MODREC/14).

2.2 Equipment and Materials

The ComTET team designed and built a submarine control room simulator (see Fig. 1) that is based upon a currently operational RN submarine. A fuller description of the building process and the simulator capabilities is provided by Roberts et al., [2]. The operator roles present in the control room was advised by Subject Matter Experts (SME's) to be representative of an operational submarine. They include; two Sonar Operator stations (SOP), two Target Motion Analysis stations (TMA), a Sonar Controller station (SC), an Operations Officer station (OpsO), a Periscope station (PERI), a Ship Control station (SHC) and an Officer of the Watch station (OOW). The OOW role was played by a member of the ComTET team who was highly trained in the use of all DW stations and had received training from RN SME's in playing the role of OOW.

An unclassified set of scenarios (see Table 1) capturing the widest range of operations submarines routinely complete were selected by SMEs and programed in DW. The scenarios chosen were Returning to Periscope Depth (RTPD), completion of Inshore Operation (INSO) and Dived Tracking (DT) of a surface vessel. The average scenario length was 45 min.

A training program was developed by the ComTET team, allowing a team of novices with no maritime experience to become competent enough to complete scenarios in DW as a functional submarine command team. Initially over 12 h of video tutorials were designed, written and recorded. This initial package was then

Fig. 1 The ComTET submarine control room simulator

Table 1 Dependant variables synopsis with analysis plan

Measure	Purpose	Design	Analysis
Command team communication (s)	Examine the command team from a socio-technical perspective to elicit how the structure of the system changes depending on demand and scenario type	3 × 2 × 8 Mixed Scenario × Demand × Role	Event analysis of systemic teamwork (EAST—examines three types of networks (Social, Information and Task) derived from communications. Network metrics shall then be statistically analysed by conducting a mixed ANOVA
NASA TLX and Bedford WL scale and SA assessment	Examine the subjective WL and SA experienced by different submarine command team operators depending on demand and scenario type	3 × 2 × 4 Mixed Scenario × Demand × Role Roles reduced as sonar, TMA and command team members are grouped	The mental demand, physical demand, temporal demand, performance, effort frustration overall score, Bedford score and SA ratings of each operator shall be statistically examined by conducting a mixed ANOVA
Test(s) of visual and verbal WM	Examine if completing scenarios in a submarine simulator impacts upon the cognitive function of operators—optimisation or deterioration	2 × 4 Mixed Time administered × Role	Tests scores when administered upon entry to simulator and after completion of multiple scenarios shall be examined by conducting a mixed ANOVA
Dual Task Study	To examine the spare capacity each operator has during the routine completion of role and how such capacity changes depending on scenario type, scenario demand and point of scenario	3 × 2 × 4 Repeated Scenario × Demand × Time Time relates to scenarios spilt into quarterly timed segments	By subtracting single-task probe reaction time and completion rate from the corresponding dual-task probe reaction time the difference can be assumed to represent changes in workload levels. The score of each operator shall be statistically examined by conducting a repeated measures ANOVA

(continued)

Table 1 (continued)

Measure	Purpose	Design	Analysis
Physiological arousal (HR, blood pressure, ECG, Temperature)	The collection of Electrocardiogram data can provide valuable objective ‘at the time’ information concerning subtle changes in cognitive workload	3 × 2 Repeated Scenario × Demand	Heart Rate derived variables can be a sensitive measure of mental load. Typical HRV analysis consists of a series of measurements of successive R-R wave interval variations. Differences in the derived metrics will be compared across scenario type and demand by conducting a repeated measures ANOVA
Task performance behavioural measures	To examine how the behavioural indicators of sub-task performance (e.g. vessel detection times) varies depending upon scenario type and demand	3 × 2 Repeated Scenario × Demand	WL is often inferred from objective measures of performance, such as time or error. The performance metrics of all operators will be compared across scenario type and scenario demand by conducting a repeated measures ANOVA

reviewed by SME's from the RN who facilitated the condensing of the tutorials into a daylong training package in which all essential tasks were trained. The final training program consisted of 6 h long video tutorials concerning the basics of bearing, course, ranging, speed, communications and operator workstation specifics (e.g. a broadband sonar tutorial or Periscope tutorial). Participants also completed a training communications game, in which disparate pieces of information (word anagrams) had to be manipulated (solved) and integrated for higher goals (sentence solving). The solved sentences then linked to a higher global objective. During the game operators were required to use a military verbal protocol such as putting out 'calls' to operatives by stating who the 'message' is from and intended for.

A wide range of data were collected during the study to assess a variety of different constructs (e.g. SA vs. WL) from a number of different perspectives (e.g. objective vs. subjective). A comprehensive list of the data collected and the purpose of each measure is provided in Table 1. A number of cognitive tests and WL measures were administered to participants. The majority of tests were chosen from the Psychological Experiment Building Language (PEBL) test battery [13]. Electroencephalogram and respiration rate was monitored using Biopacs Bionomadix system. The BioNomadix system is a wireless, non-invasive, multi-channel physiological recording platform. Its design allows for nearly unlimited freedom of movement, enabling subjects to easily relax into their experimental procedure. A total of 4 wireless receivers were used allowing the collection of ECG and respiration rate from 4 operators (TMA and Sonar operators). The blood pressure, heart rate and temperature of all operators was measured across the day using two Omron M7 monitors and a Braun Thermoscan ear thermometer. To comply with ethical guidelines, fresh ear caps and electrodes were used for every recording.

2.3 Procedure

The experiment was run over a two-day period. Informed consent was attained from participants whilst a simulator induction was completed on the first (training) day. Participants were randomly assigned station roles and provided with a daylong familiarization/training program at their workstation. The tutorials were run on the second screen at the workstations, allowing operators to watch the tutorials whilst simultaneously practising skills at the workstation on the bottom screen. Each tutorial lasted approximately 45 min—participants watched 5 tutorials in total, with adequate breaks in between. At the end of the day participants took part in 4 practice scenarios with guidance and instruction provided by DW SMEs from the ComTET team. At the mid-point of the training day, after a lunch break, participants were administered a battery of cognitive tests which took approximately 60 min to complete, heart rate, blood pressure and temperature of participants was also recorded at this point. The battery of cognitive tests included baseline scoring for the three tasks chosen to be part of the dual task assessment alongside all of the tests of WM. All versions of the tests were electronic, running on the top monitor of the workstation. The presentation order of the tasks was counterbalanced across teams.

On the second day of testing participants were guided through a final practice scenario in the morning before the formal testing began. The four participants chosen (TMA and Sonar operators) to have ECG recorded were asked to step into private fitting booths in which posters with detailed instructions concerning the fitting of the electrodes and respiration belts were present.

The integrity of the physiological data was checked after fitting of equipment. Participants were then told that the first scenario would begin—all recording devices were started and a verbal time stamp was read aloud for synchronization purposes. Participants were also presented with a secondary cognitive task on the top screen of their workstation. Participants were instructed that anytime they were not required to be actively engaged with tasks relevant to submarine command team performance (e.g. no work had been allocated), they should try and complete as many trials on the cognitive task as possible, whilst maintaining accuracy.

Each scenario began with an OOW briefing outlining the mission objectives. The OOW led the direction of the scenario tactically, as would occur operationally. Once the command team had completed the mission objective(s) the end of the scenario was called. Immediately after the end of the all scenarios participants would be presented with electronic versions of the NASA TLX, Bedford scale and SA scale, the presentation order of these scales was counterbalanced across teams. After a short break for refreshments and debrief regarding the previous scenario participants were asked to sit back at their workstation and the second scenario would begin. From the third scenario onwards, immediately prior to the end of the scenario, two participants would have their blood pressure, heart rate and temperature recorded.

2.4 Results

The ComTET team is currently in the process of completing the pre-processing and analysis of data from all 10 teams that completed phase two of ComTET, the baseline testing. A selection of results are presented below providing an early indication regarding the direction of the work. Firstly, the NASA TLX and Bedford WL scores from 6 teams (48 participants) are presented from all 6 scenarios completed. Secondly, measures of HR and temperature recorded at baseline during training and during the completion of scenarios from the same 6 teams (48 Participants) are presented.

The temporal load ($F_{2,\ 62} = 3.22$, $p < 0.05$, $\acute{\eta}_p^2 = 0.10$), performance ($F_{2,\ 62} = 14.04$, $p < 0.05$, $\acute{\eta}_p^2 = 0.31$) and frustration ($F_{2,\ 62} = 3.04$, $p < 0.05$, $\acute{\eta}_p^2 = 0.09$) of the command team was statistically significantly affected by scenario type. Scenario demand statistically significantly affected the mental load ($F_{1,\ 31} = 46.24$, $p < 0.01$, $\acute{\eta}_p^2 = 0.60$), physical load ($F_{1,\ 31} = 39.67$, $p < 0.01$, $\acute{\eta}_p^2 = 0.56$), temporal

Table 2 Means and SDs of workload scores in high and low demand scenarios

Measure	Scenario	Low demand mean ± SD	High demand mean ± SD
NASA TLX mental	RTPD	7.03 ± 3.94	11.29 ± 6.66
	INSO	7.75 ± 4.72	10.51 ± 4.84
	DT	9.14 ± 5.34	10.49 ± 5.47
NASA TLX physical	RTPD	4.45 ± 3.75	7.05 ± 5.27
	INSO	5.20 ± 4.07	6.80 ± 4.73
	DT	5.86 ± 4.32	6.97 ± 4.99
NASA TLX temporal	RTPD	7.97 ± 4.25	12.14 ± 6.34
	INSO	8.20 ± 5.05	10.40 ± 3.84
	DT	10.06 ± 5.22	11.51 ± 4.70
NASA TLX performance	RTPD	16.40 ± 3.48	12.91 ± 5.48
	INSO	14.11 ± 3.69	12.37 ± 5.04
	DT	15.80 ± 3.13	16.48 ± 2.58
NASA TLX effort	RTPD	9.46 ± 4.18	12.54 ± 5.49
	INSO	9.17 ± 5.22	11.17 ± 3.65
	DT	11.02 ± 4.53	11.23 ± 5.01
NASA TLX frustration	RTPD	6.14 ± 3.46	10.20 ± 6.59
	INSO	8.11 ± 4.93	10.48 ± 5.41
	DT	7.48 ± 4.65	8.14 ± 4.45
BEDFORD WL	RTPD	2.97 ± 1.38	5.60 ± 2.99
	INSO	3.29 ± 2.08	5.07 ± 2.50
	DT	3.77 ± 1.93	4.87 ± 2.38

load ($F_{1,\ 31} = 34.33$, $p < 0.01$, $\acute{\eta}_p^2 = 0.53$), performance ($F_{1,\ 31} = 11.13$, $p < 0.01$, $\acute{\eta}_p^2 = 0.26$), effort ($F_{1,\ 31} = 19.05$, $p < 0.01$, $\acute{\eta}_p^2 = 0.38$), frustration ($F_{1,\ 31} = 23.75$, $p < 0.01$, $\acute{\eta}_p^2 = 0.43$) and Bedford score ($F_{1,\ 31} = 86.82$, $p < 0.01$, $\acute{\eta}_p^2 = 0.74$) of the command team (see Table 2). Operator role statistically significantly impacted upon the mental workload ($F_{3,\ 31} = 4.03$, $p < 0.05$, $\acute{\eta}_p^2 = 0.28$), and Bedford score ($F_{3,\ 31} = 2.80$, $p < 0.05$, $\acute{\eta}_p^2 = 0.21$) of the command team. A number significant interactions were observed the most notable being a statistically significant between scenario demand and operator role on effort ($F_{6,\ 62} = 19.94$, $p < 0.01$, $\acute{\eta}_p^2 = 0.20$) and Bedford score ($F_{6,\ 62} = 19.94$, $p < 0.01$, $\acute{\eta}_p^2 = 0.20$).

The HR ($F_{1,\ 42} = 11.22$, $p < 0.01$, $\acute{\eta}_p^2 = 0.21$) and temperature ($F_{1,\ 42} = 179.02$, $p < 0.01$, $\acute{\eta}_p^2 = 0.81$) of participants was statistically significantly affected by participation in scenarios (see Table 3). Operator role statistically significantly affected the temperature of the command team ($F_{3,\ 42} = 3.28$, $p < 0.05$, $\acute{\eta}_p^2 = 0.18$), a significant interaction between time measure was taken and role ($F_{3,\ 42} = 5.07$, $p < 0.01$, $\acute{\eta}_p^2 = 0.26$) was also observed.

Table 3 Means and SDs of HR and temperature before and after scenarios

Measure	Role	Baseline mean ± SD	After scenarios mean ± SD
Heart rate	Command	74.00 ± 11.77	37.25 ± 7.25
	Picture	70.91 ± 9.37	81.25 ± 8.53
	Sonar	72.08 ± 11.42	73.16 ± 13.82
	Peripheral	75.70 ± 12.70	84.70 ± 16.70
Temperature	Command	36.30 ± 0.47	37.25 ± 0.56
	Picture	35.37 ± 0.57	37.30 ± 0.33
	Sonar	36.00 ± 0.54	37.08 ± 0.27
	Peripheral	35.67 ± 0.81	37.32 ± 0.15

3 Discussion

The current work aims to provide a comprehensive understanding of submarine command team functionality across a variety of operations routinely completed by submariners. It is clear that safe submarine operation is complex, requiring the integration of vast amounts of data from different instruments by different operators [1, 3, 7]. Once complete the current work will offer insight into how the integration of information to inform a tactical picture changes with different operational requirements, building upon previous submarine command team studies which examined one task type or one scenario type [1, 3, 5, 6].

The primary objective of the baseline study is to design and conduct a set of rigorous experiments that provide a comprehensive understanding of submarine command team functionality and capacity using both subjective and objective data. There is not a great volume of work published in the submarine control room domain due to economic and security issues [2]. The current work has demonstrated that it is possible to conduct cost effective, unclassified experiments with high statistical power. The ComTET team has successfully managed to recruit and train 80 participants to complete various operations as part of a submarine command team. The success of the training package is a notable feather in the cap of the ComTET project to date. Furthermore, the statistically significant differences in both subjective (e.g. NASA TLX) and objective (e.g. HR and Temperature) indicators of WL indicate that the scenario design has been successful in terms of creating a high and low demand operational environment. Such success is further evidenced by the fact that difference WL scores were observed depending on scenario type, particularly as the WL of particular operators changed as a result of scenario type (i.e. the WL of PERI is higher during an INSO compared to DT). This indicates that the design of scenarios effectively captures different operation capacities, however when complete such work may also offer insight into where capacity may exist within the command team to accommodate future sensors and/or reducing crewing that is required [2–4].

The current research builds upon previous work investigating WL and SA in submarine control rooms [8, 9]. In the current work the subjective WL of operators is being considered from a sociotechnical perspective across a greater selection of sub-tasks (e.g. track management, sonar detection and periscope operation) and overarching operational goals (i.e. RTPD, DT and INSO). The subjective WL scores (NASA TLX and Bedford) highlight that the WL of the command team is different depending on scenario type. This indicates that future submarine control rooms may be optimised by increasing command team flexibility, where the layout, roles and technological support offered to operator's change depending upon the particular operation being completed.

3.1 Conclusions and Future Work

There is a drive for submarines to have increased capacities, whilst operating in difficult conditions and utilizing a wider array of complex instruments to facilitate operations [1–3, 7]. For this to be possible a thorough evaluation of submarine control room design and command team performance is required—including evaluations of optimizing new technologies, layouts, ways of working and interfaces. Once all analysis of data is complete the current work will provide an excellent baseline describing how submarine command teams currently function. The fact that statistically significant findings are observable with just over half the data (6 of 10 teams) included in the analysis offers great hope for findings with high statistical power—a key objective of the ComTET project. It is anticipated that once complete this body of work will offer insight into future ideas for submarine control room improvements, providing a baseline for comparison whilst also providing a methodological template for future research in this domain. This will provide the catalyst for phase 3 of ComTET—examining future concepts.

References

1. Stanton, N.A.: Representing distributed cognition in complex systems: how a submarine returns to periscope depth. Ergonomics **57**(3), 403–418 (2014)
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)
3. Dominguez, C., Long, W.G., Miller, T.E., Wiggins, S.L.: Design directions for support of submarine commanding officer decision making. In: Proceedings of 2006 Undersea HSI Symposium: Research, Acquisition and the Warrior, pp. 6–8 (2006)
4. Beevis, D., Vallerand, A., Greenley, M.: Technologies for workload and crewing reduction (No. DCIEM-TR-2001-109). Defence and Civil Institute of Environmental Medicine, Downsview, Ontario (2001)
5. Mansell, T., Tynan, A., Kershaw, D.: Investigating The Network Enabled Conventional Submarine: An Australian Perspective. Undersea Defence Technology (UDT) Pacific (2003)

6. Chalmers, B.: A work-centred approach to optimizing human-system integration in command and control for the VICTORIA class submarine, In: Proceedings of Undersea Human Systems Integration Symposium 2010, American Society of Naval Engineers (2010)
7. Shattuck, L.G., Miller, N.L.: Extending naturalistic decision making to complex organizations: a dynamic model of situated cognition. Organ. Stud. **27**(7), 989–1009 (2006)
8. Kirschenbaum, S.S.: Analysing submarine decision making: a question of levels. In: Salas, E., Klein, G.A. (eds.) Linking Expertise and Naturalistic Decision Making. Psychology Press, Abingdon, pp 189–207 (2001)
9. Loft, S., Morrell, D.B., Huf, S.: Using the situation present assessment method to measure situation awareness in simulated submarine track management. Int. J. Hum. Factors Ergon. **2** (1), 33–48 (2013)
10. Baddeley, A.: The episodic buffer: a new component of working memory? Trends Cogn. Sci. **4**(11), 417–423
11. Hutchins, E.: Cognition in the Wild. MIT Press, Cambridge (1995)
12. Stanton, N.A., Salmon, P.M., Walker, G.H.: Let the reader decide: a paradigm shift for situation awareness in sociotechnical systems. J. Cogn. Eng. Decis. Making **9**(1), 44–50 (2015)
13. Mueller, S.T.: A partial implementation of the BICA cognitive decathlon using the psychology experiment building language (PEBL). Int. J. Mach. Conscious. **2**(2), 273–288A, 2000 (2010)

Passenger Safety During Rescue and Evacuation from Passenger Vessels: A Holistic Concept for the Integration of Technical Assistance

Silvia Höckling, Alexander Kunz and Florian Motz

Abstract Worldwide the popularity of cruise holidays is constantly growing. In the last 10 years the annual passenger capacity has doubled to more than 20 million. With the deployment of modern cruise ships with capacities of up to 8000 persons (passenger and crew) new challenges arise for the international cruise industry with regards to a safe and effective evacuation in case of an emergency. This paper describes project work funded by the German Ministry of Education and Research (BMBF) to develop innovative methods and technical solutions for supporting a fast and complete evacuation of all passengers. The focus of this project is on optimizing the workflow and communication of the emergency organization onboard by providing technical support for the crew member's key functions such as counting and identifying passengers at their assembly stations.

Keywords Human factors · Evacuation management · Cruise ship · Ship evacuation · Decision support · People localization

1 Introduction

The cruise industry has been expanding rapidly in the past years. It is expected that 15 new ships will be delivered from 2016 to 2017 that will add around 40,000 lower berths to the worldwide passenger capacity [1]. This trend has been driven by increasingly larger ships with many new on-board amenities, as well as shore-based activities, offered to match the customer's demand. The ships being built in the next

S. Höckling (✉) · A. Kunz · F. Motz
Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Fraunhoferstr. 20, 53343 Wachtberg, Germany
e-mail: Silvia.Hoeckling@fkie.fraunhofer.de

A. Kunz
e-mail: Alexander.Kunz@fkie.fraunhofer.de

F. Motz
e-mail: Florian.Motz@fkie.fraunhofer.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_37

few years are essentially small cities which are already seen to represent a destination in themselves [2].

Maritime disasters in the past years have shown that there is a strong need for timely and effective evacuation of large passenger ships during emergencies. Major challenges during emergencies involve evacuation, mustering (head count of people) and abandonment procedures. For instance, accidents such as the fire on the Star Princess in 2006 showed major problems in communication and mustering of all 2690 passengers. The investigation report describes that roll calls “had to be repeated several times and took between 2 and 3 h to complete” which contributed to the situation where one deceased passenger could not be identified until over 6 h after the alarm was sounded [3]. Moreover, it was not possible to check all cabins since busy telephone lines made it impossible to request additional master keys and to inform the responsible coordinator.

A safe evacuation of cruise ships significantly depends on accurate and fast roll calls at assembly stations supplemented by searching all areas of the ship. Each crew member contributes to the evacuation and mustering procedures in various ways, for instance in ensuring that passengers proceed directly and in an orderly manner to the muster stations, preparing life boats, assisting persons with handicaps, etc. Their most difficult and time-critical emergency duties are, however, crowd management, proper and correct head counting and locating the whereabouts of missing persons onboard.

Several EU funded projects reflect the urgent need for crew support using indoor positioning systems. For instance, in the LYNCEUS project localization technology was initially tested in lab and in small scale pilots. In the follow-up project LYNCEUS2MARKET launched in June 2015, the declared aim is to implement the first market replication of products such as localizable life jackets, bracelets and cabin key cards and intelligent decision support systems [4]. In MONALISA 2.0, another EU funded project, a pilot system for people tracking was tested onboard of Ruby Princess, a cruise ship operated by Princess Cruises. In a testing phase of several months 150 crew members have been tracked in safety critial crew areas such as the engine room [5].

There are already attempts by the cruise industry to implement tracking systems on newly built cruise ships. Royal Caribbean offers RFID bracelets, known as WOWbands, on the Anthem of the Seas launched in April 2015 [6]. RFID enabled bracelets allow onboard purchases and stateroom access and can be optionally used instead of the traditional board card. At this stage the system is not used to localize passenger in case of emergency. However, with already existing luggage tracking features, localization of passengers via RFID could be the next step with the opportunity to monitor passenger movement around the decks. This technology is likely to be rolled out on every new Royal Caribbean ship but it is doubtful, if RFID retrofitting of existing ships is technically and economically feasible.

Any intelligent system that makes the gargantuan task of keeping track of several thousand people simpler will likely replace clipboards and paper lists. Although tracking solutions using RFID technologies, as well as barcode scanning techniques, are available there is still no proven system established on the market that fulfills all the complex requirements in the given context. These requirements can be summarized as follows:

1. Tracking technology should be based on wireless technology and easily be integrated in new and existing passenger ship infrastructures.
2. Low-cost and robust tracking devices should be localized in real time, and not hindered by possible interferences of large steel structures or masses of people.
3. Tracking sensors should be able to be embedded in wearable items which can be scanned quickly and without contact to allow counting and identifying its owner.

In the research project SIREVA, funded by the German Ministry of Education and Research (BMBF), a holistic concept for supporting evacuation processes is being developed that not only addresses the technical requirements listed above but also issues regarding data protection and protection of privacy, acceptance of passenger wearing tracking devices, and others.

In the end, however, a key factor for a successful implementation will not only be to which extent tracking systems meet the technical and economical requirements but also how well the system supports the common understanding between the different actors involved in emergency situations. One area of research in the SIREVA project is facilitating communication and cooperation between decision makers, monitoring and control authorities and distributed teams in order to improve reaction time, the reliability of the process and situational awareness (SA).

This paper describes and analyzes critical roles and problems in the organizational structure of emergency management that need to be addressed when specifying user requirements for an evacuation management/support systems. Furthermore, human factors techniques and methods that have been applied in the first phase of the project are presented including requirements analysis, modelling of design solutions for human machine interfaces (HMIs) and iterative evaluation of the prototypes with end users [7].

2 Understanding and Specifying Characteristics of Users, Tasks and the Organizational Environment

Various methods were applied to understand emergency management as it is practiced onboard various cruise ships of one of the world's largest cruise ship operators. These methods included reviewing available documentation material, conducting observations, and interviewing crew members. The focus in the analysis phase was on identifying characteristics of the individual roles and tasks and the organizational, technical and physical environment which defines the context that will apply to the future system.

First, emergency plans and training materials, such as work flow charts, job descriptions, checklists, and other working material were analyzed. Additionally, observing emergency drills with crew and passengers enabled the capture of typical work processes of each position in the organizational structure. As a result, the degree of workload, as well as sources of error and the effects of exceptional incidents on the workflow could be determined. Additional interviews conducted with representatives of all significant roles in the organizational structure resulted in more detailed understanding of work processes and communication behavior during emergency drills. Based on the information gathered a task analysis was conducted to understand the current system and the information flows. High level tasks were broken down into their subtasks and have been systematically documented according to the methodical approach of task decomposition by Kirwan and Ainsworth [8]. The information gathered through document analysis, workplace observations and interviews was organized into organizational and communication charts for further analysis. The following paragraphs give an overview of the most significant results.

The individual tasks of crew members vary depending on the allocated role in the organization and the type of alarm being triggered. During the first alarm phases the completeness and readiness of teams for evacuation and mustering is determined in order to be prepared for crew and passenger mustering, and the evacuation of the decks in the alarm phases following thereafter.

The hierarchical organizational structure defines a clear separation of responsibilities and a fixed communicational structure and chain of command. As shown in the organigram (Fig. 1) the command center on the bridge is the strategic lead and

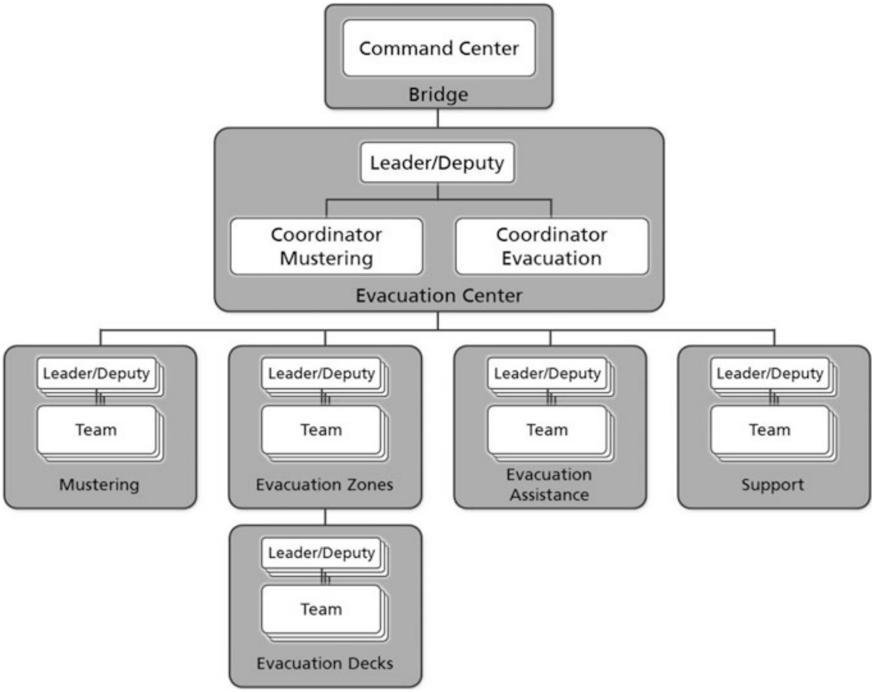


Fig. 1 Organization chart

the highest instance in the hierarchical structure. The evacuation center directly sub-ordinate to the command center collects and filters all status reports of the distributed teams and forwards the main information to the command center.

The teams in the lower hierarchical levels are organized by areas of responsibilities (evacuation, mustering, and assistance) and report solely to the evacuation center. Only an indirect exchange of information between these teams is possible via the evacuation center. Long communication channels in such hierarchical structures encourage effects like delays in coordination and communication processes. This issue could be observed on several cruise ships during emergency drills. The hierarchical organization reaches its limits if exceptional or unexpected incidents occurred in addition to standard situations. Due to the inflexible and long communication channels efficient exchange of status messages between teams is quickly affected.

Within such hierarchical structures it is typical that fixed roles and responsibilities are assigned to the crew. A consistent and complete system of task completion using paper checklists and clipboards trains the working routine of crew members and avoids human inattentiveness and forgetfulness. However, this manual process promotes long transmission times of information between the teams and control and evacuation centers. Since only one message can be received or redirected after another, waiting times due to blocked communication channels are unavoidable.

In addition, status information is predominantly verbally communicated in person, by telephone or radio. This increases the information density especially in the evacuation center where all information is accumulated. Information from many different communication channels needs to be perceived, processed and filtered in an environment of high acoustic noise level, which is distracting and increases the potential for errors.

The physical environment is an additional aspect that needs to be taken into account when specifying system requirements and designing user interfaces. Each organizational unit has a primary duty station where the teams meet after the alarm has been sounded. The command and evacuation center are located in office spaces, while the subordinate teams meet on open decks, in public spaces of the vessel or in crew areas. It is necessary to provide usable technical support for both teams using PC working station, and teams working without a fixed workplace.

The above mentioned main characteristics of the organizational structure, the individual tasks and responsibilities, and the technical environment must be considered in requirements specifications.

3 Identifying User Needs and Specifying System Requirements

The information gathered throughout the analysis phase provided indication for the potential use of computer-based support in emergency management. The concept for computer-based support solutions are described here.

Indoor positioning of crew and passenger makes the complex task of counting persons on the assembly stations easier and assists in evacuating danger areas and searching for missing persons. Additionally, capacity bottlenecks on assembly stations as well as in lifeboats and rafts could easily be identified. The system may provide in a further step decision support for an effective distribution of passenger among all remaining life-saving appliances.

A network of distributed information systems across all hierarchy levels enables a quick distribution of status messages so that information is automatically exchanged and not dependent on potentially busy telephone lines and radio channels. Only periodic manual checks of the current status may then be necessary so that the lines of communication mainly remain open for important safety-critical reports. An automatic exchange of up-to-date status information is essential to keep and maintain shared SA across all hierarchy levels and at the same time frees up the operators' resources for their core business, as described in following paragraphs.

The command and evacuation centers continually exchange information about evacuation, mustering, no-go areas, and inaccessible staircases and escape routes. The head of the evacuation center collects information from many different sources and forwards status reports to the command center. Shared SA requires an accurate common picture of the current situation. However, if information gathering and exchange consumes too many mental resources the perception and evaluation of ongoing processes is reduced which directly influences the decision making process. Especially non-standard situations, e.g. appearance of medical emergencies or unaccompanied children, create additional workload which leads to reduced SA and increased time to make the decisions. An information system used in this context is required to automatically exchange status information, so that accurate and complete information is provided and only a minimum of time is needed for collecting and transferring information.

The integration of information technology also frees up resources on the muster stations, so that the primary task of crowd control can be carried out. Calling names, checking off lists or transmitting information is then no longer required. The current status of crew and passenger mustering is directly transferred to command and evacuation center in order to ensure a continuous information flow across all hierarchy levels.

Additionally, the mustering process will be optimized in terms of also identifying passengers at muster stations to which they were not assigned. These passengers would normally be reported as "missing" on their assigned stations. However, if they are localized by tracking technology, the evacuation center would receive the correct information so that time and resources are not unnecessarily spent searching for them. An overview of all people present at muster stations also allows for monitoring the capacity of the stations and enables efficient distribution of people, if the maximum capacity of a station is exceeded. Information systems also contribute considerably to an effective management of evacuating public spaces and crew areas on the ship by providing an accurate and complete picture of the current situation. Missing persons on the muster stations can be found quickly so that crew and passenger do not need to remain for an unnecessarily long time in hazardous areas. Here too, the chain of reporting is significantly shortened as the actual state of evacuation processes is transferred automatically to all higher hierarchy levels.

Considerable time would also be saved by using tracking technology in the evacuation of passengers with handicaps. Normally teams are sent to the passengers' cabins and assist them in reaching their muster stations. Knowing their positions beforehand avoids unnecessarily long ways of the crew to the cabins, and lost time, if the passenger is not located there.

In this project user groups, mainly leader and deputies in the first three hierarchy levels, have been identified who considerably benefit from information systems in terms of reduced workload and enhanced SA. Figure 3 shows an overview of the integration of individual systems in the emergency organization. For each user group system requirements were developed that were specifically tailored to their needs and areas of responsibilities (Fig. 2).

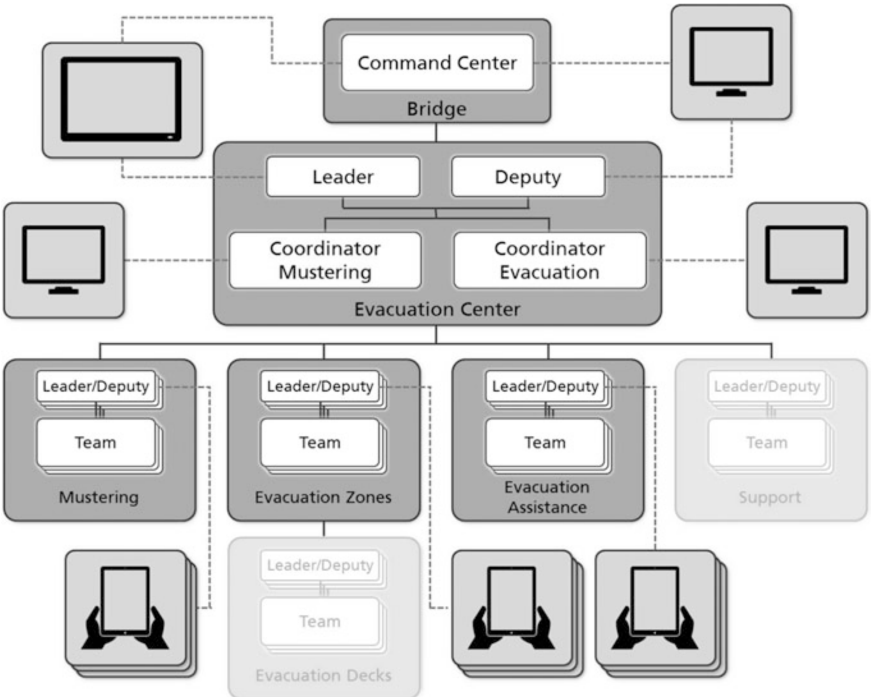


Fig. 2 Integration of information systems in the emergency organization

4 Creating User Requirements and Interface Designs

Integration of technology influences existing workflows and communications processes in the emergency management, as described in the previous chapter. Therefore, the processes have been modelled in flow charts in order to identify bottlenecks and less-obvious features and for deriving user requirements.

Figure 3 shows a workflow example for electronic identification at a muster station. Passengers are identified via the tracking sensor's unique identification number (ID) which allows determining whether a person is assigned to the station or not. If a passenger appears on the assigned station, the mobile device of the station leader lists the person as "present". Passengers not tracked at their muster station are displayed as "missing" and must be called with name and cabin number. Persons reaching their stations but carrying no tracking sensor, or a non-functioning sensor, will then be manually mustered.

In a further step, user scenarios have been created, describing system-specific interaction processes from the user's point of view. These scenarios show user's input and the system's output and supported the process of designing graphical user interfaces. First, simple conceptual layouts were created to communicate and discuss user requirements with project partners in workshops. HTML-mockups used for interviews with end users showed an advanced graphic design and allowed simple interactions to get detailed user feedback for the further development. Reactions and comments captured from a total of 18 interviewees were used to verify user requirements and to improve and refine the user interfaces.

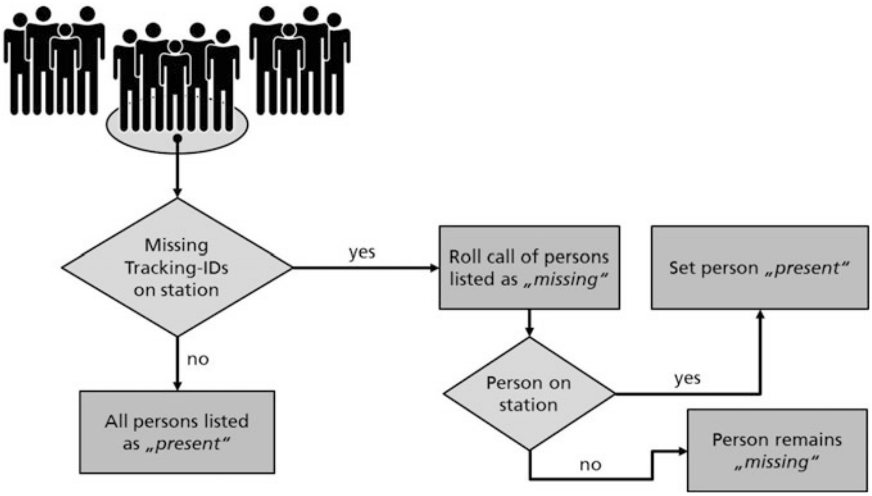


Fig. 3 Flow chart—electronic mustering process

The industrial partners in the SIREVA project used the requirement specifications and the mockups as a basis for the development of pilot systems which again will be iteratively evaluated until the end of the project in September 2016.

5 Discussion

User-centered activities carried out in an iterative fashion provided understanding of the context of use and supported in specifying user requirements, producing design solutions and evaluating the designs against requirements.

In general, it can be noted that the potential of information systems essentially is to make key processes in the workflow more effective and to support a quick flow of information among distributed teams. In an emergency situation, complex and time-critical incidents must be managed by the local teams and coordinated and monitored by the command and evacuation center. It is required to keep communication channels open and to enable short lines of communication for the development and maintenance of the necessary SA across all hierarchy levels. If task-relevant and timely information is already incorporated into the development process of situation awareness, it can lead to better decisions and results. This is particularly the case when several sudden incidents occur leaving little time for collecting and transferring information.

Through the integration of technology, resources are set free to enhance acquisition and sharing of information across the teams with the effect that actual tasks can be fulfilled more efficiently. However, the optimization level in the mustering and evacuation processes is closely related to the willingness of the passengers and crew to wear tracking sensors with the knowledge that their personal movement profiles are captured in case of emergency and during drills onboard.

Using tracking data leads to a variety of potential risks concerning privacy and data protection which must be considered. Measures must be taken so that access to personal movement data outside of emergencies and drills is restricted. The data shall only be available after appropriate authorization of the bridge crew and when the alarm for evacuation is sounded. Control mechanisms may involve logging each attempt to access these data for subsequent verification and informing the shore-side fleet operation center of the cruise line accordingly.

6 Conclusion

The potential for technical support in evacuation processes on cruise ships is to manage the information flow between the organizational units and to support decision making in time-critical work processes. Developing and maintaining SA in such complex, dynamic environment requires complete and accurate information about the current situation in order to inform effective decision processes.

This aspect is transferable to many other possible domains where people in safety-critical working areas must be evacuated in the event of immediate danger to life, for example on offshore oil rigs, power plants or in the chemical industry. The advantage of this technology lies in connecting motion profiles with personal data in order to identify the physical whereabouts of people.

The challenge is to combine the use of tracking data to the benefit of passengers and crew with the protection of their personality and data protection rights. Therefore, it is necessary to recognize risks arising from the use of positioning technologies, and to take action against misuse or illegal use by unauthorized persons.

The use of localization technologies in the evacuation of cruise ships will create a better SA by facilitating a faster flow of information and supporting essential work processes. Such a system can enhance existing evacuation management only in connection with adequate crew training. The knowledge of the individual's role and responsibility in the organization remains essential in order to allow building accurate mental models of the current situation.

References

1. Statista: New cruise ships planned worldwide from 2015 to 2017, by passenger capacity| statistic. [Online]. Available: http://www.statista.com/statistics/243857/new-cruise-ships-by-capacity/. Accessed: 07 Feb 2016
2. Rodrigue, D.: The cruise industry, The geography of transport systems. [Online]. Available: https://people.hofstra.edu/geotrans/eng/ch7en/appl7en/ch7a4en.html. Accessed: 04 Mar 2016
3. MAIB marine accident investigation branch: report on the investigation of the fire onboard Star Princess off Jamaica. [Online]. Available: https://assets.digital.cabinet-office.gov.uk/media/547c706ae5274a4290000097/Star_Princess.pdf. Accessed: 10 Jan 2016
4. Lynceus, LYNCEUS2MARKET—an innovative people localisation system for safe evacuation of large passenger ships. [Online]. Available: http://www.lynceus-project.eu. Accessed: 18 Jan 2016
5. MONALISA 2.0: People tracking system for Maritime applications. [Online]. Available: http://monalisaproject.eu/wp-content/uploads/ML2GFC_10_S1_2_Act_3.pdf. Accessed: 10 Feb 2016
6. Royal Caribbean International: Anthem of the Seas—stay smart. [Online]. Available: https://www.royalcaribbean.co.uk/our-ships/anthem-of-the-seas/staysmart. Accessed: 22 Jan 2016
7. International Standardization Organization: ISO 9241-210:2010, Ergonomics of human-system interaction—part 210: human-centred design for interactive systems. ISO Copyright Office, Geneva (2010)
8. Kirwan, B., Ainsworth, L.: A Guide To Task Analysis: The Task Analysis Working Group. Taylor & Francis, London (1992)

Part VI
Road and Rail—Vehicle Automation

What Drives Ecodriving? Hybrid Electric Vehicle Drivers' Goals and Motivations to Perform Energy Efficient Driving Behaviors

Thomas Franke, Matthias G. Arend, Rich C. McIlroy and Neville A. Stanton

Abstract Hybrid electric vehicles (HEVs) can significantly contribute to sustainable road transport, yet driver behavior has a marked effect on actual energy efficiency (i.e., the ultimate sustainability effect). The objective of the present research was to examine ecodriving motivation of HEV drivers. To this end, we recruited 39 HEV drivers with above-average fuel efficiencies (suggesting at least some degree of ecodriving motivation) and collected interview data, questionnaire responses, and fuel efficiency data. Specifically, we assessed factors that motivated drivers to drive energy efficiently as well as factors that led to reduced ecodriving behavior. Ecodriving motivation of HEV drivers was found to be particularly driven by the goals of environmental protection, cost reduction, and gamification aspects. Furthermore, relationships between drivers' most important ecodriving motivation and the level of ecodriving motivation, the achieved fuel efficiency, the level of total HEV driving experience, as well as typical HEV driving distances were examined.

Keywords Hybrid electric vehicles · Ecodriving · Motivation · Driving behavior

T. Franke (✉) · M.G. Arend
Department of Psychology, Cognitive and Engineering Psychology,
Technische Universität Chemnitz, Chemnitz, Germany
e-mail: thomas.franke@psychologie.tu-chemnitz.de

M.G. Arend
e-mail: matthias-georg.arend@s2013.tu-chemnitz.de

R.C. McIlroy · N.A. Stanton
Transportation Research Group, Faculty of Engineering and the Environment,
University of Southampton, Southampton, UK
e-mail: r.mcilroy@soton.ac.uk

N.A. Stanton
e-mail: n.stanton@soton.ac.uk

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_38

1 Introduction

Within the global agenda of decarbonization and reduction of greenhouse gas emissions electrification of road transport has become a major trend [1]. Within the last decade hybrid electric vehicles (HEVs) have become particularly widespread [2, 3]. While HEVs can significantly contribute to sustainable road transport, driver behavior has a marked impact on actual energy efficiency and, therefore, is key for the ultimate sustainability effect achieved (see e.g., [4]). Ecodriving is the term that is used to describe all those driving behaviors performed in order to increase real-world energy efficiency of a road vehicle. Research shows that both motivation and knowledge are important factors that lead to successful ecodriving behavior in HEV driving (e.g., [5]). However, which factors motivate HEV drivers to drive energy efficiently?

The objective of the present research was to examine HEV drivers' ecodriving motivations, particularly (Q1) the factors that motivate drivers to increase ecodriving efforts (including the most important ecodriving motivation, Q2) as well as (Q3) the factors that lead drivers to not drive maximum energy efficiently (i.e., other drivers' goals in conflict with ecodriving). Furthermore (Q4), relationships between drivers' most important ecodriving motivation and the level of ecodriving motivation, the achieved fuel efficiency, the level of total HEV driving experience, as well as typical HEV driving distances (i.e., mobility profile) were examined. Adopting a similar perspective as [6], the main focus was to advance understanding of the motivating factors as experienced by the drivers (i.e., by directly asking drivers via qualitative interview questions). The focus on drivers with above-average fuel efficiency (and thus a probably high ecodriving motivation) within sample recruitment enabled us to assess a particularly broad bandwidth of motivational factors.

2 Background

Besides cognitive factors such as the drivers' mental models (e.g., [7]) or knowledge (see e.g., [8]), the motivation to drive energy efficiently has been investigated as an important predictor of ecodriving behavior (see e.g., [5, 6, 9]). Usually these investigations focus on (a) how ecodriving behavior can be explained by theoretical motivational processes (see e.g., [10, 11]) and (b) how practical interventions can increase ecodriving motivation (e.g., financial reward, [12]; specific feedback types, [6, 13]).

Taking into account the exploratory character of the present research, we refer to *motivational factors* as those factors drivers report as reasons for their motivation to perform or not to perform ecodriving. Consequently, two different kinds of motivational factors are examined, motivating factors (factors related to increased ecodriving efforts) and factors conflicting the ecodriving motivation (factors related to reduced ecodriving efforts or simply factors related to not driving maximum energy efficiently).

From a theoretical perspective, motivational processes driving ecodriving behavior are multifold. One important distinction is the one between motivational (how the intention to drive energy efficiently is formed; cf. [10]) and volitional (how the existing intention to drive energy efficiently is transferred into concrete driving behavior; cf. [10]) processes. Central to motivational processes is the formation of a behavioral intention (see e.g., [14]), which—in terms of ecodriving—represents the general goal to perform energy efficient driving behaviors. The concept of behavioral intention stems from the theory of planned behavior [15], where the intention is, in terms of ecodriving behavior, the mediating factor between driver variables (attitudes, social norms, behavioral control) and driver behavior (see e.g., [10]). Hence, it is important to understand which factors form the behavioral intention to perform energy efficient driving behaviors. Consequently, our first and second research questions are (Q1) which factors motivate drivers to increase ecodriving efforts and (Q2) what is the most important ecodriving motivation.

Ecodriving motivation as defined above is the general goal to drive energy efficiently. However, in typical driving, drivers have to balance numerous goals, such as safety or time pressure (see e.g., [16, 17]). These goals and their weighting can be expected to vary between journeys and situations. This notion is specifically acknowledged by control theoretic models of (eco)driving (see e.g., [5, 17, 18]). Hence, it is also important to understand the driver goals that typically stand in conflict with ecodriving. Consequently, our third research question is (Q3) which factors lead drivers to not drive maximum energy efficiently (i.e., other drivers' goals in conflict with ecodriving).

Finally, it is also relevant to understand the relationship between the motivational structure (e.g., the most important or primary motivation) and further variables. Thus, our last research question (Q4) focuses on an exploratory quantitative analysis of the relationships between drivers' most important ecodriving motivation and (Q4.1) the level of ecodriving motivation, (Q4.2) the achieved fuel efficiency, (Q4.3) the level of total HEV driving experience, as well as (Q4.4) typical HEV driving distances (i.e., mobility profile).

3 Method

3.1 Participants

We focused recruitment on HEV drivers of the Toyota Prius (2nd gen, 3rd gen, and Prius c [in Germany sold as Yaris Hybrid]), being the most sold (see e.g., [3]) and most prototypical HEV model. From the almost 1500 Prius drivers in the www.spritmonitor.de database we invited drivers who (a) had an average fuel efficiency above the fleet-average of the vehicle model, (b) were from Germany, Austria, or Switzerland, and (c) had logged their fuel efficiency within the last 3 months. We avoided drivers who appeared to log fuel efficiency inconsistently, and sought to sample drivers across a range of above-average fuel efficiencies (i.e., from "just above average" to "top of the list"). Ethical approval was sought from and granted by the University of Southampton's Ethics and Research Governance committee (reference number 17071).

Participants in the resulting sample (N = 39) had an average age of M = 45 years (SD = 10) and an average HEV driving experience of M = 74,079 km (SD = 64,513), 92 % were male, and 56 % had a university degree.

3.2 Procedure

Telephone interviews (including questionnaire sections) were conducted ($M_{duration}$ = 48 min, SD = 8). Participants received the interview guideline before the interview and could therefore refer to the documentation as the interviewer went through the questions. The interviewer's experience with HEV driving (>6 years) facilitated the process.

After introducing the study and gaining informed consent, the interview had the following parts: (P1) ecodriving motivation, (P2) ecodriving strategies, (P3) further questions on ecodriving support, strategy development, and false beliefs, and (P4) questionnaire to assess socio-demographic and experience-related variables. The interview was audio-recorded and transcribed. The present paper focuses on section (P1) of the interview. Results regarding the other parts of the interview and further details on the methodology have been published in [5].

3.3 Scales and Measures

Interview Questions regarding Ecodriving Motivation. The three interview question to assess the data for Q1–Q3 were as follows: (1) What motivates you to drive energy efficiently? (2) What is the most important motivation for you? (3) What are the reasons for you to not drive maximum energy efficiently sometimes?

Level of Ecodriving Motivation. To assess ecodriving motivation quantitatively, a two-item scale (Cronbach's Alpha = 0.82) focusing on the behavioral intention [15] was assessed (see also [5]).

Fuel Efficiency Indicators. Two fuel efficiency indicators were computed for the analysis. First, participants exported fuel log data of their current main HEV from the spritmonitor.de database. This data included all refueling events, comprising the refueling amount (in liters) and the distance driven (in total 1.9 million km).

The data of the last 90 days was extracted (labelled FuelEfficiency.last90d.log) to standardize for seasonal variations (interviews were conducted in August 2015).

Second, participants' estimated fuel efficiency for four situations (autobahn, city, rural flat road, rural mountainous road) was assessed (in liter per 100 km, for further details on the specific questionnaire items see [5]). We computed a mean score FuelEfficiency.4sit as the average estimated fuel efficiency for the four situations (supported by a high Cronbach's Alpha = 0.87). While this score is based upon self-reported energy efficiency, it helps to counteract inter-individual differences in fuel efficiency that are caused by differences in trip profile or environmental conditions (i.e., for some drivers it is simply more difficult to achieve a good fuel efficiency because of their usual trip profile which is reflected in FuelEfficiency.last90d.log). Moreover, the self-reported energy efficiency estimations where found to be accurate in our sample (see [5]). Hence, FuelEfficiency.4sit should be conceived as the primary indicator of fuel efficiency for the present analysis (parallel to the analyses in [5]).

Finally, to eliminate the influence of vehicle model we computed the distribution parameters for each HEV model (based on fuel efficiency data from all vehicles of this model in the spritmonitor.de database) and z-standardized the fuel efficiency values of each participant on the respective distribution. Finally, all fuel efficiency values were inverted such that higher values corresponded to higher fuel efficiency.

Total HEV Driving Experience. HEV driving experience was assessed as total distance driven. The item text was: "How many kilometers did you already drive HEVs altogether?"

Typical HEV Driving Distance. As a basic indicator of HEV usage intensity we asked participants to estimate their average distance driven in a typical week with HEVs. The item text was: "How many kilometers do you drive with HEVs in a typical week? Please refer your statement to a typical week within the last three months.

3.4 Qualitative Data Analysis

We based our qualitative data analysis on thematic analysis [19]. After each interview, the interviewer and the scribe (first and second author) discussed insights and first ideas for possible codes. After familiarization with the data, the initial coding phase led to a list of codes that were relevant to the respective research question. Afterwards, the coding system was reviewed and discussed, and initial ideas for themes (i.e., thematic clusters) were revised and refined based on preliminary indicators of prevalence in the data. As only a relatively low level of abstraction of statements was targeted, this phase was less complex than for other topics in psychology (i.e., semantic rather than latent level analysis; [19]). In the final phase we again went through all transcripts and coded participants' statements with regard to the developed coding systems (i.e., clusters and sub-clusters). Within this phase some final revisions and refinements of the coding system were performed. Hence, all statements of the participants that were relevant to the respective research question were grouped into clusters and sub-clusters. Clusters group similar statements of different participants (i.e., an overarching theme that is addressed by several participants). In the results section only clusters/sub-clusters with a prevalence of $n \geq 2$ (i.e., 5 % of the sample) are described.

4 Results

4.1 (Q1) Ecodriving Motivations

Participants' responses regarding the interview question (1) What motivates you to drive energy efficiently? were grouped into seven major clusters (see Table 1). Q1C1 environmental protection and Q1C2 cost reduction where the most important motivations that drivers expressed.

Regarding Q1C1 environmental protection, many drivers (21 %) referred to the reduction of driving emissions (CO_2 [10 %], exhaust emissions [5 %], noise pollution [5 %]) and several drivers (13 %) spoke about resource conservation. Finally some drivers (8 %) stated their motivation to be the wish to take responsibility for the next generations or improve the world. Most drivers however expressed the environmental protection motivation in rather general terms (e.g., 'environmental awareness', 'make a contribution to environmental protection', 'behave in an ecologically sensible manner', 'reduce the ecological footprint').

The motivation of cost reduction (Q1C2) was less multifaceted. Drivers typically referred to the potential to pay less for fuel (i.e., for driving a certain distance) or running costs in general (e.g., also addressing the increase of durability of car parts).

The third major cluster (Q1C3) was labeled gamification because it grouped the various motivations that are usually addressed by gamified systems. Particularly, many drivers (28 %) expressed their motivation as pushing the limits (i.e.,

Table 1 Ecodriving motivations (Q1)

Cluster-ID	Cluster label	% of 39 participants
Q1C1	Environmental protection	77
Q1C2	Cost reduction	74
Q1C3	Gamification aspects	49
Q1C4	Enthusiasm about HEV technology	15
Q1C5	More relaxed driving	15
Q1C6	The car/interface motivates ecodriving	13
Q1C7	The smart and safe way of driving	8

ecodriving as a personal challenge). Furthermore, several drivers explicitly stated ecodriving to increase driving enjoyment and fun (15 %), others stated ecodriving to be a competition (10 %), sport (8 %), game (8 %) or hobby (5 %) and that their ambition motivated them to increase energy efficiency (8 %).

Apart from these three large major clusters there were also four less prevalent major clusters named by 8–15 % of the drivers. Cluster Q1C4 summarized several drivers who stated their ecodriving motivation to be driven by the enthusiasm for the HEV technology. Moreover, several drivers also stated ecodriving to be a more relaxed and elegant way of driving (Q1C5). Interestingly, only relatively few of the drivers (13 %, Q1C6) pointed out that they were motivated by the car (or more specifically the interface) to increase their ecodriving efforts. Finally, few drivers stated ecodriving to be the smarter and safer way of driving (Q1C7).

4.2 (Q2) Most Important Ecodriving Motivation

Participants' responses regarding the interview question (2) What is the most important motivation for you? were categorized with the same coding system as used for Q1. Most drivers clearly stated only one most important motivation. Two drivers however also named a second (or third) 'most important' motivation. Here only the first stated motivation entered the analysis. As shown in Table 2, the first three major clusters from the analysis of Q1 accounted for 92 % of drivers' most important motivations. Cost reduction (Q3C1) was the most important factor, directly followed by environmental protection (Q3C2) and gamification (Q3C3).

4.3 (Q3) Reasons for not Driving Maximum Energy Efficiently

Participants' responses regarding the interview question (3) What are the reasons for you to not drive maximum energy efficiently sometimes? were grouped into five major clusters (see Table 3).

Table 2 Most important ecodriving motivation (Q2)

Cluster-ID	Cluster label	% of 39 participants
Q2C1	Cost reduction	36
Q2C2	Environmental protection	31
Q2C3	Gamification	26
Q2C4	More relaxed driving	5

Table 3 Reasons for not driving maximum energy efficiently (Q3)

Cluster-ID	Cluster label	% of 39 participants
Q3C1	Time issues	67
Q3C2	Traffic compatibility	36
Q3C3	Specific driving maneuvers	23
Q3C4	Enjoyment of acceleration/speed	18
Q3C5	Long trips	13

Above all (Q3C1) time issues were the most important reason to reduce eco-driving efforts, either because of time pressure (56 %) or because of the individual wish to save time (13 %). A specific motivation herein, which was named by two drivers (5 %), was the wish to be home with the family quickly.

Regarding Q3C2 traffic compatibility, drivers stated several traffic-related factors that made them reduce ecodriving efforts (or that simply impeded ecodriving efforts), such as general bad traffic conditions (18 %), the wish to not be a traffic obstruction (13 %) or other drivers tailgating (8 %).

Regarding Q3C3 specific driving maneuvers, drivers for example named overtaking maneuvers (18 %) or dangerous situations (5 %) to be situations when ecodriving played a less important role.

Finally, several drivers named the enjoyment of fast acceleration and high driving speeds (Q3C4) and particularly long trips (Q3C5) to be factors that led to reduced ecodriving efforts.

4.4 (Q4) Relating Most Important Motivations to Level of Ecodriving Motivation, Fuel Efficiency, Total HEV Driving Experience and Typical HEV Driving Distance

To assess the relationship between the most important ecodriving motivations and the level of ecodriving motivation, fuel efficiency (FuelEfficiency.last90d.log and FuelEfficiency.4sit), HEV driving experience and typical HEV driving distance, a quantitative analysis comparing the three groups from Q2C1, Q2C2 and Q2C3 against each other was performed (see Table 3). Cohen's d for differing group sizes (d) was calculated as standardized effect size measure [20]. Statistical significance was computed via a permutation test [21]. Nonetheless, due to the small group sizes (and thus small power), we base the interpretation of results primarily on effect sizes (i.e., moderate effects, respectively $d \geq |0.50|$, are interpreted).

As indicated in Table 4, there were no significant and no effects bigger than moderate for the comparison between the group with cost reduction and the group with environmental protection as primary ecodriving motivation. This is clearly different for the comparisons of both groups with the group with gamification as primary ecodriving motivation: Gamification was, when compared to

Table 4 Relationship between most important ecodriving motivations and further variables

Variables	Cost reduction versus environmental protection		Gamification versus environmental protection		Gamification versus cost reduction	
	d	(*p*)	*d*	(*p*)	*d*	(*p*)
Level of ecodriving motivation	0.27	(0.540)	0.66	(0.174)	0.35	(0.456)
FuelEfficiency. last90d.log	−0.31	(0.435)	0.18	(0.681)	0.54	(0.201)
FuelEfficiency.4sit	−0.08	(0.836)	0.51	(0.243)	0.66	(0.127)
Total HEV driving experience	−0.32	(0.430)	−1.11	(0.001)	−1.16	(0.001)
Typical HEV driving distance	0.22	(0.610)	−0.74	(0.102)	−0.69	(0.107)

environmental protection, positively related to the level of ecodriving motivation and FuelEfficiency.4sit but negatively related to both, total HEV driving experience and typical HEV driving distance. The same applies to the comparison of gamification to cost reduction, except for FuelEfficciency.last90d.log, which was additionally positively related to gamification. Summed up, gamification was generally positively associated with fuel efficiency and negatively related with HEV driving experience and the typical HEV driving distance (i.e., gamification is by tendency less often the most important ecodriving motivation for drivers' driving more km per week).

5 Discussion

5.1 Summary of Results

The objective of the present research was to identify and understand HEV drivers' ecodriving motivations as well as factors that lead drivers to not drive maximum energy efficient. Results regarding Q1 indicate that the goals to protect the environment and to reduce costs were important factors for drivers as were aspects of gamification such as seeing ecodriving as a challenge (pushing the limits), competition, sport or game. Regarding Q2, drivers' most important ecodriving motivations were cost reduction, environmental protection and gamification aspects. Furthermore, regarding Q3, time issues, traffic compatibility, implementation of specific driving maneuvers, and enjoyment of the car's acceleration and speed were reported to be factors that conflicted the ecodriving motive. Finally, regarding Q4, comparisons between the resulting three major groups regarding the most important ecodriving motivation (cost reduction versus environmental protection versus gamification aspects) revealed the following: First, there were no relevant differences regarding the level of ecodriving motivation, achieved fuel efficiency, HEV driving experience, and mobility profile between the groups with cost reduction and environmental protection as primary motivation. Yet, gamification was, compared to both other groups, related to higher fuel efficiency, less total HEV driving experience and a mobility profile comprising shorter typical weekly driven distances (criterion for interpretation: moderate effect sizes, only effect of total HEV driving experience was also statistically significant).

5.2 Implications

Although some limitations have to be taken into account when interpreting the results of the present study (specifically the relatively small sample size and the correlational design of our study regarding Q4), some interesting tentative conclusions can be drawn.

While the relatively 'classical' ecodriving motivations related to cost savings (i.e., individual utility) and environmental protection (i.e., social utility) stand out as the most important motivational factors, the cluster of gamification aspects is also particularly strong (i.e. prevalent). This pattern and particularly the positive relationship of the gamification motivation with fuel efficiency gives some support to the notion that gamification approaches to ecodriving can indeed have a considerable potential (i.e., in relation to approaches focused on costs or the environment). However, the results regarding the relation of gamification and driving experience could be interpreted as giving some indication that motivations related to gamification are less prevalent or play a less important role (less often act as primary ecodriving motivation) with increased driving experience. Therefore, using different motivational factors at different states of experience could enhance the long-term stability of ecodriving-motivation. However, further more controlled studies would be needed to allow a definite conclusion about these patterns of results.

All in all, the present study provides a further step towards a comprehensive understanding of (a) those factors HEV drivers perceive as motivating their use of energy efficient driving behaviors as well as (b) those factors that have the potential to conflict with ecodriving. For designing optimal intervention approaches and advanced driver assistance systems (i.e., ecodriving support systems) these factors should be taken into account, for example by providing drivers with precise information to easily balance the goals of energy efficiency and time efficiency (e.g., how much energy does it cost to save five minutes on this trip). Ultimately, such design approaches could contribute to a greater sustainability effect of HEVs.

Acknowledgments This research was partly supported by a DAAD grant to the first author and an ERASMUS+ grant to the second author. We gratefully thank Prof. Dr. Josef Krems for providing parts of the research infrastructure; Juliane Ketzer and Christian Rosomm for their assistance with processing the interview data; and Christiane Attig for assistance in manuscript preparation.

References

1. Dijk, M., Orsato, R.J., Kemp, R.: The emergence of an electric mobility trajectory. Energy Policy **52**, 135–145 (2013)
2. Al-Alawi, B.M., Bradley, T.H.: Review of hybrid, plug-in hybrid, and electric vehicle market modeling studies. Renew. Sustain. Energy Rev. **21**, 190–203 (2013)
3. U.S. Department of Energy.: U.S. HEV Sales by Model. http://www.afdc.energy.gov/data/10301 (2015)
4. Barkenbus, J.N.: Eco-driving: an overlooked climate change initiative. Energy Policy **38**, 762–769 (2010)
5. Franke, T., Arend, M.G., McIlroy, R.C., Stanton, N.A.: Ecodriving in hybrid electric vehicles—exploring challenges for user-energy interaction. Appl. Ergon. **55**, 33–45 (2016)
6. Stillwater, T., Kurani, K.S.: Drivers discuss ecodriving feedback: goal setting, framing, and anchoring motivate new behaviors. Transp. Res. Part F **19**, 85–96 (2013)
7. Pampel, S.M., Jamson, S.L., Hibberd, D.L., Barnard, Y.: How I reduce fuel consumption: an experimental study on mental models of eco-driving. Transp. Res. Part C **58**(D), 669–680 (2015)
8. McIlroy, R.C., Stanton, N.A.: A decision ladder analysis of eco-driving: the first step towards fuel-efficient driving behaviour. Ergonomics **58**, 866–882 (2015)
9. Mansfield, L.R., Guros, F., Truxillo, D.M., Macarthur, J.: Individual and contextual variables enhance transfer for a workplace eco-driving intervention. Transp. Res. F **37**, 138–143 (2016)
10. Lauper, E., Moser, S., Fischer, M., Matthies, E., Kaufmann-Hayoz, R.: Psychological predictors of eco-driving: a longitudinal study. Transp. Res. F **33**, 27–37 (2015)
11. Schießl, C., Fricke, N., Staubach, M.: Identification and analysis of motives for eco-friendly driving within the eco-move project. IET Intell. Transp. Syst. **7**, 46–54 (2013)
12. Lai, W.: The effects of eco-driving motivation, knowledge and reward intervention on fuel efficiency. Transp. Res. D **34**, 155–160 (2015)
13. Dogan, E., Bolderdijk, J.W., Steg, L.: Making small numbers count: environmental and financial feedback in promoting eco-driving behaviours. J. Consum. Policy **37**, 413–422 (2014)
14. Cristea, M., Paran, F., Delhomme, P.: The role of motivations for eco-driving and social norms on behavioural intentions regarding speed limits and time headway. Int. J. Social Behav. Educ. Econ. Bus. Ind. Eng. **6**, 1486–1491 (2012)
15. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Dec. **50**, 179–211 (1991)
16. Dogan, E., Steg, L., Delhomme, P.: The influence of multiple goals on driving behavior: the case of safety, time saving, and fuel saving. Accid. Anal. Prev. **43**, 1635–1643 (2011)
17. Summala, H.: Towards understanding motivational and emotional factors in driver behaviour: comfort through satisficing. In: Cacciabue, P.C. (ed.) Modelling Driver Behaviour in Automotive Environments, pp. 189–207. Springer, London (2007)
18. Fuller, R.: Driver control theory: from task difficulty homeostasis to risk homeostasis. In: Porter, B.E. (ed.) Handbook of Traffic Psychology, pp. 13–26. Academic Press, London (2011)
19. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**, 77–101 (2006)
20. Hothorn, T., Hornik, K., van de Wiel, M., Zeileis, A.: Implementing a class of permutation tests: the coin package. J. Stat. Softw. **28**, 1–23 (2008)
21. Cohen, J.: Statistical Power Analysis for the Behavioral Sciences, Rev edn. Academic Press, New York (1977)

A Countdown to Manual Driving: How Do Drivers Get "Back-in-the-Loop"?

Annika Larsson

Abstract In automated driving, transitions of control from and to the driver are an important safety issue; studies have shown that drivers are slower to respond to hazards immediately after resuming control. This study explores the use of countdowns to mitigate transition problems before drivers are pushed back into control in a simulated driving task. In order to test the effectiveness of the countdowns, immediately after one of the four countdowns, the driver had to react to a stationary vehicle (TTC 5 s) revealed by a lead vehicle changing lanes. Driver responses were logged and analysed. The results indicate that drivers respond more alike if they have the opportunity cognitively to get back into the loop before they need to respond. However, the signal needs to be perceived by the driver, otherwise it can be missed. Drivers thus need support during the switch to manual control, as a focus on other tasks can cause them to miss signals from the vehicle. If such support cannot come from the vehicle, it needs to come from the infrastructure.

Keywords Automated driving · Take-over-request · Tactical responses · Out-of-the-loop

1 Introduction

When the vehicle itself drives for parts of a journey, transitions of control from and to the driver pose a potential risk. Studies have shown that drivers are slower to respond to critical scenarios immediately after resuming control (e.g. [1–3]), which could be risky at higher speeds. We believe that one reason for difficulties when resuming control is that drivers do not monitor the road in a structured way when automation is driving the vehicle, for the simple reason that they do not need to. Drivers switch from needing to monitor traffic to delegating that to the vehicle. When the need to drive and monitor is removed, drivers instead perform other tasks

A. Larsson (✉)
Autoliv Development AB, Wallentinsvägen 22, 44783 Vårgårda, Sweden
e-mail: Annika.Larsson@Autoliv.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_39

(see e.g. [4]) which they come to see as their new main tasks. Drivers even delay transitions back to manual control, if possible. Merely keeping their hands on the wheel does not appear by itself to lead to structured monitoring. Gold et al. [1] found that keeping the hands on the wheel when in automated mode did not lead to a faster response to a critical situation. In our view, goal-directed monitoring of traffic was still lacking, which in turn affects how drivers are able to respond to risky situations as well as even to signals prompting them to resume control.

It is likely that changing the task from driving to something else produces a cognitive cost [5]. The theory of task switching (ibid) suggests that activating another set of rules, associated with the new goal, is effortful. This effort will then show as longer deliberation as well as slower responses until the switch is completed. Therefore, it may be effortful to resume the driving task after having been in automated mode. Helping drivers switch their task goal ahead of resuming physical control of the vehicle could therefore provide help in handling hazardous situations when back in control again.

Resuming control while moving is not necessarily simple or straightforward. Following a signal to resume control, it can take drivers up to over 10 s merely to put their hands on the steering wheel again [6]. Thus, some drivers may not resume control early enough to be able to handle risky situations immediately after a transition of control. Banks and Stanton [7] found that 22 % of their participants failed to resume control of an automated vehicle following an auditory system warning. Rather, the drivers required a prompt by a safety driver to resume control and put their hands back onto the steering wheel.

Some studies have also examined transition prompts with pre-warnings, such as [8], and [9]. Willemsen et al. [8] tested both a countdown to manual control and an instant pass-back to manual via sound with visual indication. However, they did not compare responses to critical situations for the two indications. [9] transitioned control to the driver after a total of 7 s, consisting of a two-second spoken message followed by a five-second visual countdown. Participants were also asked to confirm the transition by pressing a button on the steering wheel, all providing time for a mental switch of control ahead of the physical switch. In contrast to the earlier research cited on switching from automated to manual control, [9] found no significant difference in response times to critical events after automation compared to baseline. One difference is their use of a countdown ahead of the actual pass back to manual control, as the other studies previously mentioned all pushed drivers into control immediately when issuing a beep sound of some sort along with a visual indication.

Providing a countdown is one way to allow drivers to anticipate when they will receive control again. Indeed, [10] found a trend that anticipated transitions of control lead to a somewhat shorter time to take control and to get back into the driving loop. The benefit of anticipation or countdowns may be different given what drivers are doing immediately before, though. [11] also found that a scheduled transition was beneficial for drivers who were looking at video when responding to a critical event, compared to a non-scheduled transition. Radio listeners were faster to respond though, both with scheduled and non-scheduled transitions. There was no significant effect of the scheduled strategy for radio listeners, but the standard deviation of response times became smaller. We speculate that this is due to drivers being able to both listen to the radio and somewhat stay in the driving task, whereas when watching video the task switch becomes more complete. Drivers may also consciously choose to keep performing a non-driving task they have started, if first provided with a more optional resumption signal, until required to resume control immediately [5].

Drivers appear to take about five seconds to be able to handle more demanding driving situations again. Mok et al. [12] provided drivers with an auditory message, at the start of which control was immediately handed back to them 2, 5 or 8 s before having to manually negotiate a curve. In the 2 s condition, all participants failed to negotiate the curve, whereas in the 5 and 8 s condition all participants managed to do so. The same also applies for responding to hazards. In Gold et al. [13], driver strategies changed from a fairly even distribution of steering, braking and steering and braking at a five second TTC transitioning from automated driving, through a mix of steering and steering and braking (7 s TTC from automation, as well as 5 s TTC from manual), to only steering for 7 s TTC from manual. Thus, available time also appears to have an effect on the tactical side of driving, not only on response times or similar operational measures (see Michon [14]).

In this study, we intend to investigate whether a countdown will cause drivers to resume control in a different way, compared to when being handed control immediately. In particular, we focus on whether countdowns make a difference to drivers' tactical strategies for resolving a critical situation just after control transition.

2 Method

This study explores the use of countdowns to ease drivers back into manual control in a simulated driving task.

The countdowns tested were 0, 5, 10 and 30 s long. At the start of the countdown, an information sound was played, and an informational display counting the seconds to transition was shown. At five seconds to the end of the countdown, the same information sound was played again (the first sound for the 5 s countdown). At the end of the countdown (the only sound for 0 s), a distinct warning sound was played. At the end of this warning sound, the driver was pushed back in control of the car.

2.1 Participants

39 participants (21 male), recruited from a sample of employees at Volvo Cars participated in the study, as part of their normal working day. Participants held a

valid driving license and ranged in age from 25 to 60 years (mean 40 years, sd 8.8 years), and drove on average 300 km per week. Thirteen participants, with at least one year of experience with ACC, were randomly distributed between the groups.

2.2 Simulator

The study was conducted in a fixed-base simulator with a 180° field of view, in the Volvo Cars HMI lab. Side mirror screens were unfortunately out of order during the study, but a rear view mirror screen displayed the scene behind the participant's vehicle. The participant controlled the steering wheel as well as gas and brake pedals, gear shifting was automatic. In automated mode, the simulator disregarded any input from the wheel or pedals. To indicate automated mode, the instrument cluster displayed a black screen with just the current speed in the same location where the speedometer is located during manual drive. Automated driving was based on an ACC-type functionality, set to a speed of 95 km/h. In automated mode, the participant followed a lead vehicle travelling at 90 km/h for parts of the drive.

The simulated road consisted of a two separated two-lane roads with moderate traffic and gentle curves, so that the road was not completely straight. Drivers could not drive into oncoming traffic. Participants were instructed to keep to the road speed, 90 km/h. Participants drove manually for around one minute before the vehicle started driving itself. They were then shown a message on the simulator wall that automation was active. Participants were in automated mode for five minutes before manual control was returned, at the same time as the lead vehicle changed lanes. This loop was repeated four times for each participant. The critical scenario only occurred once for each driver, after the lead vehicle had switched lanes.

During the practice session, participants were introduced to the Android mobile phone game "Dots" by Playdots, Inc. The game was played on a Samsung Galaxy S4 mini. Participants were instructed to play the game during automated driving. The game was set to a time limit of one minute, during which participants were instructed to score as high as possible by connecting dots on the screen. After the minute was up, participants were instructed to play again, and repeat for the entire drive while in automated mode. At the start of each drive, if the participant did not start playing the game, the test leader asked them to do so.

2.3 Procedure

Participants were welcomed in the simulator, and asked to adjust the seat and the steering wheel so that they were comfortable. They were then allowed to drive for about 5 min, so that they were at ease with driving the simulator. Participants were trained in having control passed back to them from the vehicle, as well as to recognize the accompanying auditory and visual signals used to indicate that automated driving had ended. Each participant heard and practiced the signals at least 3 times for each type of signal.

For the test, each participant drove four drives in the simulator, trying all four countdowns in a balanced order. In order to test the effect of the countdowns, immediately after one of the four countdown, the driver had to react to a stationary vehicle (TTC 5 s) revealed by a lead vehicle changing lanes. Drivers' responses were logged and analysed. The study was conducted in a between-group design, split by countdown type.

2.4 Critical Scenario

The participant's vehicle was driving without a lead vehicle for the majority of the drive, switching to a lead truck during the last minute of automated driving to conceal the possibility of the critical scenario occurring (Fig. 1).

During the critical scenario, the lead vehicle, a truck that had been obscuring the road in front, changed lanes to reveal a stationary car in the right lane. Simultaneously, the left lane filled with other cars, meaning the participant had to be on the lookout for obstacles to that side as well. In order to avoid the stationary vehicle in front, the participant had to brake or to steer into the left lane, avoiding the other cars there (Fig. 2).

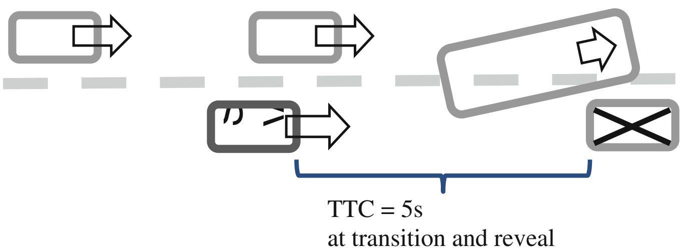


Fig. 1 Critical scenario, subject vehicle in the right lane, lead vehicle changing lanes in front of the stationary vehicle

Fig. 2 Critical scenario as seen in the simulator

3 Results

Steering input of over 4°, and pedal input of 10 % of max was used to distinguish conscious input as well as signal from noise in the data. Smaller steering input than 4° was difficult to tell apart from noise in the simulator. Results indicate that a longer countdown resulted in participants using the steering wheel by itself to a lesser degree. Instead, they tended to both steer and brake, or only brake, in order to avoid the stationary vehicle. For the 0 s transition with only the warning sound, two participants did not move to avoid the stationary vehicle. This was not the case for any of the countdowns (see Fig. 3).

At the moment of transition to manual, pedal use was very similar across the four conditions. Approximately half of the participants in each condition did not press any pedal at all, whereas the other half pressed the gas pedal (Fig. 4).

One of the participants in the 0 s group who did not make any attempt to avoid the stationary vehicle, did not use any of the vehicle controls after the control transition. The other participant in the 0 s group who did not attempt to avoid the stationary vehicle was not pressing any pedals at transition either, but started pressing the gas pedal close to the stationary vehicle. No steering above the chosen threshold of 4 degrees was detected, but smaller steering adjustments to the left can be seen in the data just after the participant depressed the gas pedal. So, this participant may have tried to steer away from the stationary vehicle.

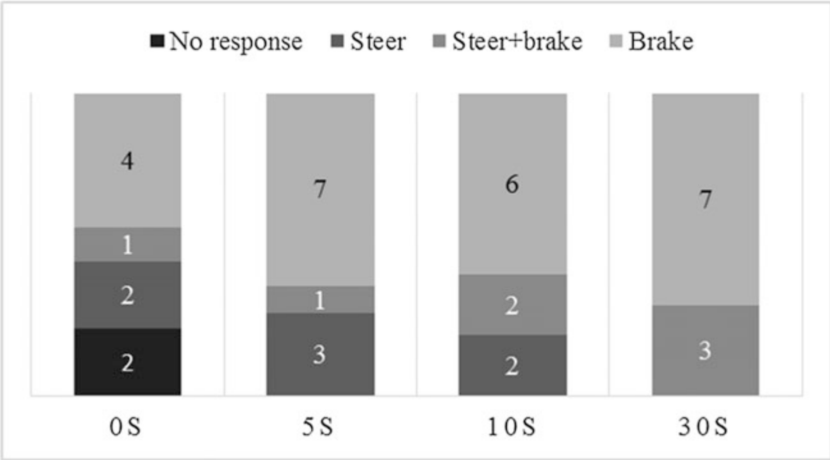


Fig. 3 Response types, ordered by countdown

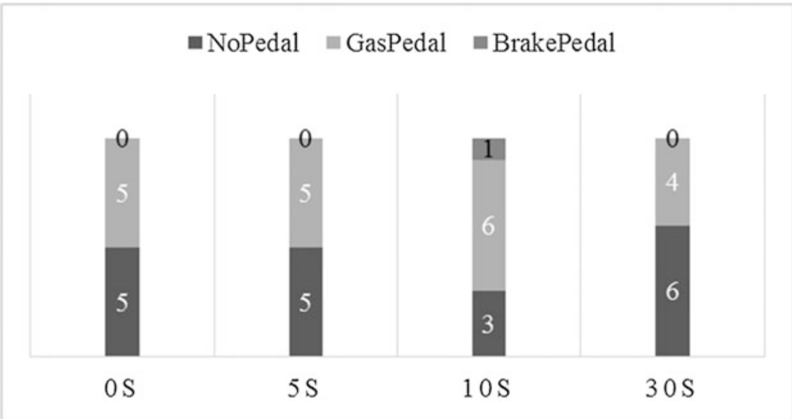


Fig. 4 Pedal use at transition to manual control across the four groups

4 Analysis

The transitions and the critical scenario in this study all provided drivers with 5 s time-to-collision (TTC) to the stationary vehicle. Five seconds is enough to respond, but can be difficult if one has recently been driven by an automated vehicle (e.g. [1]). For this five second take-over-request and TTC, it was apparent that drivers in the 0 s group did not always manage to resume control and perform an avoidance maneuver. One participants in the 0 s condition did not act to avoid the stationary vehicle at all, another in the same group only acted very vaguely, whereas all participants in the other countdown groups performed clear avoidance maneuvers.

It could be assumed that the lack of response would appear as a general problem also at the transition point for the 0 s group compared to the other conditions. Yet, pedal use at the transition point is similar across all four conditions. Instead, it may be that the drivers in the 0 s condition who did not perform an avoidance maneuver did not notice that they were back in control of the vehicle, as they did not need to confirm the switch (cf. [9]). The fact that approximately half of the participants across all conditions did not press any pedal at transition, whereas the other half pressed one of the pedals, warrants further investigation.

Similarly to what was found in [13], the results of this study indicate that driver strategies to respond to a critical situation when back in manual control appear to become increasingly alike if more time is available to deliberate. The results indicate that countdowns can cause this streamlining effect on tactical driving strategies even without longer manual control. Thus, it appears that a countdown to manual driving can start the task switching process earlier. This may also be the reason for [9] not finding any significant difference in response time to critical events for drives with countdowns from automated driving compared to baseline manual driving.

Other explanations are of course also possible. Perhaps previous research that has found that drivers are slower to respond after automated driving also includes some kind of double-checking on the part of the driver, to see whether they are really in control again. As previously mentioned, these other studies generally have provided the driver with one single beep before pushing them back in control. Such a warning could also take drivers a while to identify in itself, and they may even miss the signal. For these reasons, drivers need to be given the opportunity to become aware of the transition to manual control, and not miss it if they focus on a non-driving task.

4.1 Conclusions

If it is not possible to provide the driver with a warning that automation will no longer be controlling the vehicle in time before a control switch, mitigating systems or infrastructure may need to be in place.

Providing not one, but several temporally spaced and sufficiently salient transition signals appears to be important for drivers to be able to handle potential hazards after a period of automated driving. Handing control to a driver without asking for confirmation may also be risky, as the driver may fail to notice the pass-back signal. If it is not possible to provide the driver with a warning that automation will no longer be controlling the vehicle in time before a control switch, mitigation is necessary.

If drivers cannot be supported during or immediately after the switch, one way to handle the risk of drivers not responding to hazards could be only to allow that level of automation on roads with low speeds. Thus, infrastructure could mitigate difficulties of transitioning to manual control from automation, so that the risk of severe accidents can be reduced.

Acknowledgments Data was collected in collaboration with Volvo Car Corporation.

References

1. Gold, C., Damböck, D., Bengler, K., Lorenz, L.: Partially automated driving as a fallback level of high automation. In: Tagung Fahrerassistenzsysteme. Münich, 28–29 Nov 2013
2. 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. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, **58**(1):2063–2067 (2014) 10.1177/1541931214581434
3. 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: IEEE 18th International Conference on Intelligent Transportation Systems, pp. 2458–2464 (2015)
4. Blanco, M., Atwood, J., Vasquez, H.M., Trimble, T. E., Fitchett, V.L., Radlbeck, J., Morgan, J.F.: Human factors evaluation of level 2 and level 3 automated driving concepts. (Report No. DOT HS 812 182). Washington, DC: National Highway Traffic Safety Administration, (2015)
5. Rubinstein, J.S., Meyer, D.E., Evans, J.E.: Executive control of cognitive processes in task switching. J. Exp. Psychol. Hum. Percept. Perform. **27**(4), 763–797 (2001)
6. 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. In In 6. Tagung Fahrerassistenzsysteme, pp. 1–26. München, (2013)
7. Banks, V.A., Stanton, N.A.: Keep the driver in control: automating automobiles of the future. Applied Ergonomics, **53**:389–395 (2016) 10.1016/j.apergo.2015.06.020
8. Willemsen, D., Stuiver, A., Hogema, J.: Transition of control : automation giving back control to the driver. In: T. Ahram, W. Karwowski, Marek T. (eds.), Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE, pp. 2119–2127. Kraków, Poland 19–23 July 2014
9. Johns, M., Miller, D.B., Sun, A.C., Baughman, S., ZXhang, T., Ju, W.: The driver has control: exploring driving performance with varying automation capabilities. In: Proceedings of the Eighth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 400–406 (2015)
10. 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: Proceedings of the 5th Transport Research Arena (2014)
11. Blommer, M., Curry, R., Kochhar, D., Swaminathan, R., Talamonti, W., Tijerina, L.: The effects of a scheduled driver engagement strategy in automated driving. 59th Annual Meeting of the Human Factors and Ergonomics Society, pp. 1681–1685 (2015)
12. Mok, B.K., Johns, M., Lee, K.J., Ive, H.P., Miller, D., Ju, W.: Timing of unstructured transitions of control in automated driving. In: 2015 IEEE Intelligent Vehicles Symposium (IV), pp. 1167–1172. Seoul, Korea (2015)
13. Gold, C., Damböck, D., Lorenz, L., Bengler, K.: "Take over!" how long does it take to get the driver back into the loop? Proceedings of the Human Factors and Ergonomics Society Annual Meeting, **57**(1):1938–1942 (2013) 10.1177/1541931213571433
14. Michon, J.A.: A critical review of driver behaviour models: what do we know? what should we do? In: Evans, L.A., Sching, R.C. (eds.) Human behaviour and traffic safety, pp. 487–525. Plenum Press, New York (1985)

Assessing Truck Drivers' and Fleet Managers' Opinions Towards Highly Automated Driving

Natalie Richardson, Fabian Doubek, Kevin Kuhn and Annika Stumpf

Abstract Highly automated driving is on the advance and is linked to various benefits such as increased overall comfort for the driver as well as rising fuel and transport efficiency. Especially within the domain of truck driving in terms of long distance haulage, it seems as if highly automated driving systems could positively enhance the conditions for drivers and transport companies. Literature suggests that the acceptance of new technologies is a major determinant of whether a developing technology is used. This paper describes an approach aimed at assessing truck drivers' attitudes towards highly automated driving. Furthermore, fleet managers' opinions regarding the potential and limitations of highly automated driving systems were queried to investigate whether transport companies would invest in this new technology. Data was collected by an online and paper-based questionnaire. The results reflect the major areas of acceptance and doubts towards highly automated driving from the drivers' as well as the fleet managers' perspectives. Both groups are found to be the most concerned about legal liability issues and the general safety and reliability of such technology. Comfort and safety seem to have the biggest influence on the acceptance of highly automated driving. Truck drivers were concerned about reduced driving pleasure as well as being redundant. Results show that the majority of truck drivers do not have a clear idea of highly automated systems. In contrast, fleet managers claim to have an idea of the system.

Keywords Highly automated driving · Acceptance · Truck driver · Fleet manager

N. Richardson (✉) · F. Doubek · K. Kuhn · A. Stumpf
Lehrstuhl Für Fahrzeugtechnik, Technische Universität München, Boltzmannstrasse, 1585748 Garching, Germany
e-mail: richardson@ftm.mw.tum.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_40

1 Introduction

In the future, transportation demands can be expected to increase in line with trade growth [1], indicating a rising importance of truck haulage. To meet the requirements of a fast growing economy, transport companies need to enhance fuel and transport efficiency while taking a maximum safety level into account. Highly automated driving is considered promising in meeting these requirements. Highly automated driving is defined as a system able to perform the driving task completely by itself [2]. Permanent system monitoring by the driver is not necessary, though he must be able to take over control within a predefined take over time if required by the system [2]. This advancing technology seems to have the capability to strike a balance between growing mobility demands, road safety and environmental protection [3, 4]. Especially in the domain of truck driving in terms of long distance haulage, it seems as if highly automated driving systems could positively enhance the conditions for drivers and transport companies.

By enhancing fuel efficiency, thereby decreasing gas emissions and traffic congestion, transport companies might be able to increase their efficiency and profitability [3]. While commercial haulage operators might decrease the overall operational costs, truck drivers would be the major beneficiaries when it comes to driving comfort and road safety [4].

While higher levels of automation offer benefits, these might be undermined by several downsides. Relevant issues mainly relate to the driver's changing role, resulting in an exclusion from the actual driving task [5]. As an outcome, phenomena, such as a reduced situation awareness leading to inappropriate levels of trust in the automated system, might increase [5–7].

The acceptance of new technologies is of primary importance and a major determinant in whether a developing technology is used [8]. Therefore, the introduction of highly automated driving in trucks will only be beneficial if drivers are willing to use the system. Previous research in the automotive domain indicated that acceptance towards highly automated systems is significantly affected by the user's extent of technical affinity [5, 9]. If automated systems are not accepted by drivers, the positive impact of automation on the transportation industry might be mitigated [10, 11]. However, there is no clarification of the current acceptance level of truck drivers and transport companies towards the usage of highly automated systems within trucks. Current research has already investigated car drivers' trust as well as acceptance regarding highly automated systems [11]. As trucks are characterized as commercial transport vehicles and therefore workspaces, findings on the acceptance of car drivers towards highly automated driving cannot be transferred onto the field of truck driving easily.

2 Research Question

This study investigates truck drivers' as well as fleet managers' acceptance towards highly automated driving to determine acceptance enhancing elements. Furthermore, affinity for technology was determined. Information about concerns and needs were collected. In addition to that, forwarding agents' perspectives on benefits and concerns regarding highly automated driving were analyzed.

3 Method

The quantitative method used was a questionnaire. Data was collected by applying it either as an online or as a paper-based questionnaire. This approach reflects the method of explorative research attempting to establish a foundation for formulating hypotheses leading to future studies. Two surveys were created, for truck drivers as well as transport companies. The online questionnaire was implemented using the open source tool LimeSurvey [12].

3.1 Questionnaire Construction

Drivers' questionnaires were divided into the following sections: demographic information such as age, gender, number of years in employment as a truck driver, mileage as well as driving frequency. Furthermore, technical affinity and truck drivers' general opinions towards the idea of using highly automated driving systems were assessed. For measuring truck drivers' affinity towards technology, the TA-EG (Technology Affinity—Electronic Devices) [8] questionnaire, consisting of 19 items, was used. The resulting scores between 1 and 5 were divided into the four categories enthusiasm (5 items), competence (4 items), negative (5 items) as well as positive attitude (5 items). In addition, the survey recorded information about potential benefits and concerns regarding the implementation of highly automated driving, as well as necessary conditions under which truck drivers would like to be supported by such systems.

Fleet managers' surveys also focused on the attitude towards highly automated systems within trucks and investigated thoughts and ideas of a potential shift in drivers' work tasks and environments. To investigate whether transport companies would invest in this new technology, the questionnaire further analyzed benefits and limitations of highly automated driving systems from forwarding agents' perspectives.

Both subject groups were informed about the aim of the survey study and that they would need approximately 20 min to complete the questionnaire. A brief definition of the term highly automated driving was provided in written form within the survey. Participants were informed that all data recorded for the study was anonymized.

3.2 Item Analysis

For the purpose of measuring truck drivers' technical affinity, the standardized TA-EG questionnaire [9] was used. Questions regarding concerns, benefits and usage of highly automated systems were translated into German using questions from Kyriakidis et al. [11] on the public opinion on automated driving. An empirical examination of the questionnaire should be conducted to check the reliability of items regarding the general acceptance and attitude towards higher automation. In order to verify whether the test items meet their purpose, an item analysis was conducted using the statistical analyzing software SPSS [13]. In total, eight items, selected to measure the drivers' attitudes towards highly automated driving systems, were taken into account for analysis. Item difficulty was found to vary between 2.62 and 3.31 (scale from 1 = disagree strongly to 5 = agree strongly). The major share of the item difficulties is above the average rating of 3. This means that the majority of the items the subjects agreed to were difficult to measure psychometrically. The scientific merit of each item was valued using Cronbach-α for standardized items (Cronbach-α = 0.855). The inter-item-correlation spread from 0.051 to 0.757 (mean 0.425). The adjusted part-whole selectivity came to r_ix > 0.30, except for one item ("if the truck drives highly automated, other tasks can be pursued") with r_ix = 0.235. Hence, the item was excluded from further analysis.

3.3 Procedure

Data was collected by using either an online or paper-based questionnaire. A newsletter referring to both questionnaires was sent out by the Bavarian Association of Shipping Companies to its subscribers. Both online questionnaires were published on different online community forums. Truck drivers were approached using a list of test subjects provided by the Institute of Automotive Technology of the Technische Universität München. Furthermore, a paper-based version of the truck driver questionnaire was handed out to drivers at two regular driver meetings organized by the traffic police of Bavaria.

A link to the questionnaire addressing transport companies was posted by the German Logistics Association on their official Facebook site as well as in the research and development forum of the associated Xing group. The questionnaire was also sent out to various transport companies based on the general e-mail contact information found on the companies' internet websites.

3.4 Participants

A total of 69 subjects replied to the driver questionnaire. Incomplete questionnaires (N = 22) were removed and not used for further analysis. Another 8 respondents no longer involved in active driving were also excluded from the analysis. In total, 39 complete response sets from the driver's side were taken into account for analysis. The responses were gathered between December 21, 2015 and February 17, 2016. The sample consisted of 2 female and 37 male drivers aged 22–74 (mean 49.34 ± 11.07). 30.56 % of the drivers operate on long haul transport, 41.67 % on regional transport and 22.22 % on local transport (local transport: up to 50 km; regional transport: 50–150 km; long haul transport: more than 150 km per day). On average, the subjects of the inspected sample had been operating as truck drivers for 27 years (± 11.96).

Concerning the transport company questionnaire, 17 responses were received of which 10 were completed. The companies employ a minimum of 3 and a maximum of 100 drivers. The results indicate that 60 % lease, 30 % rent and 80 % buy trucks. 80 % of the inspected companies stated that their trucks are in use for a range of 4–6 years, while 20 % indicated a range of 7–9 years.

4 Results

Descriptive statistics were calculated for each of the factors and items of the questionnaires. In addition, linear models (LM) were used to analyze the data. LM are especially appropriate for the estimation of item influence on the given influence of other variables, which is not possible with univariate testing methods. Therefore, LM were calculated between the averaged drivers' attitudes towards highly automated driving and the items regarding why a driver would want to let the system take over the driving task.

4.1 Descriptive Statistics

Resulting TA-EG (scale ranging from 1 = strongly disagree to 5 = strongly agree) scores of truck drivers (Fig. 1) are located around 3.16 (± 0.39), reflecting a moderate overall affinity of drivers towards technology. The scores varied between 3.39 (± 0.76) representing the subscale competence, and 2.99 (± 0.45) describing positive attitude towards technology in general. Intermediate scores for the subscales negative attitude (reversed, mean 3.28 ± 0.62) and enthusiasm (mean 3.01 ± 0.63) support the tendency of a rather medium attraction of truck drivers towards new technologies.

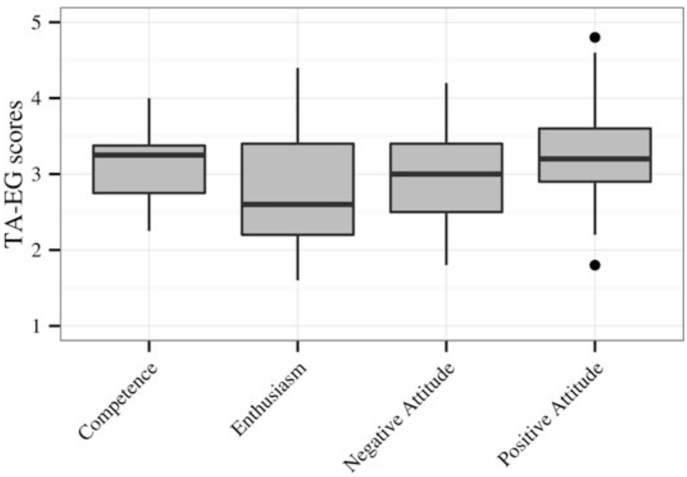


Fig. 1 TA-EG scores of truck drivers

Analyzing the initial question addressing the general comprehension of the concept and of highly automated driving systems, 43.6 % of truck drivers were not familiar with the terminology. Truck drivers showed themselves neutral towards the statement that the idea of highly automated driving is fascinating (mean 3.05 ± 1.19, on the scale from 1 = strongly disagree to 5 = strongly agree). Regarding driving pleasure, respondents neither assumed that highly automated driving will be enjoyable (mean 2.97 ± 0.96), nor did they strongly agree that highly automated driving would be easier than manual driving (mean 3.03 ± 0.90).

Furthermore, truck drivers fear losing their job due to being replaced by automation (25.6 %). Drivers adopted a neutral attitude towards the statement that driving in highly automated mode might be a potential relief in monotonous demanding driving situations (mean 3.31 ± 0.92). Although 53.8 % of all respondent drivers would hand over the driving task to the system when manual driving is monotonous, only 7.7 % would like to drive in highly automated mode in order to be able to engage in a secondary task.

51.3 % of truck drivers named higher safety as one of the reasons why they would like to drive in a highly automated truck (Fig. 2). Further insights into truck drivers' attitudes towards highly automated driving are provided by open questions. 47 % of the participants stated that the main reason for choosing their job was "driving pleasure". 17.65 % mentioned family backgrounds for instance "truck company is family owned" or "other family members have been truck drivers" or other forms of forced reasons. The following examples can be given: "unemployment" or "no alternatives". 8.8 % were tempted by the salary. Additional acceptance related statements are: "If the truck drives by itself, it should forgo the driver" or "automation will not mitigate the current situation for the drivers, it will replace them".

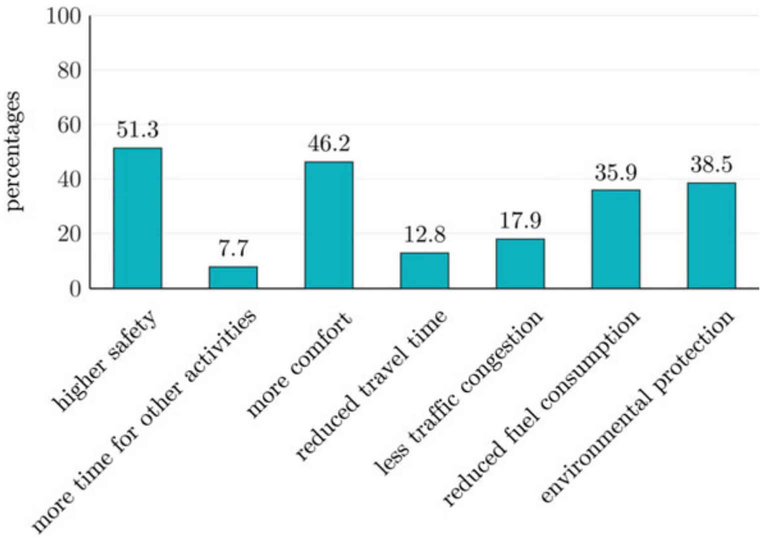


Fig. 2 Truck drivers' answers (in percentage) indicating why they would like to drive in highly automated mode

90 % of fleet managers have already heard of highly automated driving systems before. From the forwarding agencies' perspectives, the main reasons for choosing the profession "truck driver" nowadays are financial aspects instead of choosing a desired profession. Moreover, the companies mentioned doubts regarding safety issues, cybercrime and liability. The fact that drivers could pursue other occupational activities is evaluated positively by truck companies and may promote acceptance. Examples are vocational training and occupational retraining or editing documents. Also drivers can comply with and make optimal use of statutory rest periods.

Asking about concerns (Fig. 3), a widespread introduction of automated systems might bring up that both groups are notably worried about the fact that drivers might lose the joy of driving and the feeling of being in control (truck drivers: 46.2 %, fleet managers: 40 %). Truck drivers adopted a rather neutral attitude (mean 3.31 ± 1.00) towards the statement that highly automated driving might prevent severe accidents. 38.5 % of truck drivers (fleet managers: 60 %) indicated that they have substantial doubts about the safety and general reliability of such systems. Both truck drivers and fleet managers showed themselves most concerned about legal liability issues in the event of a highly automated truck being involved in an accident (truck drivers: 51.3 %, fleet managers: 60 %).

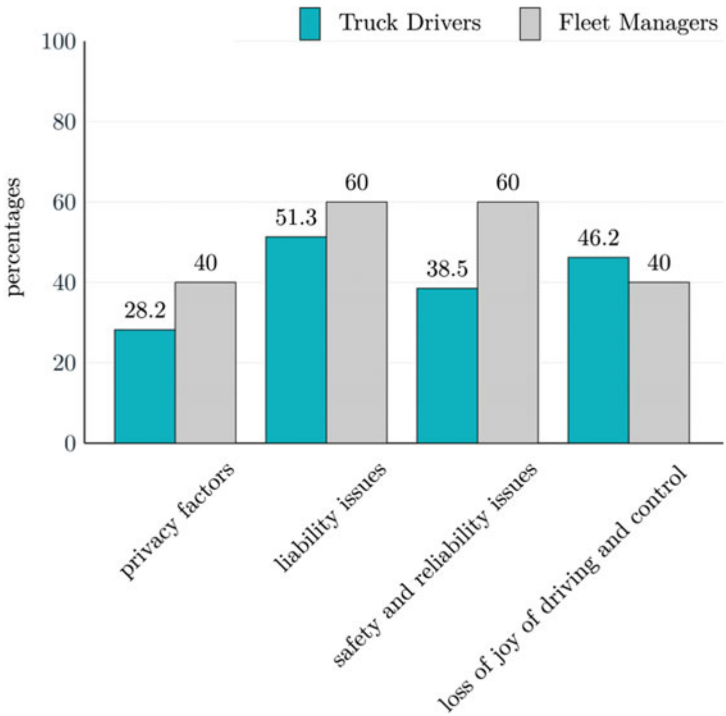


Fig. 3 Respondents concerns (in percentage) about the idea of introducing highly automated driving systems on a widespread scale

4.2 Linear Models

Linear models were run to determine the relationship between the averaged values of drivers' attitudes towards highly automated driving (response variables) and the values of every item concerning why a driver would want to let the system take over the driving task (explanatory variables) (Fig. 4). Linear models (α = 10 %) show that subjects thinking that highly automated driving is more comfortable have a significant higher attitude (mean +0.45, SD 0.19, p-value 0.0256) towards highly automated driving. Calculated effect size with a mean of +0.45 also indicates a strong as well as relevant effect, with the effect uncertainty (SD 0.19) being low in comparison to the other cases.

In addition, subjects thinking that highly automated driving is safer show a significant higher attitude regarding highly automated driving (mean +0.39, SD 0.2, p-value 0.0601). The calculated effect size also shows a strong, relevant effect with effect uncertainty again being low in comparison to other covariates.

Other comparisons between drivers' attitude and explanatory variables such as reduced traffic jams (mean +0.02, SD 0.25, p-value 0.927), reduced driving times (mean +0.06, SD 0.27, p-value 0.811), lower fuel consumption (mean 0.02, SD

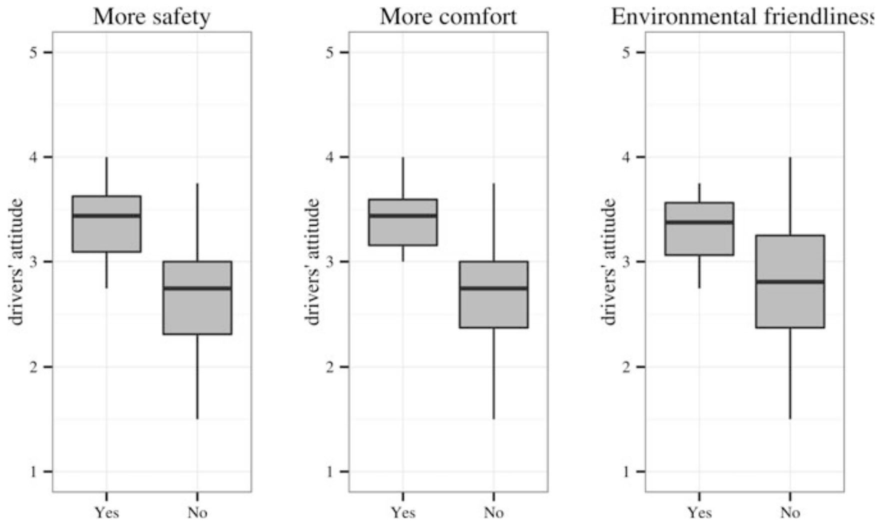


Fig. 4 Descriptive analysis showing drivers' opinion regarding highly automated driving and why a driver would want to let the system take over the driving task

0.24, p-value 0.9435) or a higher environmental friendliness (mean +0.13, SD 0.25, p-value 0.61) did not reveal significant results.

5 Discussion

This paper describes an approach aiming to assess truck drivers' attitudes towards highly automated driving. In order to set a baseline and derive implications for enhancing system acceptability on behalf of the driver, affinity for technology as well as concerns and needs towards system usage were inquired. Fleet managers' opinions regarding the potential and limitations of highly automated driving systems were queried in order to investigate whether transport companies would invest in this new technology.

Contrary to previous findings in the automotive domain indicating that acceptance towards automated systems is affected by the user's extent of technical affinity, could not be confirmed within this study [8]. Almost half of the respondent truck drivers (43.6 %) answered that they did not have a clear idea of the term "highly automated driving systems". This fact might lead back to truck drivers' average affinity towards new technologies, leading to problems regarding the imagination of driving in a highly automated mode. Furthermore, it must be considered whether using the TA-EG questionnaire is suitable for inquiring affinity for technology in terms of highly automated driving since the TA-EG was originally developed for electronical devices [9]. Considering future studies, an evaluation of the TA-EG must take place in order to determine whether it is suitable for measuring affinity for the technology of highly automated driving.

Recent studies have documented public opinion on automated driving technology. Most of these studies focused on measuring the public opinion of car drivers [4, 11, 14, 15]. Differences between car and truck drivers lie in the fact that most people buy cars for private reasons, whereas trucks are commercial vehicles serving a different purpose. Therefore, comparability between studies relating to car and truck drivers is limited.

Kyriakidis et al. [11] stated that car drivers find manual driving the most enjoyable mode of driving, yet they found the idea of highly automated driving fascinating. In contrast, truck drivers show a neutral attitude towards enthusiasm. Furthermore, car drivers were found to be most concerned about software hacking and misuse, but also about legal issues, similar to the findings by Schoettle and Sivak [15]. The results of this study confirm this fact for truck drivers as well as fleet managers. Additionally, open questions show that truck drivers are even more worried about reduced "driving pleasure" and the feeling of being replaced by automation systems.

Significant results between drivers' attitudes towards the concept of highly automated driving systems and benefits drivers could see in the implementation of such systems were found.

The more positive drivers' attitudes towards highly automated driving are, the more importance is attached to factors such as comfort and safety. Comfort and safety seem to have the biggest influence on the acceptance of highly automated driving. Most commercial haulage operators (90 %) have a clear picture of the concept of highly automated driving systems and suspect the major benefits in higher safety and reduced fuel consumption. In order to enhance efficiency, fleet managers could imagine a possible shift in truck drivers' work tasks and environment by providing them with secondary administrative tasks being performed during highly automated driving mode. The results show that fleet managers are able to put themselves into their employees' positions, as some doubts mentioned correspond to drivers' doubts. Therefore, it can be hypothesized that improvements in drivers' acceptability might also increase the acceptance level of forwarding agents.

Some limitations of this research should be noted. The sample size was small, but relevant effects were still found. However, a larger sample size would have ensured a more representative distribution of the population and would have been considered more representative of the group of people to whom results will be generalized or transferred. Hence, continuing studies should recruit a larger number of participants in order to draw a more representative conclusion in terms of population effects.

Online surveys offer advantages. According to this study, the main advantage was the ability of the internet to provide access to potential participants who would be difficult to reach through other channels [16]. Nevertheless, disadvantages are prevailing. As drivers are deterred by the topic of highly automated driving, recruitment was difficult. Invitations to participate in a survey on community bulletin boards, discussion groups, and chat rooms are often interpreted as rude or offensive [17], or considered as spam-mail [18]. Furthermore, self-selection bias is associated with online survey research [19]. One cannot assess how truthfully a respondent participates and how much consideration is put in. Further difficulties lie in the assessment of implicit information, such as feelings or emotions [20]. Another problem relates to the fact that respondents might not be aware of the context the questionnaire is set in. In this case, the concept of highly automated driving needs to be understandable to truck drivers and fleet managers. To foster an understanding of highly automated systems, suitable methods, other than describing, need to be applied and evaluated.

6 Conclusion

The results indicate different points of view regarding the introduction of highly automated driving. Truck drivers are most concerned about reduced driving pleasure and "being made redundant". Comfort and security have been indicated as promoting acceptance. Fleet managers see the positive effects in economic efficiency, profitability, safety, fuel consumption and possible secondary tasks. On the other hand, managers are worried about their drivers' attitudes towards the system. Therefore, fleet managers' acceptance might rise with increased acceptability on behalf of truck drivers. The least prominent areas of concern are software hacking/misuse and data transmitting issues. This explorative research establishes a foundation for formulating hypotheses and will lead to future studies. To further investigate whether driving simulator studies, enabling the driver to experience highly automated driving, will influence drivers' opinions regarding highly automated driving a driving simulator study is currently being conducted. Results will give further insights regarding possible changes. Also trust and acceptance towards the highly automated system are under investigation.

Acknowledgments Natalie Richardson was the initiator of the idea of this article and contributed to development of the questionnaire. Fabian Doubek, Kevin Kuhn and Annika Stumpf contributed to the development and to data collection. Markus Lienkamp revised the manuscript critically for important intellectual content. He gave final approval of the version to be published and agrees with all aspects of the work. The research was conducted with basic research funds of the Institute of Automotive Technology.

References

1. Ribeiro, S.K., Kobayashi, S., Beuthe, M., Gasca, J., Greene, D., Lee, D. S., Wit, R.: Transportation and its infrastructure. Institute of Transportation Studies (2007)
2. Gasser, T.M., Arzt, C., Ayoubi, M., Bartels, A., Bürkle, L., Eier, J., Lotz, C.: Rechtsfolgen zunehmender Fahrzeugautomatisierung, p. 83. Unterreihe Fahrzeugtechnik, Berichte der Bundesanstalt Für Strassenwesen (2012)

3. Anderson, J.M., Nidhi, K., Stanley, K.D., Sorensen, P., Samaras, C., Oluwatola, O.A.: Autonomous vehicle technology: a guide for policymakers. Rand Corporation (2014)
4. Howard, D., Dai, D.: Public perceptions of self-driving cars: the case of Berkeley, California. In: Paper presented at the 93rd Annual Meeting TRB, Washington, DC (2014) http://www.danielledai.com/academic/howard-dai-selfdrivingcars.pdf
5. Kauer-Franz, M., Franz, B., Abendroth, B.: Akzeptanz teilautomatisierter Fahrzeugführung. Erkenntnisse aus sieben Jahren Forschung. Institut für Arbeitswissenschaft, Technische Universität Darmstadt, Rollen, Aufgaben, Emotionen (2015)
6. Merat, N., Jamson, A.H., Lai, F.C., Carsten, O.: Highly automated driving, secondary task performance, and driver state. Hum. Factors: J. Hum. Factors Ergon. Soc. **54**(5), 762–771 (2012)
7. De Winter, J.C., Happee, R., Martens, M.H., Stanton, N.A.: Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. TransP. Res. Part F: Traffic Psychol. Behav. **27**, 196–217 (2014)
8. Davis Jr, F.D.: A technology acceptance model for empirically testing new end-user information systems: theory and results, Doctoral Dissertation, Massachusetts Institute of Technology (1986)
9. Karrer, K., Glaser, C., Clemens, C., Bruder, C.: Technikaffinität erfassen – der Fragebogen TA-EG. Der Mensch im Mittelpunkt technischer Systeme **8**, 196–201 (2009)
10. Hoogendoorn, R., van Arem, B., Happee, R., Espinoza, M.M., Kotiadis, D.: Towards safe and efficient driving through vehicle automation: the Dutch automated vehicle initiative (2013)
11. 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)
12. Schmitz, C.: LimeSurvey: an open source survey tool. LimeSurvey Project Hamburg, Germany (2012) http://www.limesurvey.org
13. IBM Corp. Released.: IBM SPSS statistics for windows, Version 22.0. Armonk, NY: IBM Corp (2013)
14. Casley, S.V., Jardim, A.S., Quartulli, A.M.: A study of public acceptance of autonomous cars (Bachelor of Science), Worcester Polytechnic Institute, Worcester, MA, USA (2013) http://www.wpi.edu/Pubs/E-project/Available/E-project-043013-155601/unrestricted/A_Study_of_Public_Acceptance_of_Autonomous_Cars.pdf
15. Schoettle, B., Sivak, M.: A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia (2014)
16. Garton, L., Haythornthwaite, C., Wellman, B.: Studying online social networks. In: Jones, S. (ed.) Doing Internet Research: Critical Issues and Methods for Examining the Net, pp. 75–105. Sage, Thousand Oaks, CA (1999)
17. Hudson, J.M., Bruckman, A.: "Go away": participant objections to being studied and the ethics of chatroom research. Inf. Soc. **20**(2), 127–139 (2004)
18. Andrews, D., Nonnecke, B., Preece, J.: Electronic survey methodology: a case study in reaching hard-to-involve internet users. Int. J. Hum Comput. Inter. **16**(2), 185–210 (2003)
19. Stanton, J.M.: An empirical assessment of data collection using the Internet. Pers. Psychol. **51**(3), 709–725 (1998)
20. Ackroyd, S.: Data collection in context. Longman Group United Kingdom (1992)

Will There Be New Communication Needs When Introducing Automated Vehicles to the Urban Context?

Victor Malmsten Lundgren, Azra Habibovic, Jonas Andersson, Tobias Lagström, Maria Nilsson, Anna Sirkka, Johan Fagerlönn, Rikard Fredriksson, Claes Edgren, Stas Krupenia and Dennis Saluäär

Abstract In today's encounters with vehicles, pedestrians are often dependent on cues in drivers' behavior such as eye contact, postures, and gestures. With an increased level of automation, and the transfer of control from the driver to the vehicle, the pedestrians cannot rely on such cues anymore. The question is: will there be new communication needs to warrant safe interactions with automated vehicles? This question is addressed by exploring pedestrians' willingness to cross the street and their emotional state in encounters with a seemingly automated vehicle. The results show that pedestrians' willingness to cross the street decrease with an inattentive driver. Eye contact with the driver on the other hand leads to calm interaction between vehicle and pedestrian. In conclusion, to sustain perceived

V.M. Lundgren · A. Habibovic · J. Andersson · T. Lagström · M. Nilsson (✉)
Viktoria Swedish ICT, Lindholmspiren 3A, 417 56 Gothenburg, Sweden
e-mail: maria.nilsson@viktoria.se

V.M. Lundgren
e-mail: victor.malmsten@viktoria.se

A. Habibovic
e-mail: azra.habibovic@viktoria.se

J. Andersson
e-mail: jonas.andersson@viktoria.se

T. Lagström
e-mail: tobias.lagstrom@viktoria.se

A. Sirkka · J. Fagerlönn
Interactive Institute Swedish ICT, Acusticum 4, 941 28 Piteå, Sweden
e-mail: anna.sirkka@tii.se

J. Fagerlönn
e-mail: johan.fagerlonn@tii.se

R. Fredriksson
Autoliv Research, 447 83 Vårgårda, Sweden
e-mail: Rikard.Fredriksson@autoliv.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_41

safety when eye contact is discarded due to vehicle automation, it could be beneficial to provide pedestrians with the corresponding information in some other way (e.g., by means of an external vehicle interface).

Keywords Pedestrian · Automated vehicle · Willingness to cross · Emotional state · Perceived safety · Wizard of Oz

1 Introduction

The research on automated vehicles (AVs) is accelerating for every day. Functions that support drivers in various traffic situations (e.g., Adaptive Cruise Control, Pedestrian Safety) are already on the market, and it is foreseen that this trend will remain. Companies from Volvo Cars to BMW and Google suggest that vehicles where the driving task is highly, or fully, automated will arrive soon [1]. Several states in the US, Europe, and Asia have already allowed testing of AVs in real traffic, and several companies are performing extensive testing there.

AVs are expected to bring multi-faceted benefits to the society including improved safety, reduced congestion, lower emissions, higher productivity, and improved mobility [2]. However, to obtain these benefits several challenges must be addressed, including the communication between AVs and pedestrians in their surroundings—a topic that has so far received limited attention in the research world [3].

Today, this communication occurs through interpretation of vehicle dynamics and other vehicle-centric cues using signs, lights and sound. Another key-part of the communication occurs through gestures, postures, and eye contact between drivers and pedestrians. Establishing eye contact serves as a confirmation that a driver and a pedestrian have noticed each other [4]. However, with the ever-increasing level of automation in vehicles, the driver role will change. The driver will not be representing the actions of the vehicle nor maneuvering it; he/she will be able to engage in other tasks such as reading and sleeping. With this transfer of control, there is a risk that pedestrians will not be able rely on these driver-centric cues anymore.

The question is: *Will there be new communication needs to ensure safe interactions when automated vehicles are introduced in the urban context*? This

C. Edgren
Volvo Car Corporation, 405 31 Gothenburg, Sweden
e-mail: claes.edgren@volvocars.com

S. Krupenia
Scania CV AB, 151 87 Södertälje, Sweden
e-mail: stas.krupenia@scania.com

D. Saluäär
Volvo Group, 405 08 Gothenburg, Sweden
e-mail: dennis.saluaar@volvo.com

question is addressed here by exploring pedestrians' perceived safety and how it is affected by different driver/vehicle behaviors. For this, two studies were conducted: a field experiment and a questionnaire study. These studies are described in the following sections along with the data analysis methods and the results.

2 Method

The field experiment was carried out to capture how pedestrians experience encounters with a manual vehicle (i.e. driver behaviors that are common today) and an automated vehicle (i.e. driver behaviors that may be common in the future) in a realistic traffic environment. The focus was on their (un)willingness to cross the street as an indicator of the perceived safety. Also, the reasoning behind their (un) willingness was explored as well as their emotional experience. A Wizard of Oz approach [5] was used to create encounters with the automated vehicle. To enlarge the sample size, the experiment was complemented by a questionnaire with a similar focus. The following sections describe the field experiment and the questionnaire in more detail.

2.1 Field Experiment

Participants. To participate in the study, the pedestrians needed to be familiar with the test location and frequently travelling by foot. In total, 13 pedestrians were recruited (7 male, 6 female) through direct contact at Chalmers University of Technology. They were in the age range of 20–29 years. A majority of them (N = 12) held a driver's license. About half of them (N = 7) did not report any eyesight correction.

Experiment Vehicle and Location. To create the Wizard of Oz set-up, a dummy steering wheel was installed in a right-hand steered vehicle (a Volvo V40) and the real steering wheel was hidden. This way, it appeared to be a standard left-hand steered vehicle seen from the perspective of the pedestrians (see Fig. 1). The fake driver on the left-hand side interacted with the pedestrians and seemingly drove the vehicle, while the person on the right-hand side was actually driving. The experiment was carried out at the Chalmers University of Technology on a dead-end street that is sporadically used to access the campus (see Fig. 2). There is a curb along one side of the street, but there is no zebra crossing.

Procedure. The test leader informed the pedestrian that he/she would interact with a vehicle, without mentioning that the vehicle would be automated. The pedestrian then filled in a background form (e.g., age, gender, driver's license, eyesight), and took a given position at the curb. To get familiar with the experiment, the pedestrian always experienced an introductory encounter where the (fake) driver tried to make eye contact with him/her. Next, the field experiment was executed

Fig. 1 The exterior (**a**) and the interior (**b**) of the experiment vehicle

Fig. 2 The experiment location

following the procedure described in Table 1 with three blocks of vehicle and driver behaviors (*A*, *B*, and *C*). The first two blocks were executed in a random order, while block *C* was always presented last to avoid excessive order effects. The driver behaviors illustrated in Fig. 3a-c were also presented in a random order within the blocks. After each encounter the pedestrian completed a Self-Assessment Manikin (SAM) questionnaire [6]. After finishing all encounters, the pedestrian participated in a semi-structured interview.

The total time of the experiment was approximately 30 min, depending on the length of the interview. The SAM scores were documented on paper, while the rest of the data were audio recorded and later transcribed for further analysis.

2.2 Questionnaire Study

Participants. The study involved 50 pedestrians who were selected randomly from public areas, mainly, when they were using public transportation. They were in the following age-range: 18–20 (8 %), 21–30 (40 %), 31–40 (30 %), 41–64 (16 %) and 65 and above (6 %). In total, 23 of 50 were women and 40 held a driver's license.

Pictures. The pictures were illustrating the same vehicle as in the field experiment (see Sect. 2.1). They were visualizing the driver behaviors (a)–(c) that were included in the field experiment as well as two other behaviors: (e) driver is looking

Table 1 The procedure for the field experiment study

Vehicle behavior	Driver behavior	Pedestrian task	Data
A: Motion Approaches the pedestrian on the near-side with a speed of ca 7 km/h without slowing down	Eye contact Phone Newspaper	1. Stands at the curb ca 0.5 m from the roadway at an "imaginary" zebra crossing and observes the approaching vehicle. Indicates when it feels uncomfortable to cross by turning towards the test leader 2. Completes the SAM-questionnaire 3. Answers the question: What is your decision to not cross the street based on?	Willingness to cross Motivation for the willingness SAM-questionn. Other comments
B: Standstill Approaches the pedestrian on the near-side with a speed of ca 7 km/h and then stops ca 3 m from the pedestrian	Eye contact Phone Newspaper	1. Stands at the curb ca 0.5 m from the roadway at an "imaginary" zebra crossing with the back towards the roadway. The vehicle is approaching and when it stops, the pedestrian gets the signal to turn around and answer the following question: Would you cross now? 2. Completes the SAM-questionnaire 3. Answers the question: What is your decision to cross/not cross the street based on?	Willingness to cross Motivation for the willingness SAM-questionn. Other comments
C: Same as block *B* and *A* (random order)	No driver	Same as pedestrian task in block *B* and *A* (random order)	Same as in block *B* and *A*

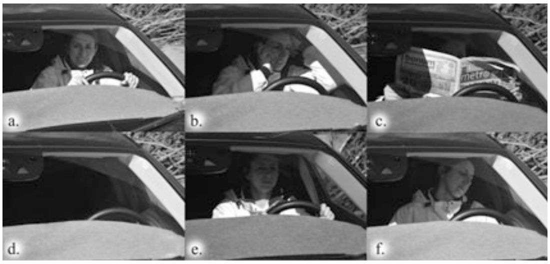


Fig. 3 The driver behaviors that the pedestrians experienced in the field experiment (**a**–**d**) and in the questionnaire (**a**–**c** and **e**–**f**)

straight ahead and (f) driver is sleeping (see Fig. 3). Behavior (e) was added to further investigate the influence of eye contact, while behavior (f) was used as a similar, but less futuristic, alternative to behavior (d). Each of the 5 pictures was evaluated by 10 pedestrians.

Procedure. A form containing background information of the study, description of the participant's role, SAM-questionnaire, and willingness to cross was given to each pedestrian. The test leader showed one of the pictures in Fig. 3 while describing the following scenario: *You are walking in a city center and you are about to cross the street at a non-signalized zebra crossing. A vehicle has just stopped and you encounter this* [shows picture] *driver behavior. Would you cross the street immediately?* The participants were asked to complete the SAM-questionnaire and encouraged to elaborate on their opinion.

2.3 Analysis Method

This section describes the analysis methods that were used to analyze the data collected in the field experiment and in the questionnaire study.

Analysis of Willingness to Cross the Street. The "would cross"/"would not cross" data were summarized and sorted by the corresponding driver behavior in order to compare the effects of these behaviors on the pedestrians' willingness to cross the street.

Analysis of Emotional State. SAM is a nonverbal assessment method that measures the valence, activity, and control associated with a person's affective reaction to stimuli [6]. Each of these dimensions is illustrated by five figures (see Fig. 4a). To identify the pedestrians' emotional state, the valence and the activity scores were combined using the Circumplex model of Affect [7]. According to this model, humans have a limited set of interrelated emotions (see Fig. 4b). The SAM scores can be plotted into this "affective space" consisting of the valence and activity dimensions [6]. As an example, an emotional reaction such as calmness is the combination of pleasantness and low activation (lower right quadrant).

Based on this, the average SAM scores were plotted using the valence dimension on the x-axis and the activation dimension on the y-axis. In addition, the scores were sorted in accordance with the driver and vehicle behaviors. The size of the circles in the plot visualizes the third dimension of SAM (control).

Analysis of Reasoning behind Willingness to Cross the Street. The recorded semi-structured interviews capturing the pedestrians' reasoning behind their (un) willingness to cross the street and their emotions were transcribed, and then imported into ATLAS.ti® qualitative research software. A qualitative assessment of the transcripts was performed in which a process of data reduction and "open coding" were performed [8]. The identified expressions were translated into English, and grouped to identify common themes and motives.

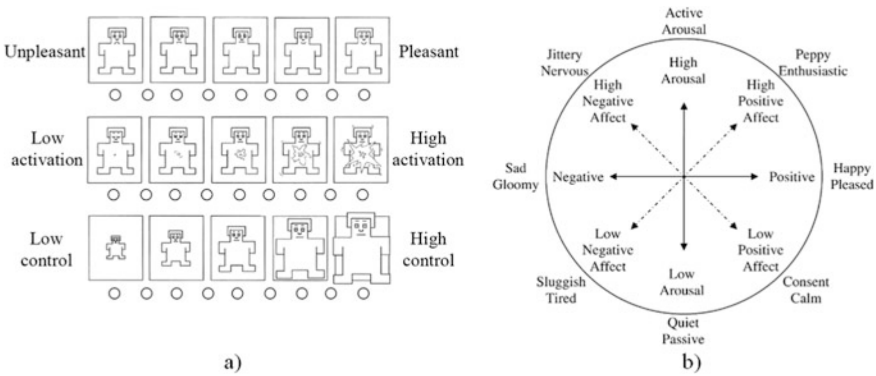


Fig. 4 The Self-Assessment Manikin (SAM) form with its three dimensions: valance, activation, and control (**a**), and the circumplex model of affect illustrating human emotions (**b**)

3 Results

This section presents results from the field experiment and the questionnaire in terms of the pedestrians' willingness to cross the street, the reasons behind their decisions, and their emotional experience. The familiarization encounter in the field experiment is omitted from the analysis.

3.1 Results Based on the Field Experiment

Willingness to Cross the Street. In the encounters with the standstill vehicle, all pedestrians (N = 13 of 13) stated that they would cross the street when they got eye contact with the driver0F[1] (see Fig. 5). When the driver was talking on the phone, the willingness to cross was reduced (N = 10 of 13). It was further reduced for the encounters where the driver was reading a newspaper (N = 5 of 13), and when there was no driver present in the vehicle (N = 5 of 13).

Reasoning behind the Willingness to Cross the Street. As shown in Tables 2 and 3, the pedestrians' decisions to cross, or not to cross, the street were motivated by both vehicle-centric and driver-centric cues. The importance of these cues varied depending on the vehicle speed and the activity that the driver was involved in.

In the encounters with the standstill vehicle, the fact that the vehicle is not in motion, and that it is standing at a rather large distance, was taken into account by several pedestrians who stated that they would feel safe to cross the street. However, these cues were less frequently mentioned when the pedestrians could get eye contact with the driver. Seven of the pedestrians who got eye contact did not

[1]The *driver* refers to the fake driver (i.e. not the person who was in fact driving the vehicle).

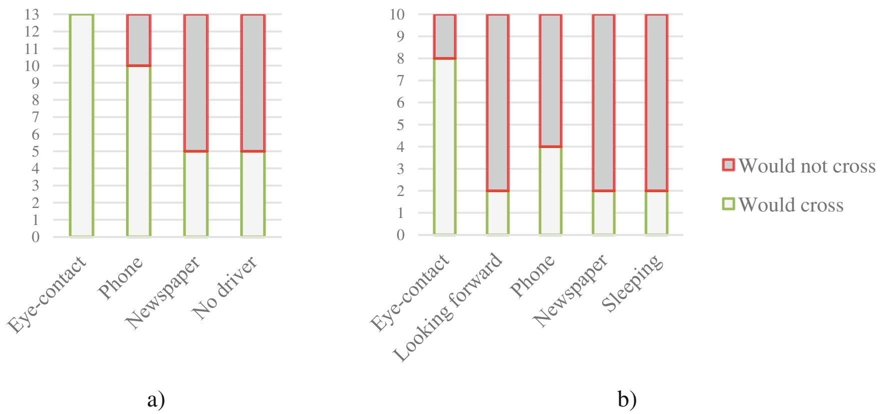


Fig. 5 Willingness to cross based on the field experiment (**a**) and the questionnaire (**b**)

Table 2 Reasons for not crossing the street when encountering a vehicle in motion

Driver behavior	Reason for not crossing the street (# of statements)
Eye contact	Not slowing down (1) Short distance, not slowing down (3) Short distance, could not see what the driver is doing (1) Short distance, it felt safe for a long while due to (eye) contact and low speed (7)
Phone	Talking on the phone, looking aside, no sign that I am noticed (3) No eye contact, talking on the phone, looking aside (1) Inattentive, talking on the phone (3) Short distance, talking on the phone (2) Not slowing down although I've been waiting for a while (1) The driver could suddenly accelerate (1)
Newspaper	Reading newspaper, could not trust the driver (4) Not slowing down, reading newspaper, inattentive, cannot have noticed me (4) Reading newspaper, cannot have noticed me (2) Reading newspaper, inattentive, would not be able to stop (1) The driver could suddenly accelerate (1) Short distance (1)
No driver	No driver, no indication that I am noticed (3) No driver (4) No driver, could not anticipate changes in the speed (1) No driver, unclear who is in control of the vehicle (1) Short distance, no driver (1) Trusted the vehicle until it was too late to cross (1) Difficult to trust the vehicle (1)

Table 3 Reasons for (not) crossing the street when encountering a standstill vehicle

Driver behavior	Reason for crossing the street (# of statements)	Reason for not crossing the street (# of statements)
Eye contact	Eye contact, showing that I am noticed (4) Eye contact, showing that I am noticed, large distance (3) Eye contact, standstill (4) Large distance, standstill (1)	
Phone	Large distance, enough time to cross (2) Large distance, unsafe as the driver is talking on the phone (3) Standstill, the driver is talking while waiting that I'll cross (3)	Talking on the phone, inattentive, could accelerate while I am crossing (1) No eye contact, talking on the phone, looking aside, unclear if I am noticed (1) No eye contact, talking on the phone, unclear if the driver stopped for me (1)
Newspaper	Standstill, the driver is reading while waiting that I'll cross (1) Standstill, the engine sound does not indicate any acceleration (1) Standstill, the driver is reading and will not accelerate (1) Standstill (1)	No eye contact, reading newspaper (1) Reading newspaper, inattentive, cannot have noticed me (3) Reading newspaper, inattentive (3) Short distance (1)
No driver	It is a zebra crossing, the vehicle stopped to let me cross (1) Standstill, large distance (1) Standstill, the passenger trusts the vehicle (1) Standstill, if it is a driverless vehicle it will let me cross (1) The driver left the vehicle, and it cannot accelerate (1)	No driver, short distance (1) No driver (3) No driver, no sign what will happen (1) No driver, couldn't get contact with the passenger (1) Unclear who is in control (1)

mention any vehicle-centric cues when motivating their decisions. In addition, the pedestrians who stated that they would not cross the street, motivated this by the lack of eye contact with the driver.

In the encounters with the vehicle in motion, the pedestrians commonly motivated their choice to not cross the street by the fact that the vehicle came too close without slowing down. Another frequent motivation was the lack of the contact with the driver and that the driver was inattentive (i.e. focusing on non-driving related tasks). In the encounters where the pedestrians got eye contact with the driver, they stated that they felt safe for a longer while compared with other encounters. The importance of the driver-centric cues was also emphasized in the encounters where there was no driver; several of the pedestrians stated that they would not cross the street since there is no driver to give them confirmation that they have been noticed.

Emotional Experience. The analysis of the average SAM-scores for the encounters with the vehicle in motion where the pedestrians got eye contact with the driver shows a fairly pleasant experience with low activation level, resulting in a

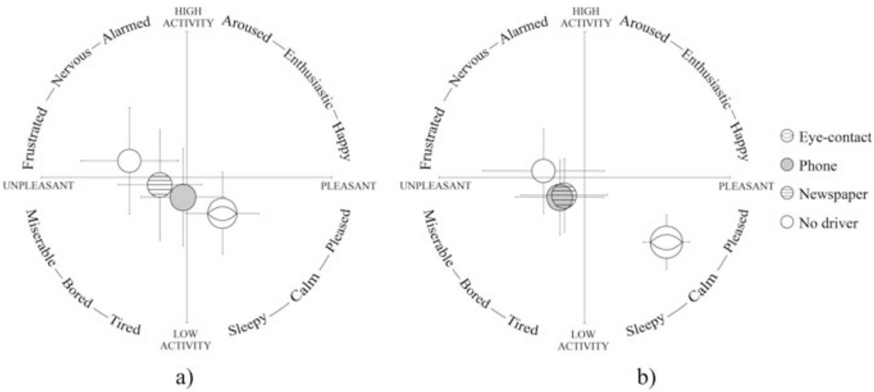


Fig. 6 Average SAM-scores of the pedestrians' experiences in the encounters with the vehicle in motion (**a**) and the standstill vehicle (**b**). The size of the score shows the level of control

calm emotional state (lower right quadrant in Fig. 6a). The phone and newspaper encounters generated a slightly unpleasant experience. The encounters with no driver resulted in an unpleasant experience with a higher activation level than in the other encounters (i.e. the pedestrians felt somewhat frustrated). Overall, the level of control is similar for all encounters, though it is somewhat higher for the encounters with eye contact (see the size of the data points in Fig. 6). Based on the interviews, this could be explained by the fact that the pedestrians experienced that the vehicle speed was low at the same time as they were standing on the curb that is regarded as a safe zone. Also, some pedestrians stated that they felt in control in all encounters since they could decide when to cross and when to not cross the street.

The emotional differences between the encounters are in general in line with the pedestrians' statements in the interviews. All pedestrians (N = 13 of 13) stated that (eye) contact with the driver, and the driver behavior in general, made the greatest difference in their experience. When asked about the safest encounter, all pedestrians (N = 13) stated they felt most safe when they got eye contact with the driver. Only 3 of 13 pedestrians mentioned explicitly that the encounters were experienced differently due to the vehicle speed. However, the importance of the vehicle speed for the emotional experience is demonstrated by the fact that several pedestrians (N = 7) stated that the safest encounter was when the vehicle was in standstill.

Furthermore, the interviews also gave some explanation for the unpleasant experience in the encounters without a driver. To start with, the pedestrians had no previous experience with self-driving vehicles, and encountering a vehicle without a driver was a surprising and unfamiliar event. When asked about the most unsafe encounter, more than half of the pedestrians (N = 8 of 13) referred to the one that did not involve any driver.

All pedestrians experienced the encounters without any driver as involving a self-driving vehicle. The encounters where the driver was talking on the phone, or reading newspaper, were in general not experienced as involving a self-driving vehicle. However, more than half of the pedestrians (N = 8 of 13) stated that even if they knew in advance it was a self-driving vehicle involved in these encounters they would expect to get some confirmation from the person behind the wheel (i.e. the fake driver).

3.2 Results from the Questionnaire

Willingness to Cross the Street. The results from the questionnaire show that a majority of the pedestrians (N = 8 of 10) would cross the street when encountering a vehicle where they get eye contact with the driver. On the other hand, relatively few of the pedestrians (N = 2 of 10) would feel comfortable to cross when the driver is looking forward, reading newspaper, or sleeping. Similarly, only 4 of 10 pedestrians would cross when the driver is involved in a phone conversation.

Reasoning behind the Willingness to Cross the Street. Several of the pedestrians highlighted the importance of receiving feedback from the driver, and motivated their unwillingness to cross by the lack of such feedback. The encounters where the driver is looking straight ahead, or talking on the phone, were described as common in the traffic, but unsafe since the driver is not showing that he/she is paying attention to pedestrians in the vicinity.

Emotional Experiences. The analysis of the average SAM-scores shows that the encounters with the driver giving eye contact are viewed as very pleasant with low level of activation (i.e. the pedestrians felt calm, see lower right quadrant in Fig. 7). The phone and looking forward encounters generated slightly more unpleasant experiences. In addition, the encounters where the driver was reading newspaper, or sleeping, were experienced as unpleasant and with a higher level of activation (i.e. the pedestrians felt somewhat frustrated). Overall, the average level of control was similar in all encounters, though it was somewhat higher for the eye contact encounters.

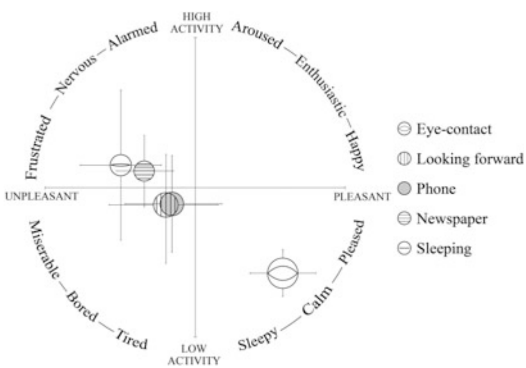


Fig. 7 Average SAM-scores of the pedestrians' experiences in the encounters with a standstill vehicle. The size of each score shows the level of control

4 Discussion

This study explored the influence of current and future driver behaviors on pedestrians' emotional state and willingness to cross the street as an indicator of their perceived safety. Our results from the field experiment and questionnaire showed similar results, i.e. the driver's attention has a clear impact on the willingness to cross the street. The driver's attention also affects perceived safety and emotional state of the pedestrians. However, we could also see that an inattentive driver does not create an equally strong negative emotional state, but there is a tendency towards frustrated and nervous emotional states for such encounters. Our results also indicate that pedestrians may expect to get confirmation from the person behind the steering wheel even if the vehicle is in automated mode. Together, these findings suggest that the communication needs of pedestrians will change when eye contact with the vehicle driver is discarded due to use of vehicle automation.

The interviews also revealed that the motion pattern of the vehicle was an important cue (cf. in 16 of 29 encounters the pedestrians reported that their decision to cross the street was affected by the fact that "the vehicle is standstill"). Today, pedestrians have the ability to interpret a vehicle's intentions based on its motion patterns, e.g., if a vehicle is slowing down, it is generally interpreted that the vehicle is yielding. These interpretations can also be made in darkness where the pedestrians cannot see the driver's face or gestures. Thus, if automated vehicles adopt the motion patterns of human drivers about to give way to pedestrians this could provide sufficient cues for a safe interaction. The motion pattern of the vehicle itself could be seen as an interface towards the pedestrian [9], telling whether it will give way or not. However, an interface that presents the vehicle's intention before it changes its speed, or position, could have a benefit in terms of judging what an automated vehicle will do next. Also, it could increase the perceived safety of the pedestrians.

On a final note, our study is based on a few traffic situations in a Swedish traffic context, and further studies are needed to draw conclusions about other traffic situations (e.g., crowded crossings, turning maneuvers) and other cultural settings.

5 Conclusion

- The study provides an understanding of how pedestrians experience and react to encounters involving manually maneuvered vehicles, and how they may experience and react to future encounters involving automated vehicles.
- The findings indicate that the communicative needs may change due to use of vehicle automation. Pedestrians' willingness to cross the street decreased when they encountered driver behaviors that are not common today, while eye contact with the driver resulted in a calm interaction. That is, pedestrians' perceived

safety might decrease when the driver is not representing the actions of the vehicle.

- To sustain a high level of perceived safety in such situations, it could be beneficial to provide the pedestrians with the corresponding information in some other way (e.g., by means of an external vehicle interface).
- The study is limited to a relatively small sample of Swedish pedestrians who are not used to automated vehicles yet, and further investigations are needed to generalize the findings.

Acknowledgments The study is carried out in cooperation between Viktoria Swedish ICT, Interactive Institute Swedish ICT, Autoliv, SAFER, Scania, Volvo Cars, and Volvo Group. The study is partly conducted within a Master's thesis project at Chalmers University of Technology. The study is co-financed by the Strategic Vehicle Research and Innovation, (FFI), the Knowledge Foundation (KK-stiftelsen), and Swedish ICT. The authors are grateful to the participants in the experiments.

References

1. Volvo Car Group: Volvo cars stakes its claim in the premium sedan segment with the long-awaited S90 (2015)
2. Fagnant, D., Kockelman, K.K.: Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations. pp. 1–20
3. Berk, B.: How Nissan's Using Anthropology to Make Autonomous Cars Safe. The Drive (2015)
4. Šucha, M.: Road users' strategies and communication: driver-pedestrian interaction. In: Transport Research Arena (TRA) 2014 Proceedings (2014)
5. Molin, L.: Wizard-of-Oz prototyping for co-operative interaction design of graphical user interfaces. In: The Third Nordic Conference on Human-Computer interaction (NordiCHI, vol. 82), pp. 425–428 (2014)
6. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry **25**(1), 49–59 (1994)
7. Russell, J.A.: A circumplex model of affect. J. Pers. Soc. Psychol. **39**(6), 1161–1178 (1980)
8. Strauss, A., Corbin, J.: Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Sage Publications, Inc., Thousand Oaks, California (1998)
9. Fast Company, http://www.fastcodesign.com/3054330/innovation-by-design/the-secret-ux-issues-that-will-make-or-break-autonomous-cars

Command-Based Driving for Tactical Control of Highly Automated Vehicles

Azra Habibovic, Jonas Andersson, Jan Nilsson, Maria Nilsson and Claes Edgren

Abstract As vehicles become highly automated, their drivers become more passive. A concern is it may take drivers out of the control loop, causing reduced satisfaction and perceived control. The study explores whether or not drivers feel the need to control tactical decisions when operating highly automated vehicles. An experiment involving 17 drivers was carried out in a driving simulator. Each driver tested two different tactical controllers, allowing him/her to give various tactical commands to the vehicle (e.g., overtake, park). The results indicate that the drivers experienced a need to affect tactical decisions of highly automated vehicles. Several of the tactical commands were found useful, especially on rural roads and highways. It also gave them a feeling of being in control of the vehicle, suggesting that command-based driving might be a way to keep drivers in the control loop.

Keywords Tactical control · Command-based driving · Highly automated vehicle · Feeling of control · Satisfaction · User experience · Wizard of Oz

A. Habibovic · J. Andersson · M. Nilsson (✉)
Viktoria Swedish ICT, Lindholmspiren 3A, 417 56 Gothenburg, Sweden
e-mail: maria.nilsson@viktoria.se

A. Habibovic
e-mail: azra.habibovic@viktoria.se

J. Andersson
e-mail: jonas.andersson@viktoria.se

J. Nilsson
Semcon Sweden AB, 417 80 Gothenburg, Sweden
e-mail: jan.nilsson@semcon.com

C. Edgren
Volvo Car Corporation, 405 31 Gothenburg, Sweden
e-mail: claes.edgren@volvocars.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_42

1 Introduction

With an ever-increasing level of automation in vehicles, the drivers' role will transfer from acting an active driver to being a passive passenger. Along the increased degree of automation, a gradual transition from manual operation toward supervision of the automated vehicle takes part [1]. An apparent risk is decreased situational awareness, described as "out-of-the-loop performance" by Endsley and Kiris [2]. Consequently, drivers may experience loss of satisfaction and of control.

The major premise in this study is that allowing drivers to affect tactical decisions of their highly automated vehicles may increase satisfaction and feeling of control. According to Michon [3], driving can be divided into three levels of control: (a) strategic, (b) tactical, and (c) operational. At the strategic level, driver decisions concern long term planning and overall goals of a trip. The tactical level describes controlled action patterns at a shorter timespan, such as maneuvers of the vehicle, while the operational level represents automatic (skill-based) action patterns in the millisecond timespan. When a highly automated vehicle is operated in automated mode, it takes ownership of the tactical control along with the strategic control (cf. NHTSA's automation level 3 [4]). Introducing command-based driving that allows drivers to affect tactical decisions in such situations could be a way to sustain the feeling of being in control.

The study explores whether or not drivers feel the need to control tactical decisions when operating highly automated vehicles, and in that case, which tactical decisions are important to control, and under which conditions they are needed. The topic is addressed by conducting an experiment in a driving simulator by means of a Wizard of Oz approach [5]. The drivers give the tactical commands by using a controller interface. Each driver tested two of three available controllers. However, to avoid possible order effects, this paper focuses on their experience after using the first controller.

2 Method

The study was carried out in a fixed-base driving simulator by applying a Wizard of Oz approach to make the test drivers believe that they were experiencing a highly automated vehicle. The test driver issued the following commands by using a controller interface, and the automation (a Wizard of Oz driver) performed them in a timely manner in sync with the surrounding traffic situation: (a) change lane to the left/right, (b) overtake, (c) take the next exit, (d) park, and (e) turn left/right.

2.1 Tactical Controllers

Three master thesis project teams were assigned the task of designing and developing their own prototypes for tactical control of highly automated vehicles, based on the concept ideas drawn from the preceding discussions and previous experience. The results, in form of *Controller A–C*, were then used in the simulator experiment.

2.2 Controller A

The *Controller A* is based on a tablet-PC application. The prototype is built upon Android, but the underlying idea is to allow cross-platform availability on nomadic devices in any comfortable position for the driver. In the experiment, the tablet was mounted in a stationary position in front of the center stack.

The commands are given through touch gestures on the tablet's touchscreen, supported by visual cues. The feedback to the driver is communicated visually. The top of the screen contains general journey information and the vehicle status, while the bottom of the screen allows for tactical input within the circle (Fig. 1). The circle contains a vehicle icon at a road intersection. A command is given by placing a finger upon the vehicle icon, and dragging it to one of the possible circular alternatives on the road. Once a command is given, the vehicle icon remains in the same place until the command has been executed. The driver can easily abort the command by dragging the vehicle icon out of the active circle. However, if the execution of a given command has started, abortion is prohibited due to safety reasons.

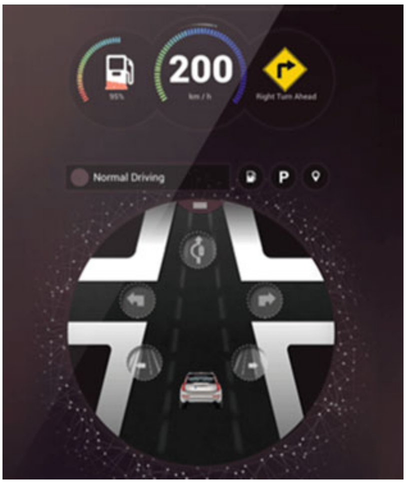


Fig. 1 *Controller A* where tactical commands are given by drag-and-drop interaction on the tablet's touchscreen. The feedback to the driver is visual

Apart from the tactical commands, *Controller A* also allows access to infotainment, climate, and other vehicle functionality. This part of the prototype is only represented by a placeholder to demonstrate the concept.

2.3 Controller B

Controller B uses gestural input on a touch surface. The prototype was built using the same hardware as *Controller A* and was placed horizontally on the center console between the seats. To give commands, the driver swipes the pre-defined single-finger gestures on the input surface (Fig. 2, left). Once the command is given, a corresponding icon is displayed in the instrument cluster (Fig. 2, right), and a sound is played, providing feedback to the driver. The feedback icons resemble the shapes of the gestures in order to facilitate learnability and recognition. If the system fails to identify the given gesture, an error message is displayed and played, encouraging the user to try again. A specific gesture is provided in order to abort a confirmed command, and abortion is possible until the command has started its execution.

The conceptual model for *Controller B* is based on analog gestures [6] resembling the movement of the vehicle. The traditional lack of signifiers for gestural interfaces [7] is mitigated by only allowing for a small amount of tactical commands. Visual gesture hinting on the input surface was tested during the design process, but was excluded from the final prototype since it failed to assist the drivers.

2.4 Controller C

In *Controller C* the gear stick is used for input of commands. Autonomous mode in a vehicle with a traditional combustion engine powertrain, requires automatic gear shifting. Furthermore, the gearstick is unoccupied while driving autonomously, which allows the hardware to be used for other purposes. The main idea behind *Controller C* is therefore to utilize the gearstick, using the conceptual model of a joystick. Input is given by moving the stick using different gesture patterns (Fig. 3).

After being issued, the command is validated by the vehicle, and either accepted or rejected. If a command is accepted, the gearstick provides haptic and auditory

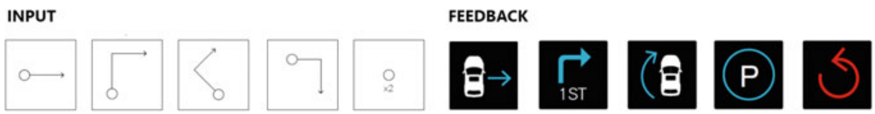


Fig. 2 Finger gestures (*left*) and visual feedback in the instrument cluster (*right*) in *Controller B*

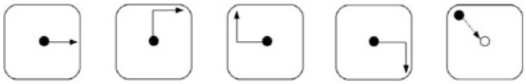


Fig. 3 The patterns for giving different commands using *Controller C*

feedback, and stays in the given position until the command is executed. If a command is rejected, haptic and auditory feedback is provided while the stick moves back to the original position in the center. In order to abort a given command, the driver simply moves the joystick back to the original position, unless the execution already is in progress, in which case abortion is prohibited.

2.5 Driving Simulator and the Wizard of Oz Approach

The experiment was carried out in one of Volvo Car Corporation's driving simulators, an advanced fixed-base simulator (Fig. 4). The traffic environment incorporated a typical Swedish highway (ca 6 km) and a rural road (ca 3 km), as well as two small cities (ca 3 km). Depending on the environment, the speed limits varied between 50 km/h (city), 70 km/h (rural road) and 90 km/h (highway).

To create a highly automated vehicle experience, a Wizard of Oz approach was used. That is, a human Wizard of Oz-driver, who used an additional steering wheel and pedals to control the vehicle, provided the self-driving functionality. The feedback from the controller interfaces was also managed using a Wizard of Oz approach. When the test driver's tactical command was registered, an assistant provided appropriate feedback in a timely manner.

2.6 Participants

In total, 17 drivers (9 male, 8 female) participated in the experiment. They were recruited via e-mail contacts. The criteria for participating in the study were that

Fig. 4 The driving simulator and the Wizard of Oz driver who acted as the vehicle automation

they were holding a driver's license, and that they were not directly involved in research and development of automated vehicles. The age and gender distribution was equal across the three different interfaces. Furthermore, a great majority of the drivers stated that they are driving a passenger vehicle on daily basis.

2.7 Procedure

Each participant evaluated two of the three controller concepts (*Controller A–C*). The pairwise experimental setup was a compromise to extend the length of each test, yet avoid a prolonged test procedure for each participant. The experiment procedure was divided into seven steps depicted in Fig. 5.

- *Step* 1: *Introduction to the study and simulator.* Initially, the test drivers were asked to complete a background questionnaire and then they were informed of the purpose of the study and asked for written consent.
- *Steps* 2 *and* 5: *Introduction to the controllers.* After the initial test drive, the drivers learned how to use the tactical controller that was to be tested first.
- *Steps* 3 *and* 6: *Test drives.* Both test drive 1 and 2 were carried out on the same road (ca 12 km), and each of them took approximately 12 min to complete. To achieve a comparable evaluation of the controllers, the test drivers were prompted by a text message to give a specific command to the vehicle.
- *Steps* 4 *and* 7: *Debriefing and evaluation.* After each test drive, the participants completed two questionnaires and a set of interview questions. The interview was semi-structured and a video recording of the test drive was also used to make it easier to discuss different events in the drive.

In this paper we present the results from steps 1–4 (i.e. the first test drive). In these steps, *Controller A* was tested by 7 drivers, *Controller B* was tested by 5 drivers, and *Controller C* was tested by 5 drivers.

2.8 Data

The data about the test drivers' demographics (age, gender, etc.) and their experiences and opinions regarding driver assistance systems and automated driving

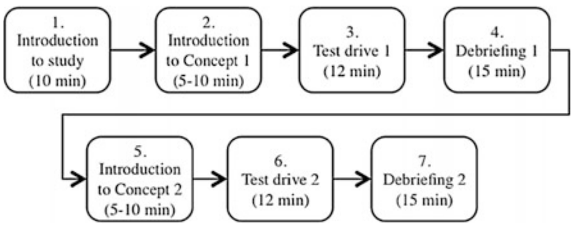


Fig. 5 The experiment procedure

were collected prior to any test drives by means of a background questionnaire. To assess the test drivers' experiences after each test drive, they were asked to complete two questionnaires and to participate in an interview. The first one was based on the User Experience Questionnaire (UEQ) [8], which probed the drivers' user experience. The other one was based on a somewhat modified Technology Acceptance Model (TAM) [9]. The TAM questionnaire consisted of 11 statements where the drivers rated their assent/dissent using a Likert-scale. The semi-structured interview probed the drivers' experience of travelling in an automated vehicle, how they experienced issuing the tactical commands, and if the concepts fulfilled their expectations.

As this paper does not focus on the design of the controllers, the results are based on the background questionnaire, interviews, and the TAM-questionnaire.

3 Results

3.1 Attitudes Towards Automated Driving

The analysis of the background questionnaire shows that the majority of the drivers had experienced at least one type of the driver assistance systems prior to the experiment. Approximately 88 % of them stated that they had experienced ACC (Automated Cruise Control). Also, a great majority of the drivers stated that they find driver assistance systems useful, and that they would like to have a vehicle that can brake and steer itself (see Fig. 6). Furthermore, a great majority stated that they would like to have the opportunity to affect decisions that such a vehicle makes.

These findings are in line with the drivers' opinions expressed in the interviews following the first test drive (i.e. after testing the first of two controllers). The analysis of the interviews shows that a great majority of the drivers (N = 13 of 17) were in general positive about the automated driving. They felt it was innovative and exciting, and that they could trust the vehicle although this feeling would be amplified with longer experience. However, a few drivers found it difficult to

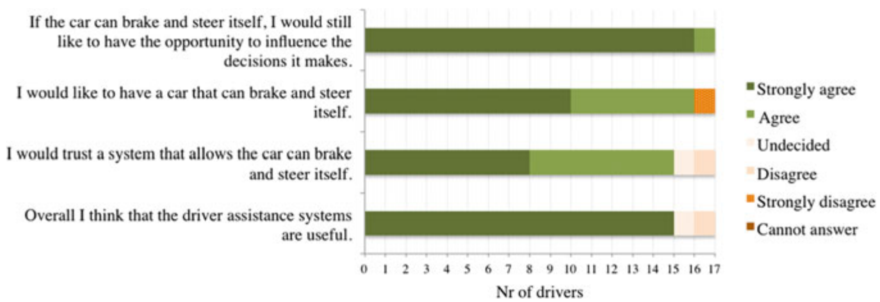


Fig. 6 Drivers' preferences prior to the experiment

imagine how the automated driving, and the tactical control, could work in real traffic. Also, two of the drivers stated that it was somewhat scary in the beginning to hand over control to the vehicle. As such, the drivers may need time to get used to highly automated driving. One of the drivers had a negative attitude towards the automated driving even before starting the experiment, and she could not find anything favorable in the name of automated driving after the experiment either. Two drivers felt a bit dizzy, but it could not be established if it was due to the simulator itself, or the automated driving.

3.2 The Need for Tactical Control

The data analysis reveals that drivers feel a need to affect tactical decisions of their vehicles while driving in highly automated mode. All drivers, except the one who had a negative attitude toward automated driving in general, stated in the interview that at least one of the tactical commands was useful (Fig. 8). Several of them (N = 7 of 17) found all commands useful. They found the available tactical commands as a bare minimum in order to feel in control of the vehicle while in the automated mode. This is also demonstrated in the TAM-questionnaire where all drivers stated that the tested controller contains functions that they would like to have (see the "usefulness" dimension in Fig. 7 where the mean values for the different controllers are in the range 6–6.8 on the Likert scale 1–7). In addition, the drivers would use the given controller in their everyday driving if they had access to a highly automated vehicle (see the "intention to use" dimension in Fig. 7 where the mean values for the different controllers are in the range 6–7). These two ratings were similar for all controllers; however, the drivers who used *Controller B* were slightly more convinced.

The results from the TAM-questionnaire imply also that having the ability to affect tactical decisions gives drivers a feeling of control when the vehicle is in the automated mode. More specifically, the mean value for the "control" dimension on

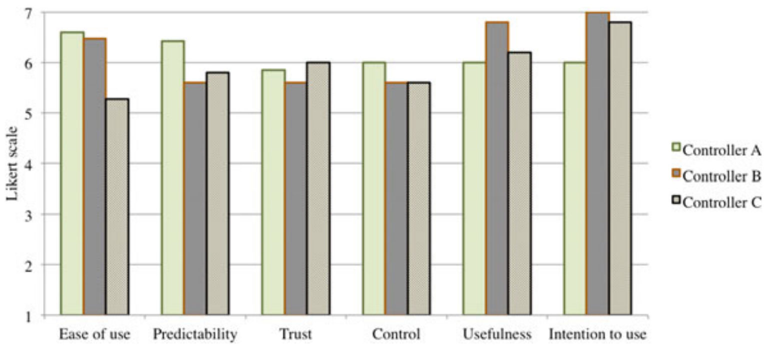


Fig. 7 Mean values of the TAM-questionnaire for *Controller A–C*

the Likert scale was in the range 5.5–6 for the different controllers, suggesting that the drivers experienced a rather high level of control irrespectively of the interface type (Fig. 7). However, it should be noted that *Controller A* gave a slightly higher feeling of being in control than the two other controllers.

When asked in the TAM-questionnaire if they could trust the controller and if it behaved in a predictable way, the drivers answered positively on both statements (see the "trust" and "predictability" dimensions in Fig. 7 where the mean values for the different controllers are in the range 5.7–6.4 on the Likert scale 1–7). The interviews clarified that many of them were in fact referring to both the automation and the controller. Again, this suggest that the drivers trusted the automation in general, but also that all controllers were in line with their expectations (i.e. they were satisfied). In turn, one could argue that drivers expect to have the ability to affect tactical decisions of highly automated vehicles. Out of the three controllers, *Controller A* behaved in the most predictable manner.

The TAM-questionnaire shows also that all controllers were easy to use, which could have contributed to the positive attitude towards the tactical control in general (see the "ease of use" dimension in Fig. 7 where the mean values for the different controllers are in the range 5.5–6.6 on the Likert scale 1–7). The lowest rate was obtained for *Controller C*. Based on the interviews, this could be explained by the fact that the drivers did not obtain enough feedback from it. The importance of feedback and "dialog" between the highly automated vehicle and the driver is also justified by the fact that all drivers highlighted that they would like to be informed by the vehicle when a deviation in the original plan occurs and to have ability to affect the new plan.

Furthermore, the interview data shows that overtaking and/or turning commands were experienced as most useful (Fig. 8). Having the ability to instruct the vehicle that an overtaking is desirable would be useful in many situations. Different drivers have different preferences when it comes to overtaking, and from that point of view it is essential to have the ability to affect such decisions. At the same time, overtaking is viewed as a demanding and safety-critical maneuver, and the drivers believed that the automation would be better than them at performing such maneuvers. The turning commands were viewed as practical when a change in the driving plan occurs. Interestingly, some of the drivers found the turning commands useful when driving in a familiar environment, while some others stated that these

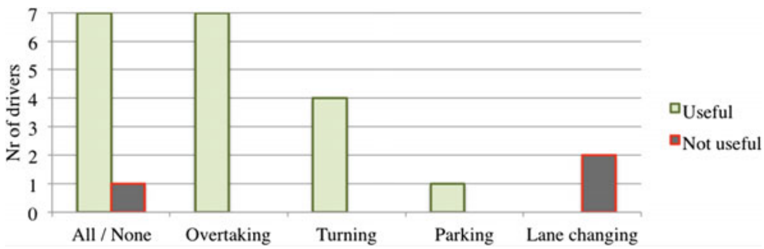


Fig. 8 Most useful and least useful tactical commands

commands would be useful when driving in an unfamiliar environment. For two of the drivers the lane-changing commands were not essential, and they suggested omitting these commands. These conclusions are independent of the controller that was tested.

On a final note, several drivers stated that they would like to have ability to affect more long-term decisions of their highly automated vehicles (i.e., on the strategic level). Examples of such commands were “take second/third exit”, “stop at the next pick-nick area”, and “find an available parking lot in a given area”.

3.3 Suitability of Tactical Control

When asked about the applicability of the tactical commands with the regard to traffic environment, several drivers (N = 9 of 17) found them as the most appropriate for rural roads and/or highways. Some of the drivers (N = 3 of 17) believed that the tactical commands are appropriate for all traffic environments. Interestingly, four drivers stated that tactical commands are inappropriate for city driving. This was motivated by the fact that the urban traffic is dynamic requiring quick decisions, and that you either let the automation drive to a given destination without affecting its tactical decisions, or you disengage it and drive manually.

4 Discussion

Overall, the drivers had a positive attitude toward vehicle automation and could imagine using such vehicles in daily driving. Also, tactical commands were in general viewed as beneficial. We could identify an overall satisfaction of having the ability to affect tactical decisions of automated vehicles, independently of the controller used. In fact, a great majority of the drivers highlighted that this functionality would be essential to gaining acceptance of highly automated vehicles. The drivers were also positive about the idea of using the given controller in their daily driving, which also shows the importance of being able to affect the tactical decisions.

The drivers experienced that the use of tactical controllers gave them a feeling of being in control of the vehicle when operated in automated mode. This indicates that allowing this functionality could be a way to keep drivers in the control loop. Several previous studies suggest that highly automated vehicles should be designed in a way that keeps drivers in the control loop and enables them to safely re-engage in manual driving ([10]). One direction for future research could be to explore the relationship between the ability to affect tactical decisions and the ability to resume control when the automation reaches its performance limits.

The future research should also investigate the difference in drivers’ experience of automation with and without access to tactical controllers. We had a few “control events" in the experiment where the automation changed the original plan without leaving the drivers possibility to affect the new plan. Basically, all drivers highlighted these events and stated that they would have preferred to be involved in the decision-making. Again, this could be interpreted that there is a need for tactical control; however, a deeper investigation should be carried out.

The study is based on a small sample size of Swedish drivers and the generalization of the results may be difficult. However, it can be seen as a first step towards investigating the role of command-based driving. The findings need to be verified in a more realistic setting, and with a larger sample size. The large-scale evaluations of highly automated vehicles that are on the way (e.g., Drive Me project [11]) could be used for such purposes.

5 Conclusions

- Enabling drivers to affect tactical decisions of highly automated vehicles has the potential to improve satisfaction and give drivers feeling of being in control. This is especially applicable for highways and rural roads where relatively few tactical commands are required.
- The tactical commands that were explored in the study were viewed as useful and essential, especially the overtaking and turning commands.
- The need for feedback from automation and a dialog with the driver when making important decisions (e.g., re-routing) is confirmed.
- The drivers expressed a need for extending the functionality to include strategic control (e.g., take third exit).
- All three controllers performed equally well in the evaluation; however, a more detailed analysis of other experiment data is recommended.

Acknowledgments The study is carried out in cooperation between Viktoria Swedish ICT, Semcon Sweden AB, and Volvo Car Corporation as a part of the project AIMMIT (supported by the Strategic Vehicle Research and Innovation (FFI). Thanks go to the master's thesis students for developing the controllers and to the participants in the experiment.

References

1. Damböck, D., Kienle, M., Bengler, K., Bubb, H.: The H-Metaphor as an example for cooperative vehicle driving. Lect. Notes Comput. Sci. **6763**, 376–385 (2011)
2. Endsley, M.R., Kiris, E.O.: The out-of-the-loop performance problem and level of control in automation. J. Hum. Factors Ergon. Soc. **37**, 381–394 (1995)
3. Michon, J.A.: A critical view of driver behavior models: what do we know, what should we do? Hum. Behav. Traffic Saf. 485–452 (1985)
4. NHTSA: Preliminary Statement of Policy Concerning Automated Vehicles. (2013)

5. Molin, L.: Wizard-of-Oz prototyping for co-operative interaction design of graphical user interfaces. In: The Third Nordic Conference on Human-Computer interaction (NordiCHI, vol. 82). pp. 425–428. ACM Press, Tampere, Finland (2014)
6. Zhai, S., Kristensson, P.O., Appert, C., Anderson, T.H., Cao, X.: Foundational issues in touch-surface stroke gesture design—an integrative review. Found. Trends Hum Comput. Interact. **5**, 97–205 (2011)
7. Norman, D.A., Nielsen, J.: Gestural interfaces: a step backward in usability. Interactions **17**, 46–49 (2010)
8. Rauschenberger, M., Schrepp, M., Perez-Cota, M., Olschner, S., Thomaschewski, J.: Efficient measurement of the user experience of interactive products. how to use the user experience questionnaire (UEQ). Int. J. Interact. Multimed. Artif. Intell. **2**, 39–45 (2013)
9. Davis, F.: User acceptance of information technology system characteristics user perception and behaviroul impacts. Int. J. Man—Machine Stud. **38**, 475–485 (1993)
10. Merat, N., Lee, J.: Preface to the special section on human factors and automation in vehicles: designing highly automated vehicles with the driver in mind. Hum. Factors J. Hum. Factors Ergon. Soc. **54**, 681–686 (2012)
11. Volvo Car Group: Volvo Car Group initiates world unique Swedish pilot project with self-driving cars on public roads, https://www.media.volvocars.com

Driving Simulator Experiment on Ride Comfort Improvement and Low Back Pain Prevention of Autonomous Car Occupants

Junya Tatsuno and Setsuo Maeda

Abstract Unlike other autonomous working vehicles such as farm and construction machinery, autonomous cars have at least an occupant, even if the driving operations are automated. Thus, the whole-body vibration problem continues to be a research topic with respect to autonomous vehicles. Several technologies have been developed to decrease whole-body vibration exposure. Most of technologies have focused on the development and improvement of automotive components such as the suspension system, automotive seats and tires. This paper examines the reduction in whole-body vibrations using autonomous functions such as lane-change control. In this paper, a subject experiment with a driving simulator is reported so as to discuss the influence of decreasing the whole-body vibration exposure on ride comfort improvement and low back pain prevention.

Keywords Whole-body vibration · ISO 2631-1 · Autonomous vehicle · Lane change · Pothole

1 Introduction

Many technologies have been developed to decrease whole-body vibration exposure during automobile movements. Research on suspension control can be considered a representative example of improving whole-body vibration exposures. Conventional suspension systems use passive springs and dampers to control spring

J. Tatsuno (✉)
Faculty of Engineering, Kindai University, 1 Takaya-Umenobe, Higashi-Hiroshima City, Hiroshima 739-2116, Japan
e-mail: tatsuno@hiro.kindai.ac.jp

S. Maeda
Faculty of Applied Sociology, Kindai University, 3-4-1 Kowakae, Higashi-Osaka City, Osaka 577-8502, Japan

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_43

motion and absorb impact. The spring and damper coefficients are fixed to satisfy the concept of each automobile model. However, the performance of conventional suspensions has a definite limit. Therefore, automotive companies have developed active suspensions [1–3]. In fact, some models on which active suspensions were installed were commercialized by automotive companies. Currently, only a few active suspensions models exist because of high costs. Nevertheless, there is continued research on active suspension system. Most active suspension research is focused on construction and improving the control method [4–6]. In contrast, the psychological effects of active suspension on vehicle occupants were not considered. Similarly, many studies report on the impact of seat design in improving whole-body vibration exposure. To date, automotive seats were mainly evaluated using the SEAT value, which is defined in ISO 10326-1:1992 [7]. The SEAT value is the transmissibility; it is calculated from the acceleration on the seat and the acceleration at the base of the seat. Because the SEAT value is derived from data obtained in laboratory experiments, several problems have been pointed out. For example, Yu and Khameneh [8] reported that subjects were unable to distinguish between three seats with different transmissibility characteristics. The methods for measuring and evaluating the whole-body vibration of vehicles are defined in ISO2631-1 [9]. The degree of comfort/discomfort is defined using the magnitude of physical vibrations. Several investigating studies were also conducted using the measurements defined in ISO2631-1 to evaluate whole-body vibration in automobiles [10–12]. As described above, various studies focused on the measurement and evaluation of whole-body vibrations on human drivers in a conventional automobile. In addition, most of these studies examined agricultural and construction vehicles [13, 14].

Let us reconsider the characteristics of autonomous vehicles. The automation of agricultural and construction vehicles most likely implies unmanned systems. However, in the case of autonomous cars, even though the driving operations are automated, vehicles still have a few occupants. Therefore, the whole-body vibration problem in autonomous vehicles continues to be a research topic. However, few studies have examined the effect of whole-body vibration on ride comfort or health effect in autonomous vehicles. This paper proposes that the autonomous vehicles could mitigate whole-body vibrations by utilizing autonomous behavior such as a lane-change function. Previous researches were conducted on the lane-change control for autonomous cars [15, 16]. In addition, a lane-changing function appears feasible given the advance in driver-assistance systems with information and communication technology (ICT).

Conventional methods achieve a reduction in whole-body vibration by improving elemental technology in the automobile. In contrast, the approach in this paper utilized the intelligent behavior of the autonomous car to attain a reduction in whole-body vibration. Generally, there is a relationship between whole-body vibrations in a traveling vehicle and the perception of ride comfort by the occupant. Moreover, it is well known that individuals driving an automobile occupationally are exposed to whole-body vibrations in the course of their jobs. These occupational whole-body vibrations often result in physical ailments such as low back pain. Therefore, it is important for the occupational driver to mitigate whole-body vibrations for low back pain prevention.

Previous research [17] performed a numerical simulation to verify the effectiveness of lane change control in decreasing the whole-body vibration exposure of occupants in autonomous cars. The paper used numerical simulations with experimental data obtained through ISO2631-1 measurement techniques. The findings revealed that if an autonomous car could skirt rough roads by the lane-change function, the degree of ride comfort could be enhanced during travel. Furthermore, the maximum acceptable exposure time could be lengthened, and this could impact health. This paper extends the work done by the previous study, and performs a subject experiment with a driving simulator to investigate the influence of decreasing the whole-body vibration exposure on ride comfort improvement and low back pain prevention.

2 Material and Method

2.1 Driving Simulator

Figure 1 shows the front and rear views of the driving simulator. Flight simulation technologies from the aerospace division of Fuji Heavy Industries and automotive technologies from Subaru automobiles were merged to develop the body of the simulator. The simulator body was mounted on a Stewart platform with six prismatic joints driven by a DC motor. This allowed the driving simulator to be moved in six degrees of freedom (6 DOF), that is, three linear movements (x, y, and z) and three rotations (pitch, roll, and yaw). The UC-win/Road Ver.5 software developed by Forum8 Co., Ltd was used to create the virtual reality environment. A three-dimensional virtual reality space could be easily created using this software.

(a) Front view (b) Rear view

Fig. 1 Photos of the driving simulator

2.2 Course Layout

Figure 2 shows the layout of the course produced in the driving simulator. In Condition 1, three bumpy sections per km were fixed on the road. In Condition 2, the number of bumpy sections per km was one. This implied that the vehicle could virtually avoid two bumpy sections in Condition 2. This experimental condition was chosen because it was considered that complete obstacle avoidance may be difficult given factors such as variable traffic flow conditions around the car, resolution of the road surface map, and lane-change control accuracy. Nevertheless, the magnitude of the whole-body vibration that the subject was exposed to decreased in Condition 2. The bumpy section consisted of 10-speed bumps, and there was no speed bump on the normal section. It was reported that varying the height of the speed bumps could control the magnitude of the vibration to which the subjects were exposed [18]. In the experiment, the height of the speed bump was configured as 5 cm given the findings of the previous study. The magnitude of the whole-body vibrations that the subject was exposed to on the bumpy section could be estimated approximately 0.8 m/s^2. However, the weighted root-mean-square

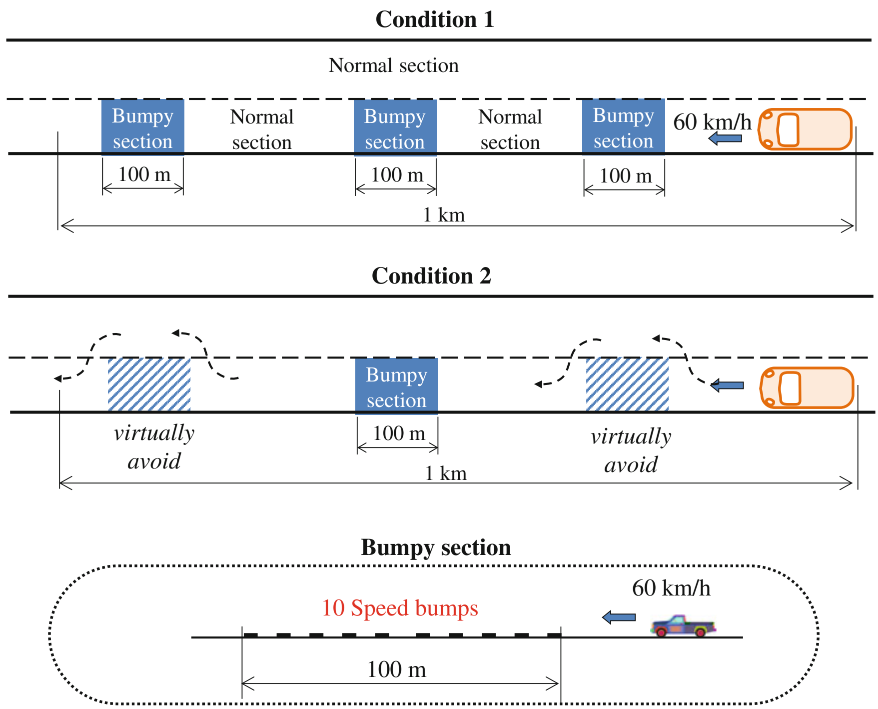


Fig. 2 Schematics diagrams of the course layout in Condition 1, Condition2, and the bumpy section

(rms) accelerations on the seat surface were measured on the basis of ISO2631-1. In this measurement, an accelerometer (Brüel & Kjær, Type 4515-B) and a vibration meter (Rion, VM-54) were used.

2.3 Participants and Task

The participants were asked to ride the driving simulator that traveled in both Condition 1 and Condition 2. Because the driving simulator had no automatic driving mode function, the experimenter operated the driving simulator with the external controller. The traveling speed and the traveling duration were 60 km/h and 30 min, respectively. Participants were given a reading task instead of a driving task.

In addition, the participants were required to rate their level of discomfort every 5 min. The subjective discomfort evaluation that consisted of the overall evaluation was adopted as shown in Fig. 3. This evaluation method was also used in the previous study [19]. The overall evaluation was the discomfort scale developed based on the basis of the Borg scale. The participants rated their overall discomfort by choosing a number between 0 and 100. Conversely, the local evaluation was used to indicate the discomfort at 5 body regions, namely upper back, lower back,

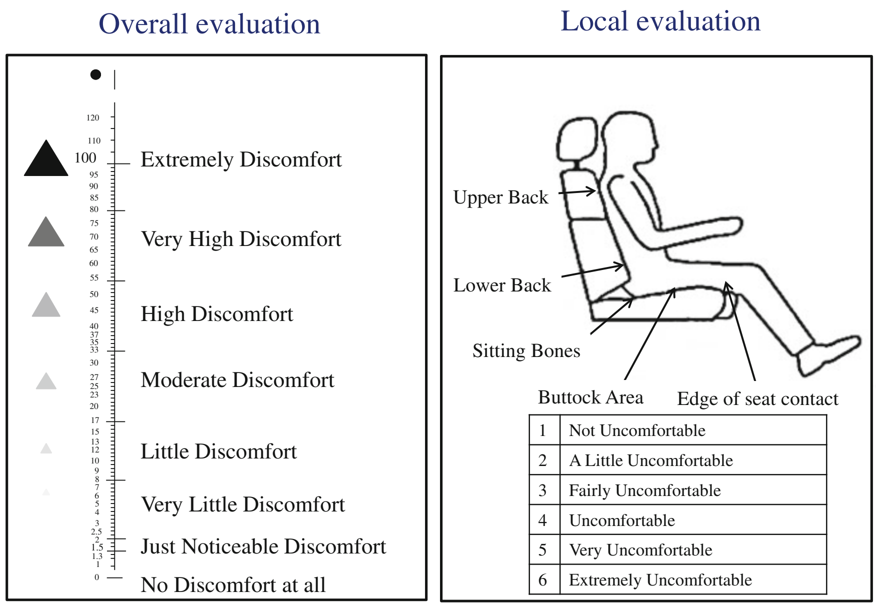


Fig. 3 The *left* part and the *right* part show the subjective discomfort scale of the overall evaluation and that of the local evaluation respectively

sitting bone, buttock area and edge of seat contact. The participants answered all local discomfort rating questions by selecting a number between 1 and 6.

The trial participants included 4 male students (mean age 21.5 ± 0.5 years) 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.

3 Results and Discussion

3.1 Objective Evaluation

The frequency-weighted acceleration of each translational axis (a_{wx}, a_{wy} and a_{wz}) measured on both the bumpy section and the normal section were summarized in Table 1. The frequency-weighting curves of W_d, W_d and W_k were used to calculate a_{wx}, a_{wy} and a_{wz} according to ISO2631-1 [9], respectively. The vibration total value of weighted rms acceleration, determined from the vibration in orthogonal coordinates, was calculated as follows:

$$a_v = (k_x^2 a_{wx}^2 + k_y^2 a_{wy}^2 + k_z^2 a_{wz}^2)^{\frac{1}{2}} \tag{1}$$

Here, the multiplying factors kx, ky and kz in the case of the ride comfort evaluation are defined in ISO2631-1 as follows:

$$[\, k_x \quad k_y \quad k_z \,] = [\, 1.0 \quad 1.0 \quad 1.0 \,] \tag{2}$$

As shown in Fig. 2, the length ratios of the bumpy section and the normal section of Condition 1 and Condition 2 are 3:7 and 1:9, respectively. We could estimate the vibration total value that the participants were exposed to at each condition as shown in Fig. 4. The graph indicated that there was a clear difference

Table 1 Frequency-weighted acceleration of each translational axis measured on both the bumpy section and the normal section. Unit is m/s^2

Participant	Section	a_{wx}	a_{wy}	a_{wz}	a_v
P1	Bumpy	0.3720	0.0701	0.6500	0.7522
	Normal	0.0166	0.0160	0.0320	0.0394
P2	Bumpy	0.4029	0.1049	0.5157	0.6628
	Normal	0.0183	0.0198	0.0259	0.0374
P3	Bumpy	0.4106	0.6560	0.2990	0.8297
	Normal	0.0108	0.0160	0.0280	0.0340
P4	Bumpy	0.4012	0.0572	0.2940	0.5007
	Normal	0.0106	0.0122	0.0183	0.0244

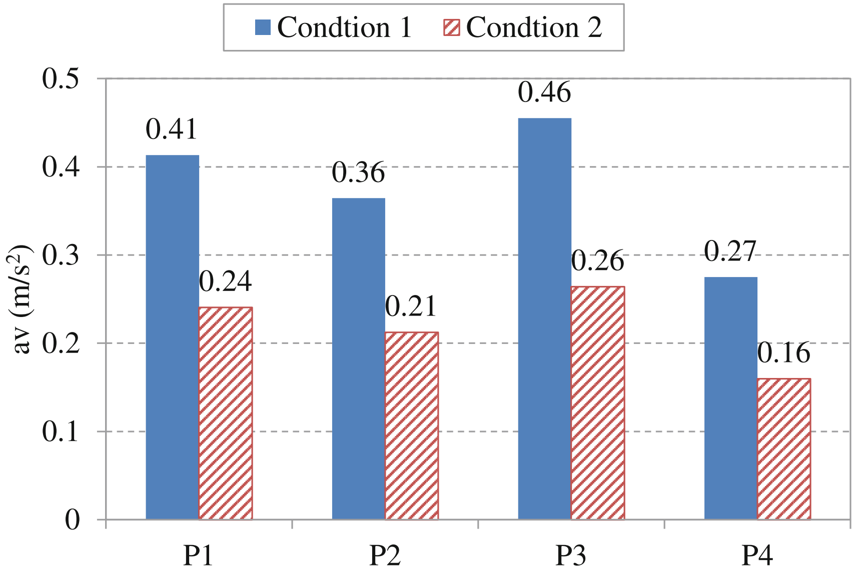


Fig. 4 A bar diagram of the vibration total value that participants were exposed to. The *solid fill* and the *hatch fill* areas indicate Condition 1 and Condition 2, respectively

between the vibration total values in Condition 1 and the vibration total values in Condition 2. The ISO 2631-1 also provided likely comfort reactions to vibration that occurred in public transport (Table 2). The usage of this scale resulted in the vibration total values of P1, P2 and P3 in Condition 1 that were classified as "A little uncomfortable". In contrast, the values of all participants in Condition 2 were rated as "Not uncomfortable". That is the rating for all participants in Condition 2 were a grade lower. This indicated that the avoiding the bumpy section was effective in improving the ride comfort of the occupants.

This paragraph discussed the experimental results from the health effect evaluation viewpoint. It is well known that individuals who drive an automobile occupationally are exposed to whole-body vibrations in the course of their work. These occupational whole-body vibrations often cause diseases such as low back pain.

Table 2 Acceleration levels and discomfort defined by ISO2631-1:1997

Acceleration (m/s^2)	Discomfort
Less than 0.315	Not Uncomfortable
0.315–0.63	A little uncomfortable
0.5–1	Fairly uncomfortable
0.8–1.6	Uncomfortable
1.25–2.5	Very uncomfortable
Greater than 2	Extremely uncomfortable

In the future, when the autonomous car is driven, the travel time will increase, because the absence of the driving task will result in less fatigue during travel. Therefore, it is important to consider the health effects of the whole-body vibrations in autonomous cars. Moreover, from the prevention of occupational diseases caused by whole-body vibration viewpoint, the daily exposure value A(8) was defined as the 8-hour energy equivalent vibration total value for workers in meters per second squared (m/s^2). This A(8) value is calculated as the highest rms value of weighted acceleration among the determined values of the three directional components, taking into the following coefficients.

$$[\,k_x \quad k_y \quad k_z\,] = [\,1.4 \quad 1.4 \quad 1.0\,] \tag{3}$$

In addition, according to "Control of Vibration at Work Regulations" [20], the exposure limit values and action values of the whole-body vibration are as follows:

(a) the daily exposure limit value is 1.15 m/s^2
(b) the daily exposure action value is 0.5 m/s^2

Thus, occupational drivers should not be exposed to a daily exposure that exceeds 1.15 m/s^2. In addition, a risk control action is necessary when the daily exposure exceeds 0.5 m/s^2. Furthermore, the maximum acceptable exposure time $T_{A(8)}$ is often used for the evaluation of the health effect of the whole-body vibrations. This may be expressed as follows:

$$T_{A(8)} = \frac{A(8)^2 \cdot 8\,\mathrm{h}}{a_v^2} \tag{4}$$

The $T_{A(8)}$ values of all participants in both conditions were derived as shown in Table 3. The value of $A(8)$ was the daily exposure action value of 0.5 m/s^2. This result indicated that the maximum acceptable exposure time $T_{A(8)}$ greatly varied between the conditions. Specifically, the $T_{A(8)}$ values of Condition 1 were approximately one-third of the $T_{A(8)}$ values of Condition 2. For example, if the car did not avoid the rough surface, the participants P3 would travel for only 8 h per

Table 3 Maximum acceptable exposure time $T_{A(8)}$

Participant	Condition	$k_x\ a_{wx}$ (m/s^2)	$k_y\ a_{wy}$ (m/s^2)	$k_z\ a_{wz}$ (m/s^2)	$T_{A(8)}$ (h)
P1	1	0.2859	0.0569	*0.3570*	15.7
	2	0.1662	0.0376	*0.2078*	46.3
P2	1	*0.3097*	0.0837	0.2833	20.9
	2	*0.1800*	0.0534	0.1649	61.7
P3	1	0.3151	*0.5034*	0.1654	7.9
	2	0.1823	*0.2912*	0.0982	23.6
P4	1	*0.3079*	0.0461	0.1618	21.1
	2	*0.1782*	0.0301	0.0946	63.0

Italic values are the highest rms value of the three directional components

Table 4 Responses of the pariticipants to subjective evaluation under Condition 1 and Condition 2

Participant	Time (min)	Condition 1						Condition 2					
		Upper Back	Lower back	Sitting bones	Buttock area	Edge of seat contact	Overall	Upper Back	Lower back	Sitting bones	Buttock area	Edge of seat contact	Overall
P1	0	3	1	2	1	1	15	1	1	1	1	1	5
	5	2	2	3	2	2	15	1	1	1	1	1	10
	10	2	3	3	3	2	20	1	2	2	1	1	15
	15	3	4	4	4	3	23	2	3	3	3	2	25
	20	3	3	4	4	4	25	2	3	4	3	3	30
	25	3	4	5	5	4	30	3	4	4	4	4	30
	30	4	5	5	5	5	40	3	4	4	4	4	40
P2	0	1	1	1	1	1	0	1	1	1	1	1	0
	5	1	1	1	2	1	3	1	1	1	1	1	2
	10	1	1	2	2	1	3	1	1	2	1	1	6
	15	1	1	3	3	1	10	2	2	3	2	2	15
	20	2	2	4	3	3	30	2	3	4	3	2	23
	25	2	2	4	3	3	40	3	3	4	3	3	30
	30	3	3	5	4	3	60	3	3	4	3	3	40
P3	0	1	1	1	1	1	0	1	1	1	1	1	0
	5	3	3	4	3	2	45	2	2	2	1	1	1
	10	4	4	4	2	2	50	3	3	2	2	2	10
	15	4	3	3	2	2	50	3	3	2	2	2	10
	20	3	4	3	2	2	50	4	3	3	2	2	30
	25	4	5	4	2	2	70	4	4	3	2	2	45
	30	4	4	3	2	2	50	3	4	3	2	2	40

(continued)

Table 4 (continued)

Participant	Time (min)	Condition 1						Condition 2					
		Upper Back	Lower back	Sitting bones	Buttock area	Edge of seat contact	Overall	Upper Back	Lower back	Sitting bones	Buttock area	Edge of seat contact	Overall
P4	0	1	1	1	1	1	1	1	1	1	1	1	1.5
	5	1	2	2	1	1	10	2	2	2	1	1	10
	10	2	2	2	1	1	25	2	2	2	2	1	15
	15	2	4	3	2	1	30	3	3	3	2	1	20
	20	3	5	3	2	2	40	3	4	4	2	2	30
	25	3	6	4	3	2	65	3	4	3	3	3	40
	30	4	6	4	3	3	75	4	6	4	3	3	65

day. In contrast, when the car avoided the rough surface (for example in Condition 2), P3 could travel all day long. Hence, intelligent behaviors such as lane-change control were highly effective in the preventing of low back pain in the occupants of the vehicle.

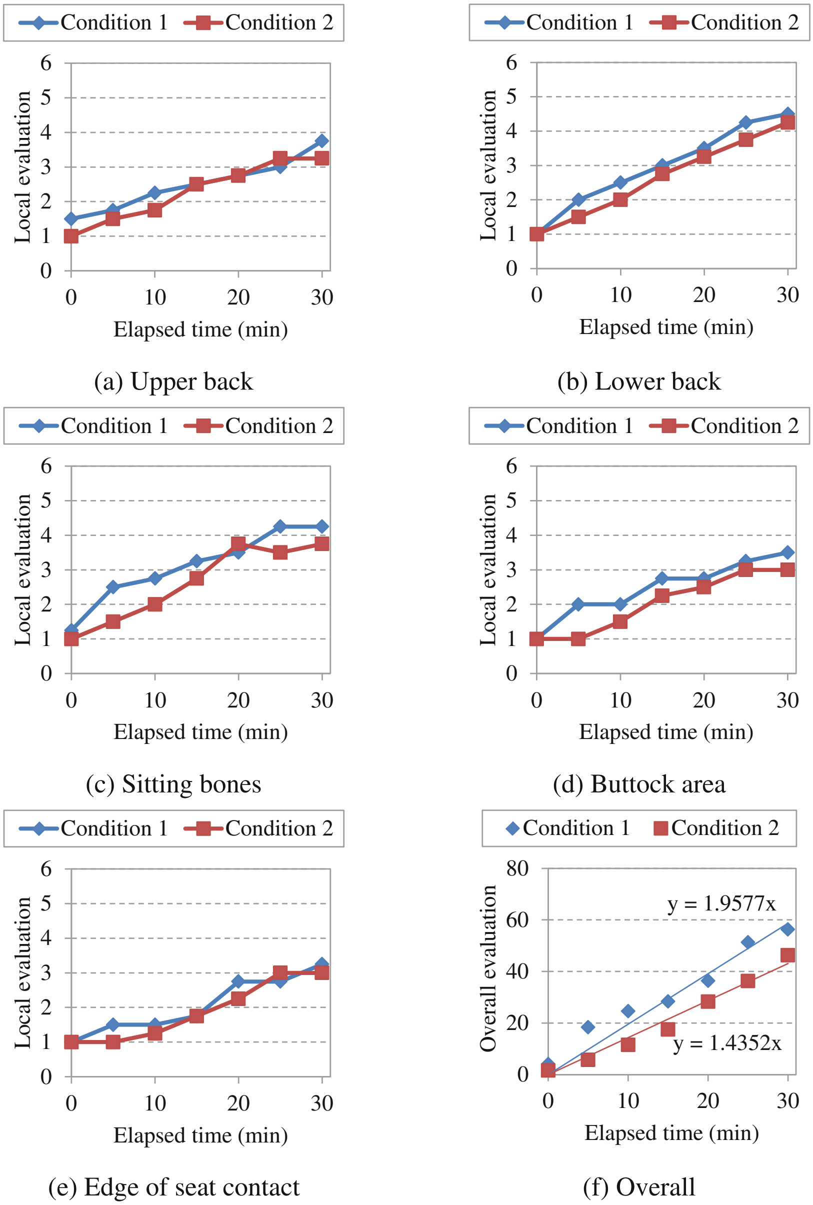


Fig. 5 Mean value of responses of the participants to the subjective evaluation

3.2 Subjective Evaluation

Table 4 shows the participant responses to the subjective evaluation under Condition 1 and Condition 2. As previously stated, the local evaluation of five body parts was recorded with the numbers from 1 to 6, and the overall discomfort rating was collected by the values from 0 to 100. The mean of all participant evaluations for every elapsed time period was calculated, and shown in the line chart in Fig. 5. The local evaluations indicated that the mean values in Condition 1 were greater than the mean values in Condition 2 with respect to the following body part: “Lower back”, “Sitting bones” and “Buttock area”. This suggested that the participants felt discomfort around lower back area, and any disease could be cause by the increased magnitude of whole-body vibrations under Condition 1. In addition, Fig. 5f depicts the mean value of overall evaluations and the regression line. The gradient of Condition 1 is 1.4 times the gradient of Condition 2. This implied that the difference between the discomfort ratings gradually increased with time. Consequently, the subjective evaluation suggested that decreasing whole-body vibrations by avoiding the rough surface is effective in improving ride comfort and also in preventing low back pain.

4 Conclusion

In this study, the driving simulator experiment was conducted to confirm the effectiveness of intelligent function that mitigates the whole-body vibrations of the occupants. The findings of the objective and subjective evaluations revealed the effectiveness of decreasing the whole-body vibrations of an autonomous car on the ride comfort improvement and low back pain prevention. However, a larger sample size is necessary. In addition, in reality, the participants may ride the driving simulator for longer periods than the time period in the experiment. Hence, as this study on whole-body vibrations mainly focused on driver ride comfort and health effects, it is important for other studies to investigate the whole-body vibration exposure of autonomous car occupants in the absence of driving tasks.

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, Proceedings of the 29th IEEE Conference on Decision and Control, pp. 2226–2231 (1990)

4. Taghirad, H.D., Esmailzadeh, E.: Automobile Passenger Comfort Assured Through LQG/LQR Active Suspension, J. Vib. Control, **4**(5), 603–618 (1998)
5. Mantaras, D.A., Luque, P.: Ride comfort performance of different active suspension systems, Int. J. of Vehicle Design, **40**(1/2/3), 106–125 (2006)
6. Kitade, J., Yoshida, K.: Design of controller for automotive semi-active suspension considering both ride comfort and driving stability (in Japanese), J. of JSME, Series C, **74** (739), 603–609 (2008)
7. International Organization for Standardization: Mechanical vibration -Laboratory method for evaluating vehicle seat vibration-, ISO 10326-1 (1992)
8. Yu, L., Khameneh, K.: Automotive Seating Foam: Subjective Dynamic Comfort Study, SAE Technical Paper 1999-01-0588 (1999)
9. International Organization for Standardization: Mechanical vibration and shock Evaluation of human exposure to whole-body vibration. International Standard, ISO 2631-1 (1997)
10. 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)
11. Setsuo MAEDA: Necessary Research for Standardization of Subjective Scaling of Whole-Body Vibration, Ind Health **43**, 390401 (2005)
12. El Falou, Wassim, Duchene, Jacques, Grabisch, Michel, Hewson, David, Langeron, Yves, Lino, Frederic: Evaluation of driver discomfort during long-duration car driving. Appl. Ergon. **34**(3), 249–255 (2003)
13. Takai, R., Yang, L., Noguchi, N.: Development of a crawler-type robot tractor using RTK-GPS and IMU engineering in agriculture. Environ. Food **7**(4), 143147 (2014)
14. Arai, T.: Robotics and Automation in Japanese Construction Industry, Proceedings of the 22nd ISARC (2005)
15. Naranjo, J.E., Gonzalez, C., Garcia, R., de Pedro, T.: Lane-change fuzzy control in autonomous vehicles for the overtaking maneuver. IEEE Trans. Intell. Transp. Syst. **4**(1), 13–22 (2008)
16. Hatipoglu, C., Ozguner, U., Redmill, K.A.: Automated lane change controller design. IEEE Trans. Intell. Transp. Syst. **9**(3), 438–450 (2003)
17. Tatsuno, J., Maeda, S.: Reducing whole-body vibration exposure of occupants in autonomous car by intelligent behavior, Proceedings of the 10th Asian Control Conference (ASCC 2015), Kota Kinabaru (2015)
18. Tatsuno, J., Nakamura, H., Yonehara, M., Takehara, S., Ishimatsu, K., Maeda, S.: Driving simulator for subjective scaling validation of the ISO 2631-1 standard, Proceedings of the 40th International Congress and Exposition on Noise Control Engineering (inter-noise 2011), DVD-ROM (2011)
19. Sammonds, G.M., Mansfield, N.J., Tatsuno, J., Fray, M., Maeda, S.: Effect of long term driving on driver discomfort in Japanese drivers,—A Simulator Study, Proceedings of 23rd Japan Conference on Human Response to Vibration (JCHRV2015), pp. 95–106 (2015)
20. Legislation.gov.uk, http://www.legislation.gov.uk/uksi/2005/1093/pdfs/uksi 20051093en.pdf

The Unknown Paradox of "Stop the Crash" Systems: Are We Really Improving Driver Safety?

Victoria A. Banks and Neville A. Stanton

Abstract This research assessed the appropriateness of the Autonomous Emergency Brake systems design using the Southampton University Driving Simulator. A total of 48 participants drove along a test route simulating testing procedures for Pedestrian Protection Systems at different levels of automation using different design strategies. It was found that whilst improvements to overall road safety was undeniably great regardless of design strategy, drivervehicle interaction patterns were affected in unexpected ways depending upon method of implementation. Contrasting design principles can therefore have varying effects on driver responses meaning that despite significant reductions in accident involvement, safety may not be improving in the way we expect. Overall, the paper concludes that we may be altering normal driver-vehicle interactions in a way that could be detrimental to driver behaviour in emergency situations opening up the debate over whether or not the benefits of automation outweigh potential costs.

Keywords Autonomous emergency brake · Driver behavior · Systems engineering

1 Introduction

In 2013, the World Health Organization declared that approximately 1.24 million people per annum die as a result of road traffic accidents with half of these (50 %) considered to be vulnerable road users; motorcyclists (23 %), pedestrians (22 %) and cyclists (5 %). Retting et al. [1] developed a number of engineering measures to reduce the number of pedestrian-vehicle collisions including speed control and separation zone measures. However, despite the intentions of these measures, 12 % of fatal collisions involving a pedestrian are thought to be caused by pedestrians

V.A. Banks (✉) · N.A. Stanton
Transportation Research Group, Engineering Centre of Excellence,
University of Southampton, Boldrewood Campus, Southampton SO16 7QF, UK
e-mail: v.banks@soton.ac.uk

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_44

expectantly entering the vehicle path [2]. This means that more automated assistance in emergency situations is likely to improve road safety further through the provision of visual and auditory warnings (e.g. Forward Collision Warnings) to give prior warning to the driver that they are in a hazardous situation and need to intervene to avoid a collision with another road user. In addition, emergency brake assistance (e.g. Autonomous Emergency Brake; AEB) is being introduced to avoid or mitigate the effects of road traffic accidents [3]. Autonomous braking intervention can be achieved in one of two ways; following a system warning that a collision is imminent (AEB Warning; AEBW) or with no overt warning to the driver (AEB No Warning; AEB nW). Strategies for implementation have been interpreted by manufacturers of vehicles in different ways and depend heavily upon a number of important factors such as the number and type of sensors available on the vehicle, the decision to warn the driver, the automated logic itself that determines when braking will be initiated and many more. The driving context does appear to influence the strategy for implementation, however, with city driving typically associated with an AEB nW system due to the certainty that autonomous assistance will be needed (i.e. greater proximity to hazards [4]). This is in contrast to inter-urban driving that is often associated with increased fitment of AEB W systems. Although it could be argued, based upon these design trends that a Pedestrian AEB system is most likely to adopt a non-warning based approach due to the propensity for pedestrians to be located in cities, it is not yet clear how such systems will be implemented. The European New Car Assessment Program in 2013 identified Pedestrian AEB as a critical safety system to be widely assessed and deployed from 2016 onwards [5] with common accident scenarios providing a baseline for the construction of test protocols (e.g. [6]). This is because it is agreed that the implementation of pedestrian detection and autonomous braking systems, regardless of design, will be advantageous in avoiding pedestrian collisions commonly caused by human error [7] and that automated assistance in such scenarios will have desirable benefits [8].

Although for some the benefits of automation may outweigh any costs, a considerable amount of research into vehicle automation has shown that drivers do not always respond in the way that engineers anticipate to automated assistance. Although in some instances driver behavioural change can be positive (e.g. if the driver is not looking at the road ahead or is distracted by other driving related tasks such as checking speed, an auditory warning could alert the driver to a problem and trigger their response), it can also be negative (e.g. drivers may become reliant on automation functionality and fail to respond as expected). For example, [9] found that increasing the level of automation can lead to complacency whilst [10] suggested that it can cause decreases to driver situation awareness which are closely related to issues of mental underload and overload [11]. The authors argue that if we are to overcome these issues, more research is needed to ensure that undesirable behavioural adaptation does not occur [12, 13] which can only be achieved if we acknowledge the *new* role of the driver in an automated driving system rather than purely focusing upon the efficacy of automation on frequency of accident involvement as a marker to determine if automated systems really do improve road safety.

For this reason, this research assessed the appropriateness of Autonomous Emergency Brake systems design using the Southampton University Driving Simulator. With one of the greatest challenges remaining to be addressed centered on whether or not the introduction of automation into driving emergencies brings about any performance increments or decrements on behalf of the driver, evidence from driver observations (through means of retrospective verbal commentary [14]), was used to further explore how driver-vehicle interaction patterns may change depending upon different system design approaches.

2 Methodology

2.1 Participants

A total of 48 participants were recruited from University of Southampton student and staff cohort. All participants held a full UK driving license for a minimum of one year and were between the ages of 18 and 65. This was to ensure that the performance decrements associated with older drivers (i.e. over 65 s) and novice drivers (i.e. drivers with less than 12 months driving experience) did not affect the results of the study.

Ethical permission to conduct the study was granted by the Research Ethics Committee at the University of Southampton.

2.2 Design and Procedure

The experimental method that that was described by Banks et al. [15] is applicable to this study. This paper details the analysis and interpretation of data that was collected in real-time by the Southampton University Driving Simulator with regard to driver control inputs (or lack thereof) in coping with critical pedestrian events which was beyond scope of the Banks et al. [15] paper.

The study consisted of four driving conditions representing different levels of automation (within Endsley and Kaber’s automation taxonomy [16]) taking into consideration the alternative methods of systems design implementation;

- Manual Control where the Driver was required to complete all of the physical and cognitive workload associated with the driving task
- Decision Support where the Driver was assisted by an auditory/visual warning to detect hazards in the roadway ahead
- Automated Decision Making represented the overarching level of automation at which an Autonomous Emergency Brake feature would sit. At this level of

automation, an automated feature is given the authority to decide and act upon events within the environment. It is within this level of automation whereby different design strategies could be pursued;

1. Escalating warning approach – automated system alerts the driver to a critical hazard in the road ahead and is coupled with autonomous braking (AEB W);
2. Non-warning based approach—automated system deliberately omits the use of a warning to alert the driver prior to autonomous braking (AEB nW). This is based on the logic that autonomous emergency braking is a last resort system and should only intervene when the moment for human intervention passes and thus there is no need for any additional feedback.

Aside from the Manual driving condition that required drivers to complete all of the physical and cognitive tasks associated with driving, the remaining three automated driving conditions (Warning, AEB W and AEB nW) were designed so that drivers received assistance in the intervention of critical emergency events. These were defined as any event that without intervention would result in an accident without human intervention. Drivers were informed that the automated systems would only intervene in critical events meaning that no warnings or brake assistance would be provided for non-critical events. Both AEB systems were design to intervene regardless of driver control inputs, in this way acting as assistance if the driver initiated emergency braking first, or capable of acting autonomously if the driver failed to respond. Automatic intervention by AEB aimed to improve on the reaction time of the driver in an effort to mitigate injury rather than collision avoidance. Thus drivers were told that AEB was not deemed a replacement for them. Instead AEB should be viewed as a 'last resort' intervention strategy, in this way, maintaining the driver within the control-feedback loops for as long as possible.

3 Results

A critical outcome measure relevant to the evaluation of automation in driving emergencies is the frequency of accident involvement. Figure 1 reveals that the potential for automation to reduce the number of accidents, regardless of strategy for implementation (i.e. use of different systems), has undeniable benefits.

At this point, the benefits of automation could lead to the ignorance of any potential subsequent behavioral implications (see [17] for a review) as on the face of it, AEB implementation seems positive with the difference between a warning based system (AEB W) and a non-warning based system (AEB nW) only having a small offset in accident involvement. However, without further analysis, we may be neglecting crucial evidence that could affect the future of AEB design implementation. For instance, Banks et al. [18, 19] proposed that the design of automation

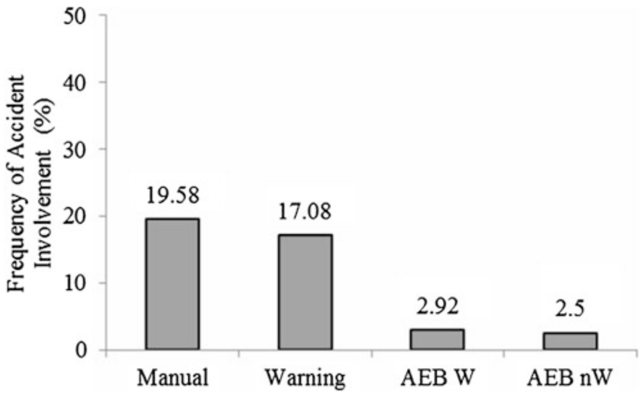


Fig. 1 Levels of automation and accident involvement observed during the study

may affect the way in which drivers approach and deal with hazardous events within the driving environment due to the workload shift that occurs when automation is introduced into the driving system. These workload shifts may lead to performance decrements on behalf of the driver. For example, skill degradation may occur if drivers delegate their control of braking to the automated subsystems (i.e. there is an unintended shift in task loading which inhibits traditional behavioural response [10, 20–22]). Drivers may feel that the onus of responsibility for reacting to hazardous situations is now shared with an automated counterpart [23] and thus, may delay their normal response as they attempt to cooperate with it [24]. Indeed, several studies have reported that performance under increased levels of automation can begin to decline as a result of lack of manual control inputs [21, 22, 25–30]. In addition, if automation is perceived to be highly reliable, drivers may not be able to monitor the system as closely as is warranted [11]. This could delay driver response if they wait for notification or confirmation of a collision risk such as that offered in AEB W.

Thematic analysis of driver verbalizations collected during the study suggested that the level of automation, and type of automation implemented, directly affected the way in which drivers interacted with the vehicle and the way in which they appraised critical braking events (see [15] for a full discussion). To explore this further, data from SUDS was used to assess the implications of automation design on drivervehicle interaction patterns. At this point, only data for AEB W and AEB nW were selected as it was in these conditions that the driver could have relinquished their full control of braking to the automation. Within the data files, it was clear to see “who” (Driver or Automation) was responsible for initiating the braking effort as the AEB feature had a clear identifier within the dataset. This raised the possibility to essentially log “who” interacted with the braking system first.

A Chi-Square test revealed that the design of the system (i.e. AEB W or AEB nW) significantly affected the frequency of “who” (Driver or Automation) initiated the braking response ($\chi^2(3) = 390.29$, $p < 0.001$, Cramer’s $V = 0.638$). Figure 2

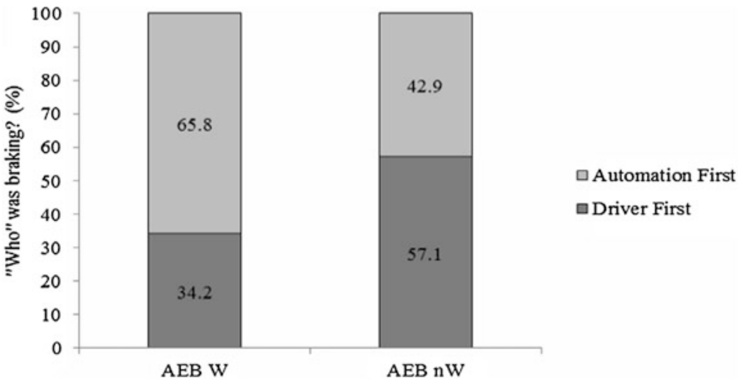


Fig. 2 "Who" is braking?

indicates that AEB nW was associated with a higher prevalence of 'Driver First' responses whilst AEB W was associated with an increased prevalence of 'AEB First' responses. 'Driver First' in this instance simply reflects that the driver responded to the critical braking event prior to AEB activation with 'AEB First' being the opposite. Out of a total 240 critical braking events in each driving condition, 34.2 % (n = 82) 'Driver First' responses were logged for AEB W in contrast to 57.1 % (n = 103) for AEB nW.

In order to discover if "who" reacted to the hazardous event in the first instance really did affect overall stopping distances, a Mann Whitney U test based upon 'Driver First' and 'Automation First' responses (regardless of AEB design) was conducted. It revealed that stopping distance was indeed significantly affected by "who" engaged the braking first ($U = 43127.5$, $z = 10.63$, $p < 0.05$). Stopping distances relating to 'Driver First' responses (*Mdn* = 19.98 m) were lower than 'Automation First' responses (*Mdn* = 23.41 m) resulting in a median difference of 3.43 m. This would suggest that maintaining the traditional role of the driver (i.e. one that reacts prior to AEB activation) remains the most desirable. It would seem that AEB nW provides the environment that is more likely to achieve this response (Fig. 2). These results simply highlight that if AEB is to remain active in the background of vehicle operation, as intended, the frequency of 'Driver First' responses should remain high. AEB nW may have encouraged drivers to monitor the road ahead much more diligently than other automated strategies, perhaps because they knew no collision warning would assist them in detecting hazardous situations. Even so, normal driver responses must have been delayed in AEB W given the higher frequency of 'Automation First' responses and that both AEB systems were designed using the same algorithm. With drivers being aware of the reliability of the system during simulation, it is possible that drivers were relying upon system activation in particular using the warning mechanism to trigger a response (similar to the purpose of a Forward Collision Warning) and thus not monitoring the environment as closely as was warranted [11].

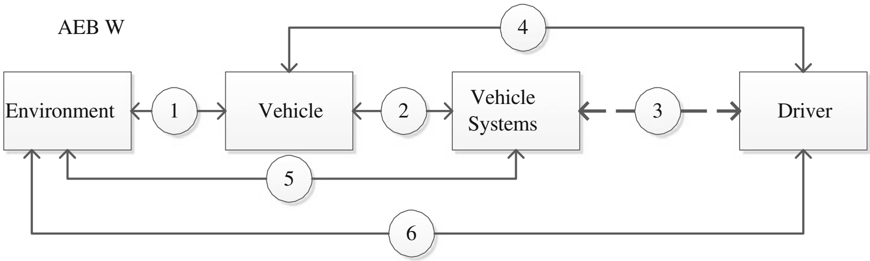


Fig. 3 Control-feedback loops in driving (Adapted from Banks et al. [20]) with loop 3 suggesting a potential decoupling between the driver and vehicle systems

4 Discussion

Assuming that automation is 100 % failsafe, common causes of vehicular accidents such as driver distraction, inattention and a lack of timely response could be offset by its implementation [31–34]. However, this study has demonstrated that whilst road safety can be improved with the implementation of AEB in emergencies, the strategy of implementation determines how far the traditional role of the driver remains and this is well worth recognizing.

It can be argued that in order to maximize the safety of drivers and other vulnerable road users, designers should be aware of how different design approaches at differing levels of automation could affect subsequent responses to critical hazards. This study has shown that despite asking drivers to react to hazardous events as they normally would, the implementation of AEB did affect driver-vehicle interaction not only with regards to their decision-making (Chap. "The Use of Modelling Tool in order to Evaluate the Dwelling Times for Trains") but also with regards to the way in which they interacted with the vehicle and braking system (Chap. "An Overview of the Factors Associated with Driver Distraction and Inattention within the South African Railway Industry"). This means that we may not be improving the safety of our drivers as their changing role has not been fully recognised. AEB implementation, regardless of design strategy at this stage, appears to weaken the control-feedback loops (Chap. "Effects of Driver Characteristics and Driver State on Predicting Turning Maneuvers in Urban Areas: Is There a Need for Individualized Parametrization?"). This could lead to drivers not being equipped to cope with hazardous situations if automation failed [35]. It also implicates the concepts of trust and complacency in automation functioning [11, 36].

It would appear that AEB implementation may lead to decoupling the link between the driver and vehicle systems within the control-feedback loops (Fig. 3) which may explain to some extent why 'Automation First' responses occurred. This decoupling was more pronounced when AEB W was used suggesting that a non-warning based AEB system is better able to preserve the traditional role of the driver. Even so, in some instances a warning based system may be preferable—especially in instances whereby the driver has failed to recognise a hazard in the road ahead.

Acknowledgments This research is supported by the Engineering and Physical Sciences Research Council (EPSRC) and Jaguar Land Rover Ltd.

References

1. Retting, R.A., Ferguson, S.A., McCartt, A.T.: A review of evidence-based traffic engineering measures designed to reduce pedestrian-motor vehicle crashes. Am. J. Public Health **93**(9), 1456–1463 (2003)
2. National Highway Traffic Safety Administration.: National Pedestrian Crash Report, DOT HS810 968 (2008)
3. Lenard, J., Badea-Romero, A., Danton, R.: Typical pedestrian accident scenarios for the development of autonomous emergency braking test protocols. Accid. Anal. Prev. **73**, 73–80 (2014)
4. Road Safety Analysis.: Stepping out, pedestrian casualties: An analysis of the people and circumstances. Commissioned by the Parliamentary Advisory Council For Transport Safety. http://www.roadsafetyanalysis.org/ (2013)
5. van Ratingen, M.: Fifteen Years for Safer Cars/AEB Fitment Survey. Proceedings of the 15th Anniversary Event on Avoid the Crash with Autonomous Emergency Braking Systems, Brussels, Belgium, 13 June 2012, Accessed 20 Jan 2015, http://www.amiando.com/FNPIBLY.html?page=760000 (2012)
6. Lenard, J., Russell, D., Avery, M., Weekes, A., Zuby, D., Kuehn, M.: Typical pedestrian accident scenarios for the testing of autonomous emergency braking systems. Proceedings of the 22nd International Technical Conference on the Enhanced Safety of Vehicles, Washington, DC. (2011)
7. Habibovic, A., Tivesten, E., Uchida, N., Bargman, J., Aust, M.L.: Driver behaviour in car-to-pedestrian incidents: An application of the driving reliability and error analysis method (DREAM). Accid. Anal. Prev. **50**, 554–565 (2013)
8. Searson, D.J., Anderson, R.W.G., Hutchinson, T.P.: Integrated assessment of pedestrian head impact protection in testing secondary safety and autonomous emergency braking. Accid. Anal. Prev. **63**, 1–8 (2014)
9. Stanton, N.A., Dunoyer, A., Leatherland, A.: Detection of new in-path targets by drivers using stop and go adaptive cruise control. Appl. Ergon. **42**(4), 592–601 (2011)
10. Parasuraman, R.: Designing automation for human use: empirical studies and quantitative models. Ergonomics **43**(7), 931–951 (2000)
11. Young, M.S., Stanton, N.A.: Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. Hum. Factors **44**(3), 365–375 (2002)
12. Stanton, N.A., Marsden, P.P.: From fly-by-wire to drive-by-wire: safety implications of automation in vehicles. Saf. Sci. **24**(1), 35–49 (1996)
13. Stanton, N.A., Pinto, M.: Behavioural compensation by drivers of a simulator when using a vision enhancement system. Ergonomics **43**, 1359–1370 (2000)
14. Ericsson, K.A., Simon, H.A.: Protocol Analysis: Verbal Reports as Data. MIT Press, Cambridge, MA (1993)
15. Banks, V.A., Stanton, N.A.: Contrasting models of driver behaviour in emergencies using retrospective verbalisations and network analysis. Ergonomics **58**(8), 1337–1346 (2015)
16. 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)

17. Sheridan, T.B., Wickens, C.D., Parasuraman, R.: Situation awareness, mental workload, and trust in automation: viable, empirically supported cognitive engineering constructs. J. Cogn. Eng. Decis. Making **2**(2), 140–160 (2008)
18. Banks, V.A., Stanton, N.A., Harvey, C.: What the crash dummies don’t tell you: the interaction between driver and automation in emergency situations. Proceedings of the IEEE Intelligent Transportation Systems for All Transport Modes 2013, The Hague, The Netherlands, 6–9 Oct 2013
19. Banks, V.A., Stanton, N.A., Harvey, C.: Sub-systems on the road to vehicle automation: hands and feet free but not ‘mind’ free driving. Saf. Sci. **62**, 505–514 (2014)
20. Jameson, A.: Adaptive interfaces and agents. In: Jacko, J.A., Sears, A. (eds.) Human-Computer Interface Handbook, pp. 305–330. NJ, Erlbaum, Mahwah (2003)
21. Miller, C.A., Parasuraman, R.: Designing for flexible interaction between humans and automation: delegation interfaces for supervisory control. Hum. Factors **49**(1), 57–75 (2007)
22. Stanton, N.A., Young, M.S., Walker, G.H., Turner, H., Randle, S.: Automating the driver’s control tasks. Int. J. Cogn. Ergon. **5**(3), 221–236 (2001)
23. Hoc, J.M., Young, M.S., Blosseville, J.M.: Cooperation between drivers and automation: implications for safety. Theor. Issues Ergon. Sci. **10**(2), 135–160 (2009)
24. Hollnagel, E., Nabo, A., Lau, I.V.: A systemic model for driver-in control. Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Park City, Utah, USA, 21–24 July, pp. 86–91 (2004)
25. Lavie, T., Meyer, J.: Benefits and costs of adaptive user interfaces. Int. J. Hum Comput Stud. **68**, 508–524 (2010)
26. 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.: Publ. IEEE Syst. Man Cybern. Soc. **30**(3), 286–297 (2000)
27. Rudin-Brown, C.M., Parker, H.A.: Behavioural adaptation to adaptive cruise control (ACC): implications for preventive strategies. Transp. Res. Part F: Traffic Psychol. Behav. **7**(2), 59–76 (2004)
28. Vollrath, M., Schleicher, S., Gelau, C.: The influence of cruise control and adaptive cruise control on driving behavioura driving simulator study. Accid. Anal. Prev. **43**(3), 1134–1139 (2011)
29. 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)
30. Young, M.S., Stanton, N.A.: Back to the future: Brake reaction times for manual and automated vehicles. Ergonomics **50**(1), 46–58 (2007)
31. Amditis, A., Pagle, K., Joshi, S., Bekiaris, E.: Driver-Vehicle-Environment monitoring for on-board driver support systems: lessons learned from design and implementation. Appl. Ergon. **41**(2), 225–235 (2010)
32. 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)
33. Donmez, B., Boyle, L.N., Lee, J.D.: Safety implications of providing real-time feedback to distracted drivers. Accid. Anal. Prev. **39**(3), 581–590 (2007)
34. 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**(2), 149–159 (1997)
35. Sarter, N.B., Woods, D.D., Billings, C.E.: Automation surprises. Handb. Hum. Factors Ergon. **2**, 1926–1943 (1997)
36. Lee, J.D., See, K.A.: Trust in automation: Designing for appropriate reliance. Hum. Factors **46**, 50–80 (2004)

Calibrating Trust to Increase the Use of Automated Systems in a Vehicle

Siddartha Khastgir, Stewart Birrell, Gunwant Dhadyalla and Paul Jennings

Abstract While accident data show that human error is the cause of most of the on-road accidents, the move towards assisting or replacing the human driver with an automated system is not a straightforward one. Industries like aerospace, manufacturing, process etc. have a high penetration of automated systems; however, the automotive industry provides new challenges due to different and more dynamic interactions between the actors (driver, vehicle and environment). To reap benefits from the automated systems, drivers have to use the automated systems. Drivers' trust in automated systems is one of the most important factors influencing drivers' use of automated systems. Trust on automated systems is a dynamic construct, which can change with experience. While discussing various factors which influence drivers' trust on automated systems, this paper discusses the changing nature of trust, i.e., calibration of trust and the possible interventions to calibrate trust on automated systems to an appropriate level.

Keywords Calibration of trust · Automated systems · Driver-automation interaction · Driver-in-the-loop

S. Khastgir (✉) · S. Birrell · G. Dhadyalla · P. Jennings
WMG, University of Warwick, Coventry CV4 7AL, UK
e-mail: S.Khastgir@warwick.ac.uk

S. Birrell
e-mail: S.Birrell@warwick.ac.uk

G. Dhadyalla
e-mail: G.Dhadyalla@warwick.ac.uk

P. Jennings
e-mail: Paul.Jennings@warwick.ac.uk

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_45

1 Introduction

Increased driver safety has been advocated as one of the potential benefits of automated systems. According to a statistical study conducted by the National Highway Traffic Safety Administration (NHTSA), causal factor for 94 % of the on-road accidents has been assigned to the human driver [1]. This has led to an increasing emphasis from manufacturers and legislators on the introduction of automated systems in transportation systems.

However, customer uptake of automated technology has been slow. Market penetration of systems like Adaptive Cruise Control (ACC) introduced as early as 1998–99 by Toyota and Jaguar has been slow and limited [2]. Automated systems have traditionally been technology driven and there is a need to re-focus automated system development on the driver rather than the system [2]. Any benefit from automated systems can be realized only if the drivers use such systems. Therefore, it is important to understand the factors that influence drivers' use of automated systems. One of the most important factors influencing drivers' use of an automated system is drivers' trust on automated systems [3, 4]. A driver will give up control of a vehicle to an automated system, only if he/she trusts the automated system [5]. To increase the use of automated systems, there is a need to design automated systems such that the drivers find them trustworthy. Development of trust in an automated system is a multi-variate process [6]. Additionally, it is important to understand the interactions of the various factors that affect development of trust in order to understand the design requirements for automated systems. This is an important aspect of research as the individual factors compete against each other to increase trust on automated system.

Automation in vehicles, if designed incorrectly, can have a negative impact also. Introduction of automation can defeat the purpose of automation itself which forms the ironies of automation [7]. One of the arguments in favour of introduction of automation has been that the removal of the human driver from the control loop will reduce the errors due to an unreliable and inefficient human driver. However, removal of the human driver makes the system prone to errors introduced in the system by the designer (another human influence) [7]. This forms the first irony. Additionally, the designer automates only the tasks he/she can automate without understanding if a task needs to be automated or not [7]. There is always a possibility that the designer's perception of the best design for a system and distribution of tasks between the driver and the automated system differs from the driver's perception; and may increase the workload of the driver [8]. This forms the second irony. This may ultimately affect the driver's trust on the system if the system is not correctly designed or tested.

This leaves the driver with the task of (1) monitoring the automated tasks (2) performing the non-automated tasks. A near-perfect automated system is more detrimental for human-machine interaction than an unreliable automated system. This is due to the rare need for human intervention in a near-perfect system, causing the human driver to be out of the loop for most periods. This means that the driver has to perform the monotonous task of monitoring the automated system, which as per Fitts list's [9] is not a task human driver is good at when compared to an automated system. This forms the third irony [7]. Fitts suggested that tasks at which machines outperform humans should be automated and tasks at which humans outperform machines should be left to manual control. Thus, driving tasks that are automated need to be carefully selected and designed in order to develop drivers' trust on the automated systems to increase their use.

2 Trust in Automated Systems in Vehicles

Trust on an automated system is one of the most important factors for the system to be used by the driver [3, 10, 11]. Therefore, in order to reap benefits of an automated system, one needs the driver to develop trust on the system so that he/she uses the system [6]. To define trust, authors have adapted the definition of trust from [5, 6, 12] with the addition of the experience dimension. The authors define trust as "*a history dependent attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability*". There are four important aspects of this definition. Firstly, task should be a goal-oriented task. Secondly, there should be a degree of uncertainty associated with the task, so that prediction of future is not deterministically possible. Thirdly, there should be a degree of risk associated with giving up manual control making the driver feel vulnerable. Fourthly, trust changes with time based on previous experiences.

Trust in itself is influenced by various factors [6] and development of trust is a dynamic process [4]. The contribution of each factor in the process of development of trust has been discussed in this section. In two experiments performed by Muir [10], it was shown that the subjective measure of trust in a system or sub-system could be correlated with the operator's use of the system. However, misplaced trust (leading to misuse or disuse) in an automated system can lead to incorrect usage of the system [13]. Many aircraft accidents have been a result of misplaced trust or over-trust of the pilots on automated systems in aircraft systems (like autopilot) [14]. Additionally, many aircraft accidents have been a result of dis-trust of the pilots on automated systems in aircrafts, e.g. pilots mistrusting the stall warning in the Air France 447 crash in June 2009 [15]. This leads to a qualitative classification of trust into three categories [13]:

1. *Mistrust* or *Over trust*
2. *Distrust*
3. *Appropriate trust*

Mistrust refers to over-trust on the capabilities of the automated systems and engaging them in situations that they are not capable of handling. Mistrust comes from lack of knowledge about the true capabilities of the system. *Distrust* refers to lack of trust on part of the driver on an automated system to perform a task even if it is fully capable of performing the required task [13]. *Appropriate trust* on an automated system comes from the correct knowledge of the working of the automated system and knowledge of its limitations. Appropriate trust guides correct use of an automated system, i.e., driver engages the automated system only for the tasks and situations that it is capable to handle and disengages automation in adverse conditions (e.g. rainy weather, fog etc.).

2.1 Types of Trust

Quantitatively, trust on an automated system in a vehicle can be classified into the following components [16]:

1. *Trust in the system*
2. *Trust with the system*

There is a subtle though important difference between the two types of trust. *"Trust in the system"* refers to system's capabilities in doing what it is supposed to do and is guided by the perceived knowledge of the capabilities (which may or may not be accurate). "Trust in cooperation with the system" or "*trust with the system*" refers to accepting system limitations and adapting system's usage to get maximum benefit from the system. Trust with the system implicitly means that the driver is aware or has "*accurate knowledge*" about the system's true capabilities and limitations and makes informed decisions on when to use, when not to use and how to use the automated system.

2.2 Factors Influencing Trust on Automated Systems in Vehicles

In the framework for trust on automated systems proposed by Riley [17], it was suggested that trust is not the sole factor responsible for an operator's use of an automated system. It is a result of multiple interactions between interdependent variables such as trust, risk (discussed in Sect. 2.2.2), self-confidence (discussed in Sect. 2.2.5) and many others. In a later version of the framework, the framework was corrected after experimental studies and an "experience" dimension was added [18]. It suggested that trust influences use of automation and is influenced by experience with automation. This idea of trust influencing use of automation garnered support from other studies [6], but has been adapted to indicate that "*trust guides—but does not completely determine—use*". This is especially true in difficult situations where complete situation understanding is impractical. However, many studies [6, 16, 19] have proved that trust in itself is also influenced by a number of factors and has different forms in itself too. Factors influencing trust have been discussed in the following sections.

2.2.1 Knowledge

Knowledge (accurate) about automation capabilities leads to the development of "appropriate trust", i.e., trust based on "accurate" facts, which in turn leads to appropriate use of the system [20]. This knowledge comprises of awareness of automation capabilities, real time information of the automation state and internal mental model of the automated system. Knowledge affects development of trust in general (both forms of trust: trust in the system and trust with the system).

A trade-off needs to be developed between providing "trustworthy" automation or "trustable" automation [6]. Trustworthy automation involves complex automated system capable of handling most scenarios. However, the complex nature of the interactions make it difficult for the driver to understand information about the automation and may lead to lack of trust [21]. On the contrary, a simpler automation, which is easily understandable and easily conveyed to the driver is more trustable and would lead to higher trust levels. A survey study of around 1000 pilots by McClumpha et al. [22] found that the primary reason for variance in pilot attitudes towards automated systems was due to their understanding of the automated systems. A driver can misinterpret a correctly working automated system due to an incorrect mental model made because of lack of knowledge of the working of the system. In an automated system, there is always a fear of giving too much information to the driver. Overflow of information has been proved to be detrimental which has led to a consensus on adopting a lean approach to information display with an option for drivers to request for more information from the system [8].

Authors propose the following classification of knowledge:

- *Static knowledge*: refers to the understanding of the working of the automated system
- *Real time*: refers to the real time information about the state of the automated system and the environment
- *Internal mental model*: refers to understanding or influence of external sources (e.g. word of mouth, media etc.)

Static knowledge influences *trust in the system* and is partially responsible for setting the initial trust level. Real time knowledge influences *trust with the system* as it informs the driver about the real time ability of the automated system. Internal mental model (along with static knowledge) is responsible for setting the initial trust level. Another dimension to knowledge is the accuracy of the information presented. If automation does not meet the initial expectations (information conveyed), it can lead to distrust on an automated system and lead to the drivers not using the automated system.

2.2.2 Consequence

While proposing the framework for trust on autonomous systems, Riley found that failure of automation didn't lead to any reluctance in the re-use of automation [18], which is contrary to the other theories of trust which show a hysteresis nature in the development of trust before and after failure of automation (discussed in Sect. 3). The difference in the findings can be explained by the lack of risk involved as a consequence of engaging automation in the former study. In order for trust on a system to be an influencing factor in the system's use, there needs to be a degree of risk attached to the deployment of the automated system [6]. Drivers must put themselves at risk by handing over control to the automated system. Different types of automated systems exist which provide varying degrees of assistance, from information to warning to active control of the vehicle. Drivers tend to trust informative and warning systems more readily than systems that take active control. This is because drivers are unwilling to give up control in critical scenarios, as they trust their own abilities more than the abilities of the automated system, even though they may not be best suited to handle the situation [5, 9]. However, driver can develop distrust even in an imperfect warning system in case it provides too many false warnings, i.e., high *frequency* of failures.

Humans have a tendency to monitor a highly trusted automation less frequently [10]. Due to lack of monitoring in a highly trusted system, trust may grow even if faults occurred. However, this is not entirely true and has a context dimension to it. Growth or fall of trust depends on the consequence of the faults. Since the driver is monitoring less frequently, his/her situation awareness is less and in case of a critical fault, his/her ability to cope with the situation would be limited. This can result in catastrophic loss of trust (i.e., rapid degradation of trust level). Thus, the *severity* of consequence dictates the fall in trust level when failures occur. Therefore, consequence has two aspects: severity of failure and frequency of failure. While severity affects both aspects of trust, frequency affects *trust with the system*.

2.2.3 Situation Awareness

Situation Awareness is a one of the most important factors for development of trust on automated systems, especially when the driver in still present in the control loop (SAE level 1–3 automation [23]). Increasing automation in systems (if not correctly designed) can lead to loss of situation awareness of the driver [24]. This can be caused due to variety of factors: humans given the role of monitoring the system, a task they are not suitable for [9], inadequate information conveyed to the driver about the state of automated system or over trust on automated system leading to complacency. Situation awareness (SA) is classified into three levels [24]:

- Level 1 SA: (Perception): perceiving critical factors in environment
- Level 2 SA: (Understanding): understanding meaning of the perceived factors (individually and collective)

- Level 3 SA: (Predicting): predicting the future state of the automated system

Level 1 SA and Level 2 SA influence the development of *trust with the system* on an automated system in a vehicle. The information of the critical factors and understanding of the information makes the driver situationally aware of the surroundings and enables him to make a decision on whether the environment is suitable for deployment of the automated system. While real time knowledge concerns the content of the information presented to the driver, level 1 SA and level 2 SA concern the manner of conveying the information to the driver, i.e., visual, audio, haptic etc., in order to make the driver situationally aware.

2.2.4 Willingness

In an experiment conducted with aviation pilots and university students as subjects, it was found that aviation pilots' average use of a non-aviation automated system in the experiment was significantly more than the average use of the same automated system by students [18]. Since aviation pilots are used to working in conjunction with automated systems, their willingness to use automation in any setting is higher than an average person. Willingness in engaging an automated system influences *trust with the system* and is developed as a result of experience or training.

2.2.5 Self-confidence

High self-confidence and low self-confidence have different interactions and influence over use of an automated system. While high level of self-confidence decreases the tendency to use automation to perform a driving related task [18], low self-confidence influences *trust with the automated system*. This can be explained as the people having low self-confidence have more trust in the abilities of the automated system than their own abilities [11].

3 Calibration of Trust

It is more difficult to recover trust after failure of an automated system as compared to the initial development of trust on the automated system [5]. In an experiment in which the real time information of Adaptive Cruise Control (ACC) state was provided to the driver, it was shown that drivers' trust on automation was greater when it was degrading than when it was recovering [21]. Additionally, trust on ACC had a strong hysteresis, with high trust in the first half of the usage and low trust in the second half. The hysteresis nature of trust evolution in an automated system is a result of driver's interaction with the automated system during its use. Different scenario interactions can lead to increase or decrease of trust on automated systems. This changing nature of trust is referred as calibration of trust. More objectively, calibration of trust has been defined as "*the process of adjusting trust to correspond to an objective measure of trustworthiness*" [4]. As a result, identification of different scenario interactions and its effect on calibration of trust on automated systems becomes relevant.

Calibration occurs with specific interactions and that leads to the development of functional specific trust [25]. Trust is an ever changing construct and the influencing factors change as the driver-automation relationship progresses [26]. Initial level of trust on an automated system depends on static knowledge, internal mental models of the driver and willingness of the driver, and is based on the performance (reliability) of the system, which evolves after process (dependability) interaction and finally stabilizes with the purpose (faith) interaction of the system [6].

Calibration of trust is a result of closed loop dynamic interactions of the driver and the factors discussed in Sect. 2.2 of this paper. The following stages characterize the process of calibration of trust on an automated system:

- *Stage A*: Initial level of trust may be high because of static knowledge (information), internal mental model of the driver and driver's willingness. As failures start to occur, there is a period of stable trust levels representing mistrust (over-trust) on the system leading to misuse.
- *Stage B*: As number of failures increase (in frequency and severity), calibration of trust occurs leading to decrease in trust levels. If the number and frequency of failures keep increasing, it may reach a level where driver has no trust on the automated system and opts for manual control. The rate of change of trust levels in stage B is dependent on the severity of the failure (which along with frequency forms the other aspect of the consequence factor).
- *Stage C*: Once the trust level has decreased considerably, it is more difficult to regain trust even if the automated system recovers or the number of failures decrease considerably. This represents distrust on the automated system, i.e., the system is not trusted even though the system is capable of performing a task.
- *Stage D and Stage E*: With further decrease in failures, trust initially recovers slowly and then rapidly. It may or may not reach back to the initial level of trust. The rate of recovery in stages D and E is dependent on the driver's experience during the failure and is governed by the consequence factor.

Depending on the type of automation, type of failure and consequence of failure, the duration and rate of change of trust in each section will be different. Figure 1 shows different paths for calibration of trust under different situations. Figure 1a represents an informative system, Fig. 1b represents a warning system and Fig. 1c represents a system that takes active control (e.g. highway chauffer). Figure 1d represents a system that takes active control and has real time knowledge (information) feedback for the driver.

Informative System: Systems that provide only information to the driver (e.g. navigation systems) have the most gradual degradation of trust in case of failures.

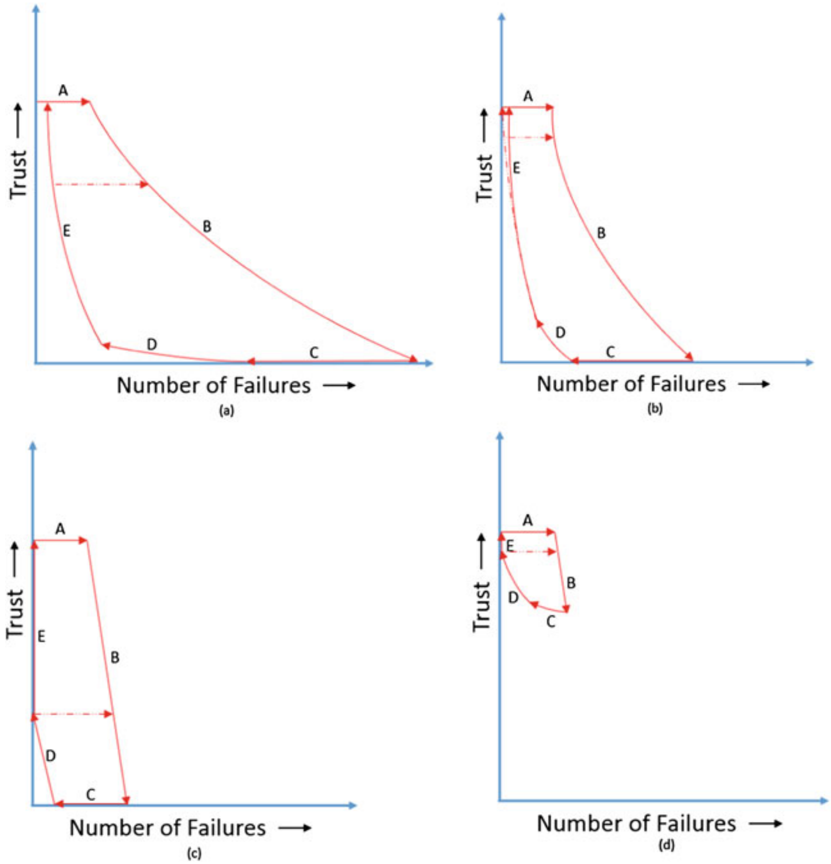


Fig. 1 Calibration of trust on automated systems as a function of number of failures. **a** Informative system. **b** Warning system. **c** System taking active control from driver. **d** System taking active control from driver with real time knowledge (feedback) intervention provided to the driver

Since the consequence of a failure is relatively less severe, drivers tend to be more accommodating with failures in an informative system.

Warning System: Warning systems have a similar trend as informative systems, however the rate of degradation of trust is more rapid and trust is regained only after substantial recovery of systems. False warnings lead to drivers developing distrust on the system and leads to degradation of *trust with the system*.

Systems taking over active control from the driver (partially or full automation systems, e.g., highway chauffer): Failures in systems taking over active control of the vehicle from the driver can lead to drastic degradation of trust on automated systems. This is because the consequence of a failure in such systems can be potentially catastrophic (fatal in some situations).

3.1 Intervention to Re-calibrate Trust on Automated Systems

While development of trust on automated systems is influenced by “*knowledge*”, “*consequence*”, “*situation awareness*”, “*willingness*” and “*self-confidence*”, intervention methods can be incorporated within automated systems to calibrate trust to appropriate levels or to prevent rapid degradation of drivers’ trust on automated systems.

Providing knowledge about the automated systems in real time is one of the intervention methods. While static knowledge helps guide the initial level of trust on an automated system, *real time knowledge* about the state of the automated system and environment can prevent degradation of drivers’ trust by re-calibrating it to appropriate level.

Figure 1d depicts calibration of trust in a system which takes over active control of the vehicle and which has real-time knowledge intervention, i.e., real-time knowledge feedback to the driver about the automated system. By providing real time information about the state and capabilities of the automated system, it is possible to stem the fall of trust on the automated system and re-calibrate to have appropriate level of *trust with the system*. Real time information feedback helps make the driver situationally aware about the current capabilities of the automated system and aids re-calibration of trust.

4 Conclusions

Automated systems in the automotive context vary in their levels of autonomy and the role assumed by the driver. While increase in the level of autonomy decreases the tasks for the driver, the driver still has a choice to use or not use the automated systems. In order to reap benefits from the automated systems, it is important for the driver to use these systems. Drivers’ trust on these systems govern drivers’ use of automated systems. Drivers’ trust is a dynamic variable and changes with drivers’ experience with the automated systems and this process of changing trust levels is referred to as calibration of trust. Calibration of trust on automated systems can be affected by factors like consequence, type of automation and knowledge. It is important to calibrate drivers’ trust to the true capabilities of the automated system for appropriate use of automated systems. This paper discusses the various factors affecting development of trust and interventions that can be adopted to increase trust and re-calibrate trust on automated systems to appropriate levels. Future work involves conducting experiments to study the effects of calibration of trust on use of automated systems.

Acknowledgments The authors would like to thank EPSRC (Grant EP/K011618/1) and WMG centre HVM Catapult for their support for this research.

References

1. Singh, S.: Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey. (Traffic Safety Facts Crash•Stats. Report No. DOT HS 812 115). Washington, DC (2015)
2. Nakajima, T.T.: Advanced Driver Assist System: Its Challenges and Solutions from the Customer's Point of View. In: SAE Convergence 2008, Detroit, Michigan. SAE Paper No. 2008-21-0031. Convergence Transportation Electronics Association, USA (2008)
3. Lee, J., Moray, N.: Trust, control strategies and allocation of function in human-machine systems. Ergonomics **35**, 1243–1270 (1992)
4. Muir, B.M.: Trust in automation: part I. Theoretical issues in the study of trust and human intervention in automated systems. Ergonomics **37**, 1905–1922 (1994)
5. Muir, B.M.: Trust between humans and machines, and the design of decision aids. Int. J. Man Mach. Stud. **27**, 527–539 (1987)
6. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**, 50–80 (2004)
7. Bainbridge, L.: Ironies of automation. Automatica **19**, 775–779 (1983)
8. Sheridan, T.B.: Human centered automation: oxymoron or common sense? In: 1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century, pp. 823–828 (1995)
9. Fitts, P.M., Chapanis, A., Frick, F.C., Garner, W.R., Gebhard, J.W., Grether, W.F., Henneman, R.H., Kappauf, W.E., Newman, E.B., Williams, A.C.J.: Human engineering for an effective air—navigation and traffic—control system. Washington, D.C., USA (1951)
10. Muir, B.M., Moray, N.: Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics **39**, 429–460 (1996)
11. Lewandowsky, S., Mundy, M., Tan, G.P.A.: The dynamics of trust: comparing humans to automation. J. Exp. Psychol. Appl. **6**, 104–123 (2000)
12. Rajaonah, B., Anceaux, F., Vienne, F.: Study of driver trust during cooperation with adaptive cruise control. Trav. Hum. **69**, 99 (2006)
13. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse. Abuse Hum. Factors **39**, 230–253 (1997)
14. Sparaco, P.: Airbus seeks to keep pilot, new technology in harmony. Aviat. Week Space Technol. **142**, 62 (1995)
15. BEA: Final report: on the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF 447 Rio de Janeiro, Paris (2012)
16. Rajaonah, B., Tricot, N., Anceaux, F., Millot, P.: The role of intervening variables in driver-ACC cooperation. Int. J. Hum. Comput. Stud. **66**, 185–197 (2008)
17. Riley, V.: A general model of mixed-initiative human-machine systems. In: Proceedings Human Factors Ergonomics Society 33rd Annual Meeting, pp. 124–128 (1989)
18. Riley, V.: Operator reliance on automation: theory and data. In: Parasuraman, R., Mouloua, M. (eds.) Automation and Human Performance: Theory and Applications, pp. 19–35. Lawrence Erlbaum Associates, Publishers (1996)
19. Parasuraman, R., Molloy, R., Singh, I.L.: Performance consequences of automation—induced "Complacency". Int. J. Aviat. Psychol. **3**, 1–23 (1993)
20. Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., Beck, H.P.: The role of trust in automation reliance. Int. J. Hum. Comput. Stud. **58**, 697–718 (2003)
21. Seppelt, B.D., Lee, J.D.: Making adaptive cruise control (ACC) limits visible. Int. J. Hum Comput Stud. **65**, 192–205 (2007)
22. McClumpha, A.J., James, M., Green, R.G., Belyavin, A.J.: Pilots' attitudes to cockpit automation. In: Proceedings of the Human Factors and Ergonomics Society 35th Annual Meeting, pp. 107–111. San Francisco, California (1991)

23. SAE International: Surface Vehicle Information Report, J3016: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems, USA (2014)
24. Endsley, M.R.: Automation and situation awareness. Autom. Hum. Perform. Theory Appl., pp. 163–181 (1996)
25. Couch, L.L., Jones, W.H.: Measuring levels of trust. J. Res. Pers. **31**, 319–336 (1997)
26. Rempel, J.K., Holmes, J.G., Zanna, M.P.: Trust in close relationships. J. Pers. Soc. Psychol. **49**, 95–112 (1985)

Cooperative Guidance and Control in Highly Automated Convoys—StrAsRob

Marcel C.A. Baltzer, Claudia Rudolph, Daniel López and Frank Flemisch

Abstract As vehicles become more and more intelligent and more and more technology to drive highly automated becomes affordable, concepts to integrate humans and these highly automated vehicles become an integral part of research and development. More functionality also increases the complexity to interact with such systems. Therefore, interaction concepts to cooperatively drive with a highly automated vehicle should focus on central goals: (1) ensure the safety and performance of the overall system, (2) establish a simple and comprehensible interaction between human and automation, (3) create acceptance for the automation system and interface. A study was conducted over the past three years in the field of highly automated military truck convoys at the Fraunhofer FKIE. In this project, FKIE focused on the formulation of use cases and the ergonomic design and evaluation. This paper presents the findings.

Keywords Human machine systems · Cognitive system engineering · Decision support systems · Highly automated trucks · Human-systems integration · Shared control · Cooperative control

M.C.A. Baltzer (✉) · C. Rudolph · F. Flemisch
Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Fraunhoferstrasse 20, 53343 Wachtberg, Germany
e-mail: marcel.baltzer@fkie.fraunhofer.de

C. Rudolph
e-mail: claudia.rudolph@fkie.fraunhofer.de

F. Flemisch
e-mail: frank.flemisch@fkie.fraunhofer.de; f.flemisch@iaw.rwth-aachen.de

D. López · F. Flemisch
Institute of Industrial Engineering and Ergonomics IAW, RWTH Aachen University, Bergdriesch 27, 52062 Aachen, Germany
e-mail: d.lopez@iaw.rwth-aachen.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_46

1 Introduction

Military vehicles like heavy trucks, tanks or excavators face the challenge of two contradicting goals: To be sufficiently guarded against enemy fire and to be safely maneuverable. Transport vehicles are the backbone of the supply troop. The assessment of safety, performance, comfort, energy efficiency and joy of use in combat vehicles is different to the assessment in the civil sector. Especially factors like combat strength, transport performance and safety against enemy fire play an essential role. Without suitable measures, those factors come at the expense of comprehensibility and accordingly safety and comfort. Furthermore, the German Federal Armed Forces face the demographic change and faces more and more complex missions with less and less staff.

In the past decades, an incremental revolution of assisted and automated driving takes place in the civil sector, also due to military research e.g. at the University of the German Federal Armed Forces or the DARPA. Even if an identical retransfer of knowledge from the civil to the military sector is not possible, a systematic translation can improve the military capacity to act.

StrAsRob, “Road transport with assistance by robots” is such a systematic research and development of assistance and automation systems—in this case for military trucks that can follow each other automatically (see Fig. 1). Whereas a long-term objective of the German Federal Armed Forces is to build convoys with a manned leading vehicle and unmanned following vehicles, StrAsRob still considers a safety driver. On the one hand, this safety driver can take over in different assistance and automation modes as well as drive manually. On the other hand, the driver will be allowed to get out of the control loop in certain clearly defined situations and take over other non-driving tasks.

In this paper, especially the interaction between driver and automation is addressed. Since R&D of the human-machine interaction profits from basic research of the past decades, we will give a short overview.

Fig. 1 StrAsRob target domain: highly automated military truck convoy

1.1 State of the Art

Assistance and Automation Systems. Although new technology is usually first developed in the military domain and later applied in the civil industry, e.g. aviation and communication, there has been important research in both military and civilian domain towards highly automated driving. At first, driving assistance systems have individually demonstrated their benefits by informing or taking over certain driving tasks and consequently reducing the workload on the driver and increasing comfort. Technical systems such as Adaptive Cruise Control (ACC) and Lane Keeping Assistance System (LKAS) significantly improve safety by constantly monitoring the traffic situation and by having a faster reaction time than that of any human.

By combining different assistance systems, an automation capable of taking over the lateral and longitudinal control of a vehicle can be developed. This has been explored since the mid-80s by a number of projects e.g. Eureka Prometheus Project that culminated in the creation of advanced test vehicles in the mid-90s, the VaMP and the VITA 2 [1, 2]. These vehicles were capable of detecting the surrounding traffic conditions and controlling the steering and velocity accordingly. The performance of an automated vehicle is dependent on the capabilities of the sensors and controllers used. This was shown by the Argo Project [3], which aimed at the development of a suitable vision and control system that could be implemented in a normal vehicle.

In the United States, the DARPA through its Grand Challenges has motivated the innovation in this same area. In 2004 and 2005, the Grand Challenges were aimed at the creation of a fully autonomous vehicle capable of completing a 142 mile uncharted desert course. The winner was able to complete the course in 6 h 53 min [4]. Furthermore, a similar challenge was held in 2007, this differed in that it took place in an urban environment. Consequently, the competing vehicles had to obey road rules and complete the course in less than 6 h. This time the winner vehicle, named Boss, was able to complete the challenge in 4 h 10 min [5].

Cooperative Guidance and Control. In parallel with the progressive improvement of the technical part or an automated vehicle, the cooperation and interaction of human drivers with such systems has also been explored. The cooperation between human and machine can be explained with Fig. 2. This cooperation between human and machine has already been described by Rasmussen [6] and in research further refined by e.g. [7–9] or [10].

Flemisch et al. [8] proposes a cooperation between driver and automation based on the H(orse)-metaphor. The H-Metaphor compares the interaction between driver and automation with that of a horse and a rider and by using a suitable interface. Based on this metaphor, a haptic-multimodal mode of interaction “H-Mode” has been developed, and the basic idea of a joint, cooperative system applied to real vehicles e.g. in HAVEit [12]. In HAVEit different European countries and industry partners worked together, to define the next steps in the development of automated vehicles with an important part of the project being the human-machine interaction.

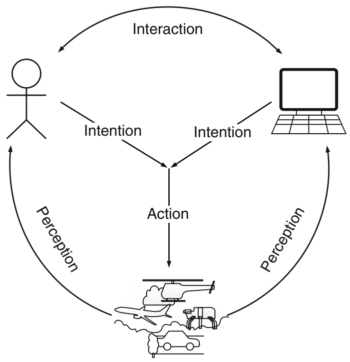


Fig. 2 Human-machine cooperation in cooperative guidance and control [11] (cf. [8])

An important realization from the abovementioned research is to include the human in the development process in order to integrate all of his useful abilities and that he can cooperate efficiently and safely with the automation (e.g. [13]). Furthermore, perspectives that go beyond human-centered design are necessary, in order to weigh up human, technical and organizational factors.

2 Assistance and Automation Modes in StrAsRob

The degree of influence of human or automation on the driving task may vary. This variation can be organized in assistance and automation modes.

Inspired by Gasser et al. [14], National Highway Traffic Safety Administration [15], SAE [16], StrAsRob has a spectrum of four different Assistance and Automation Modes (AAM), shown in Fig. 3.

In AAM1, the military driver is responsible and takes over lateral and longitudinal driving control and is assisted by various passive assistance systems, e.g. blind spot monitoring. Furthermore, the driver can activate or deactivate an emergency brake system. In AAM2, the automation actively takes over the longitudinal driving

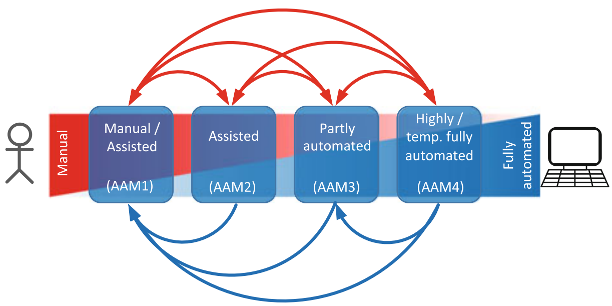


Fig. 3 Assistance and automation spectrum in StrAsRob

tasks and the military driver takes over the lateral driving tasks. He is still responsible for the overall system. In AAM3, the military driver has only a supervisory task and the automation controls the lateral and longitudinal driving tasks. The major difference between AAM3 and AAM4 is that the military driver can leave the control loop in AAM4 and can focus on other non-driving tasks.

3 The User Interface—Interaction Between Driver and StrAsRob-System

Since the different AAM come along with a different driver's level of attention, a transition from one mode to another must be secured differently. From manual control to temporary fully automated without the driver's supervision, it is of major importance that the vehicle is always sufficiently controlled.

Therefore, the driver must always know what tasks and what responsibility he has and, respectively, he must know what tasks and responsibility the automation has. Furthermore, the driver should be aware what the current capabilities of the automation are and whether it is able to take over the current driving tasks and responsibility. This awareness is called "Automation Awareness" and can be seen as part of Situation awareness (cf. [17–19]). In this context, an established term is "Mode Confusion" (e.g. [20]) and represents a special characteristic of Automation Awareness, where the user of a system thinks he is in a different automation mode than the currently active. If the military driver shows such awareness and does not feel responsible for certain tasks, a control deficit may occur. In this case, the driving tasks may be addressed insufficiently and the danger of a crash increases. The same applies to the opposite of control deficit, the so-called control surplus. Such control surplus can lead to control conflicts and needs to be avoided. A more detailed discussion can be found in [21].

In a series of participatory design workshops, a selected group of ergonomic experts, system designers, military drivers and driving teachers of the German Federal Armed Forces addressed topics of assistance and automation spectrum, transitions and automation awareness and developed a user interface. The following description of the user interface represents the final state of the iterative participatory design process.

Next to using conventional input options like steering wheel, gas and brake pedal, the driver can also interact with the StrAsRob system via buttons at the operation and display device (ODD) (see Fig. 4).

Basic requirements were that buttons needed to be designed in a way to avoid accidental input. Furthermore, the ODD needs to be usable in situations of strong concussions and movement. Therefore, in StrAsRob, the military driver selects a mode via large hardware buttons. For automation initiated transitions, e.g. when the current mode becomes unavailable, the driver needs to confirm such a mode transition with the Master Caution button flashing yellow in the respective situation.

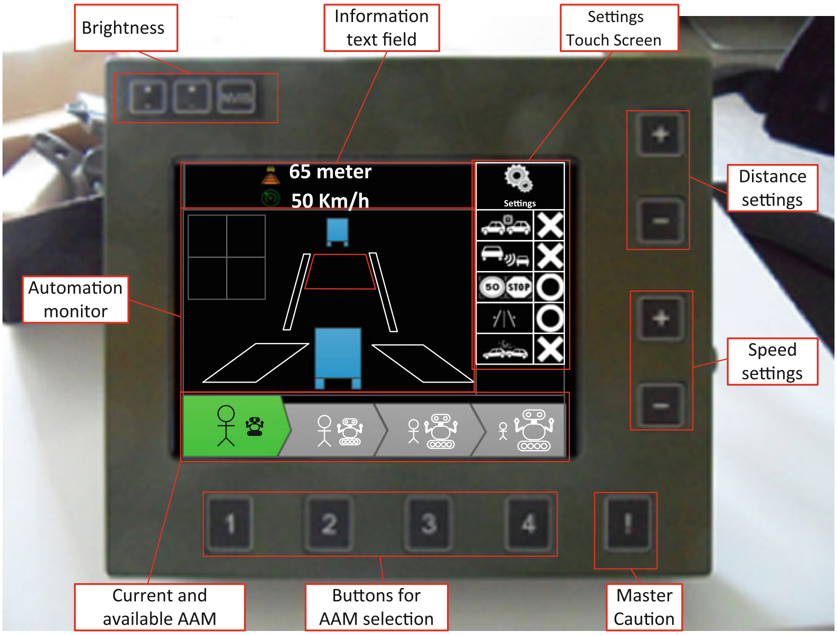


Fig. 4 The operation and display device (ODD) of StrAsRob

In order to improve the automation awareness, the current mode is highlighted in green and the respective symbol enlarged. Unavailable modes are shown in very dark grey and available modes in light grey.

4 Simulator Study

To test the user interface, a usability study was conducted in the driving simulator at FKIE. The focused research questions were: "How is the usability of the automation and interaction concept?", "How does the interaction between driver and StrAsRob system change from naïve over explained to trained run and what is the deducible training?", "How do drivers evaluate the impact of the StrAsRob system on system qualities like safety and performance?" and "Is Take-Over capability in the current implementation of the interaction concept ensured?".

4.1 Use Cases

In the following, an excerpt of the developed use cases for the interaction concept will be presented: "Use Case 3a—Start in follow mode" describes the starting. In "Use Case 3b—Standard following" maintaining the distance to the preceding vehicle of StrAsRob is structured. "Use Case 3c—Stop to standstill in follow mode" describes the stopping of the convoy while the follow mode is active. "Use Case 4a—Avoid obstacles in follow mode" describes how moving as well as stationary obstacles are avoided in StrAsRob. The "Use Case 4b—Overtaking on a motorway" was chosen explicitly because of the insufficient range of the sensors. For example, on a countryside highway, oncoming traffic would have to be considered. "Use Case 5a—Turning in follow mode" deals with the subject of turning maneuvers, where the leading and the following vehicle should have at least similar trajectories while turning. If the leading vehicle is out of sight after turning, the StrAsRob continues driving in a combination of LKAS and ACC. If the leading vehicle is out of sight during the turning process, a transition into AAM1 is necessary and the driver should catch up with the leading vehicle manually. "Use Case 6a—Split of the convoy in follow mode" addresses dangerous situations (for example hostile aircraft) so that the convoy becomes a less vulnerable target. The military driver has to take over the control and has to deliberately change to another lane. Following to a split up maneuver, "Use Case 6b1—Join with the leading vehicle" structures the following join maneuver if the danger has passed. If a foreign vehicle moves into the convoy, "Use Case 7—Foreign vehicle between leading and following vehicle", the StrAsRob subsystem autonomously transitions to a combination of LKAS and ACC and will follow the foreign vehicle with a safe distance. If the foreign vehicle leaves the convoy, the follow mode can be continued.

The use cases were combined to testing and training procedures, the so-called Kata, described in the next section.

4.2 Testing Procedures—Kata

The term "Kata" originates from the Japanese martial arts and is used to describe a sequence of moves used for competition or training. In our study, we use the word "Kata" to define a fixed set of Use Cases, which are combined and structured in a test and training form. To analyze the usability of the interaction concept in the Use Cases two fixed sequences/"Katas" were created and tested in the simulator.

Kata Base. The Kata Base is focused on the basic interaction between human and co-system. Here "Use Case 3a—start in follow mode" and "Use Case 3b—standard following" are analyzed. Moreover, it is studied how well various transitions work, whether they were initiated by the driver or by the co-system. There are three major sections to be considered: First, the start-up procedure is analyzed, followed by the driver-initiated (Di) transitions. Finally, the automation initiated (Ai) transitions are tested for usability.

Kata Advanced. In the Kata Advanced the interaction between human, co-system and environment is examined, therefore the degree of complexity is increased. "Use Case 7—Foreign vehicle between leading and following vehicle", "Use Case

5a—Turning in follow mode", "Use Case 4a—Avoiding obstacles in follow mode", "Use Case 4b—Overtaking on a motorway", "Use Case 6a—Split of the convoy in follow mode" and "Use Case 6b1—Join with the leading vehicle" are implemented.

The participants were instructed to stay on the road as accurately as possible. During the experiment, when participants were asked to switch to AAM4, they distracted themselves by reading a newspaper.

4.3 Description of the Simulator Prototype

The experiments were carried out in the simulator at the Fraunhofer FKIE (see Fig. 5).

The simulator consists of two side-by-side driving simulators. On the left, the participant simulator consists of three projectors and screens that provide the front, right and left view (Fig. 5, left). These screens are arranged at 90° so they provide a 180° view of the map. Small 11" LCD vertical displays are used for the left and right mirrors and positioned at an offset from the driver. Five separated computers running the SILAB simulation software controlled each an individual view. The sitting position and the orientation of the controls were set to better resemble that of a real truck (Fig. 5, right). Still, adjustments were possible so each participant could find a comfortable position.

The active steering wheel provides the force feedback for a more realistic feeling. The pedals are passively spring loaded and able to sense position. Two pedals are used for the simulation of an automatic transmission. Both steering wheel and pedals use a CAN-BUS interface. A military hardened Panther DK10 tablet PC from Roda Solid IT solutions positioned on the right side of the steering wheel. It was used as ODD. As an addition to the data logging from the simulation and the steering and pedal inputs, two cameras (Fig. 5) recorded each test.

Next to the driver's simulator, a confederate simulator was placed to the right and hidden behind a curtain. The confederate emulated the automation similar to the

Fig. 5 Display setup of the driving simulator. *Left* Overhead camera. *Right* Driving controls and automation display

wizard-of-oz-method [22]. Since immersion in the simulation is not high priority for the confederate, the left and right views were omitted. Left and right mirror were provided as well as an information screen displaying the automation state and vehicle speed. The driving controls, steering wheel and pedals were identical to those used by the participant, although the sitting position and control orientation had no special requirements. Data logging was also performed for the confederate with the exception of the recording cameras.

4.4 Participants and Experimental Procedure of the Study

The simulator study was organized with 12 participants (11 male, 1 female), all of them trained soldiers of the German Federal Armed Forces. At the time of the study, the participants were from 24 to 52 years old and owned driver's licenses for 6–34 years. Of those licenses, 9 also included a permission for driving trucks and 6 for driving motorcycles. The participants' yearly mileage ranges from 5000 km to more than 50,000 km. One participant uses his vehicle 1–2 times per week; the other 11 use it daily. When asked to describe their own way of driving, answers range from safe/experienced to dynamic/sporty.

At the beginning, every test person was welcomed to the FKIE-Exploroscope and the background of the study was explained to him or her. Afterwards they had to complete a demographic questionnaire followed by another questionnaire about simulator sickness. After a participant took the seat in the simulator, the seat height, recline etc. was optimally adjusted to him or her. In order to get used to the simulator the participants made a familiarization drive on a route with highways, countryside roads and city traffic. Directly after the test drive the participants had to evaluate how closely the simulation resembled reality.

The first Kata to drive was the Kata Base. Before the first run, the participants received only little explanation to ensure a naïve attitude. After this run, they had to evaluate the interaction forms on usability and got a detailed explanation of the interaction concept. With this knowledge, they did a second run of the same scenario. This time the simulation was paused twice and the corresponding questionnaire had to be completed by the participants. For another run in the same scenario, the participants received an extensive training in the interaction concept. It was the last run of the Kata Base and again, it was paused twice to fill out the questionnaires. Then the participants had to complete yet another questionnaire about simulator sickness.

After a short break, the participants drove the Kata Advanced in the same way as the Kata Base. Since the participants lost their naïve attitude after the trained run in the Kata Base, it was not possible to evaluate a naïve run again. Therefore, the first run was counted as an explained run.

4.5 Evaluation of the Simulator Study

Familiarization Drive. The results of the familiarization drive are neutral to rather good. The reason for such a reserved evaluation is the limited immersion factor of the simulator, which is normal in a static simulator (e.g.: Participant 4: The vehicle is braking visually, but the reaction cannot be felt).

Kata Base. The results of the Kata base runs are displayed in Fig. 6.

In the first, naïve run the participants drove the first two blocks of the Kata Base without any explanation of the different automation modes or the emergency transitions by steering or braking. As we can see in Fig. 6, the results were already quite good. The rather bad evaluations can be seen for the transition AAM3 to AAM4 and vice versa. From the remarks of the participants this is explainable by the limited distinguishability: In the naïve run, the participants did not know about the shift of responsibility since both modes "felt" the same. Therefore, a design adaption can help to intuitively understand such a transition. A possibility is a hands-on or hands-off symbol in the ODD.

As we can see in Fig. 6, the evaluation of the interaction design got better after explaining it. Still the transition from AAM4 to AAM1 via steering intervention got worse to neutral. The remarks criticized it to be dangerous, since the participants sometimes departed from the current lane. An explanation why this transition got even worse evaluation from naïve to explained run can be that not all participants made such a transition in the naïve run. Those that did, tried it on purpose and were mostly pleased that they were able to intervene even in partly or highly automated modes. The quite large steering torque to initialize the emergency transition was a design decision to prevent triggering it by mistake but to only make it available so the driver can still evade an obstacle that the automation did not detect. Therefore, the test design should be adjusted in order to use such an emergency transition in an emergency situation only.

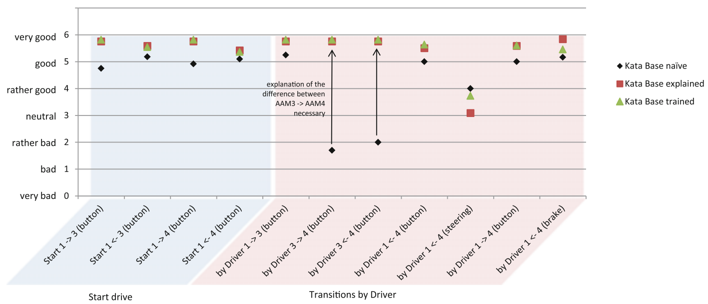


Fig. 6 Comparison of the kata base runs

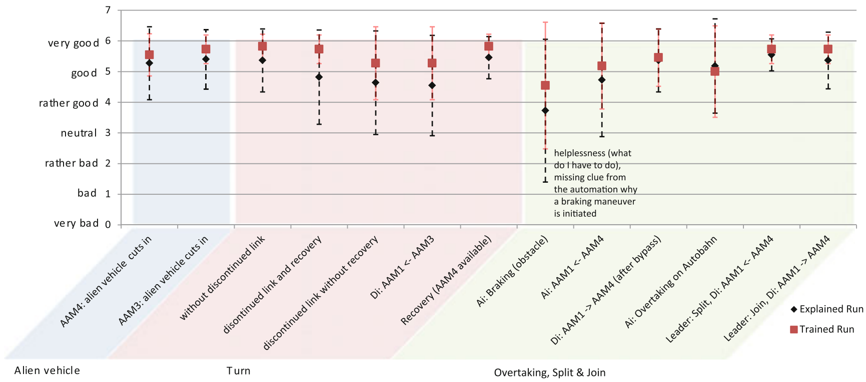


Fig. 7 Comparison kata advanced

By contrast, the emergency transition by pedal with a deflection of 30 % was accepted quite well. Transitions from AAM4 to AAM3 or AAM3 to AAM1 due to automation failure need an acoustic warning.

A further improvement of the interaction design's rating comes from an extensive training. Still, it is evident that the emergency transition by steering intervention needs the abovementioned adjustments. Comparing the differences of the evaluation of AAM3 to AAM4 and vice versa, it becomes clear that the self-explanatory character of the interaction design needs further attention.

Kata Advanced. The results of the Kata Advanced are displayed in Fig. 7.

Since the participants already received a feeling for the interaction design in the Kata Base, only an explained and a trained run were driven in the Kata Advanced. All in all, the results were promising as we can see in Fig. 7. Still the participants wanted to have more information about the current state of the automation, e.g. the link with the leading vehicle and why the automation suddenly initiated a braking maneuver. Therefore, they did not understand fast enough that they had to take over and felt a bit helpless. Furthermore, most participants wanted to have more acoustic warning when such a link broke and an acoustic information when the link was recovered.

The extensive training had again a positive influence on the evaluation of the interaction design, especially in the situations of the broken link to the preceding vehicle. Still, an improvement of the interaction design in such situations should be addressed.

5 Conclusion

StrAsRob is one of multiple possible manifestations of cooperative guidance and control of highly automated vehicles in the military sector. It takes the assistance and automation levels of the civil sector as an orientation, where an important shift of responsibility lies especially between partially and highly automated. The hypothesis that such an assistance and automation spectrum can be translated in the military sector and for convoys was completely confirmed. The overall usability was rated good to very good although improvements in the interaction design need to be addressed, such as (1) transitions from AAM4 to lower AAM should come along with an acoustic signal (2) AAM4 and AAM3 need to be more distinguishable and (3) emergency transitions need an extensive design focus.

Therefore, the proof-of-concept was successful, but before series development should be addressed, further questions need to be answered. Due to capacity constraints only essential modes were investigated.

At the end of the presented project it should be emphasized, that the participative approach for the design of Use Cases and the interaction concept by bringing together user, developer, researcher, operator, teacher and administration was one of the essential keys to success. Especially when addressing complex assistance and automation systems, existing risks can be controlled from the very beginning.

Acknowledgments This project was funded by the Federal Office of Bundeswehr Equipment, Information Technology and In-Service Support (BAAINBw). Furthermore, we want to acknowledge the good cooperative work together with the University of the German Federal Armed Forces Munich, Diehl (DBD), Rheinmetall (RLS) and the Logistic school of the German Federal Armed Forces.

References

1. Maurer, M., Behringer, R., Fürst, S., Thomanek, F., Dickmanns, E.-D.: A compact vision system for road vehicle guidance. In: Proceedings of the 13th International Conference on Pattern Recognition, vol. 3, pp. 313–317 (1996)
2. Dickmanns, E., Behringer, R., Dickmanns, D., Hildebrandt, T., Maurer, M., Thomanek, F., Schiehlen, J.: The seeing passenger car 'VaMoRs-P'. In: Proceedings of the Intelligent Vehicles '94 Symposium (1994)
3. Broggi, A., Bertozzi, M., Fascioli, A., Guarino Lo Bianco, C., Piazzi, A.: The ARGO autonomous vehicle's vision and control systems. Int. J. Intell. Control Syst. **3**(4), 409–441 (2000)
4. Thrun, S., Montemerlo, M., Dahlkamp, H., Stavens, D., Aron, A., Diebel, J., Fong, P., Gale, J., Halpenny, M., Hoffmann, G., Lau, K., Oakley, C., Palatucci, M., Pratt, V., Stang, P., Strohband, S., Dupont, C., Jendrossek, L.-E., Koelen, C., Markey, C., Rummel, C., Niekerk, J. V., Jensen, E., Alessandrini, P., Bradski, G., Davies, B., Ettinger, S., Kaehler, A., Nefian, A., Mahoney, P.: Stanley: the robot that won the DARPA grand challenge. J. Field Robot. **23**(9), 661–692 (2006)
5. Urmson, C., Anhalt, J., Bagnell, D., Baker, C., Bittner, R., Dolan, J., Duggins, D., Ferguson, D., Galatali, T., Geyer, C., Gittleman, M., Harbaugh, S., Hebert, M., Howard, T., Kelly, A., Kohanbash, D., Likhachev, M., Miller, N., Peterson, K., Rajkumar, R., Rybski, P., Salesky, B., Scherer, S., Woo-Seo, Y., Simmons, R., Singh, S., Snider, J., Stentz, A., Whittaker, W., Ziglar, J.: Tartan racing: a multi-modal approach to the DARPA urban challenge. DARPA Technical Report (2007)
6. Rasmussen, J.: Skills, rules and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Trans. Syst. Man Cybern. **3**(3), 257–266 (1983)

7. Hoc, J.-M.: From human—machine interaction to human—machine cooperation. Ergonomics **43**(7), 833–843 (2000)
8. Flemisch, F.O., Adams, C.A., Conway, S.R., Goodrich, K.H., Palmer, M.T., Schutte, M.C.: The H-Metaphor as a guideline for vehicle automation and interaction. Report No. NASA/TM-2003-212672. NASA Research Center, Hampton (2003)
9. Biester, L.: Cooperative automation in automobiles. Doctoral Dissertation, Humboldt-Universität zu Berlin (2008)
10. Flemisch, F.O., Bengler, K., Bubb, H., Winner, H., Bruder, R.: Towards a cooperative guidance and control of highly automated vehicles: H-Mode and conduct-by-wire. Ergonomics **57**(3), 343–360 (2014)
11. Flemisch, F.O., Heesen, M., Kelsch, J., Schindler, J., Preusche, C., Dittrich, J.: Shared and cooperative movement control of intelligent technical systems: sketch of the design space of haptic-multimodal coupling between operator, co-automation, base system and environment. In: 11th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, Valenciennes, France (2010)
12. Hoeger, R., Zeng, H., Hoess, A., Kranz, T., Boverie, S., Strauss, M., Jakobsson, E., Beutner, A., Bartels, A., To, T.-B., Stratil, H., Fürstenberg, K., Ahlers, F., Frey, E., Schieben, A., Mosebach, H., Flemisch, F.O., Dufaux, A., Manetti, D., Amditis, A., Mantzouranis, I., Lepke, H., Szalay, Z., Szabo, B., Luithardt, P., Gutknecht, M., Schömig, N., Kaussner, A., Nashashibi, F., Resende, P., Vanholme, B., Glaser, S., Allemann, P., Seglö, F., Nilsson, A.: HAVEit Deliverable D61.1, Final Report (2011)
13. Flemisch, F.O., Kelsch, J., Löper, C., Schieben, A., Schindler, J.: Automation spectrum, inner/outer compatibility and other potentially useful human factors concepts for assistance and automation. In: de Waard D., Flemisch F.O., Lorenz B., Oberheid H., Brookhuis K. (eds.) Human Factors for Assistance and Automation, pp. 1–16. Shaker, Maastricht (2008)
14. Gasser, T.M., Arzt, C., Ayoubi, M., Bartels, A., Bürkle, L., Eier, J., Flemisch, F., Häcker, D., Hesse, T., Huber, W., Lotz, C., Maurer, M., Ruth-Schumacher, S., Schwarz, J., Vogt, W.: Rechtsfolgen zunehmender Fahrzeugautomatisierung - Gemeinsamer Schlussbericht der Projektgruppe. (F 83) Fahrzeugtechnik, Bundesanstalt für Straßenwesen (BASt) (2012)
15. National Highway Traffic Safety Administration: Preliminary statement of policy concerning automated vehicles (2012)
16. SAE J 3016, Society of Automotive Engineers SAE: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (2014)
17. Sarter, N., Woods, D.: How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Hum. Factors **37**(1), 5–19 (1995)
18. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Factors **37** (1), 32–64 (1995)
19. Endsley, M.R.: Automation and situation awareness. In: Parasuraman R., Mouloua M. (eds.) Automation and Human Performance: Theory and Applications, pp. 163–181. Lawrence Erlbaum, Mahwah (1996)
20. Sarter, N., Woods, D., Billings, C.: Automation Surprises. In: Salvendy, G. (ed.) Handbook of Human Factors, Ergonomics. Wiley, New York (1997)
21. 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**(1), 3–18 (2012)
22. Kelley, J.F.: An iterative design methodology for userfriendly natural language office information applications. ACM Trans. Office Inf. Syst. **2**(1), 26–41 (1984)

Autonomous Vehicles in Developing Countries: A Case Study on User's View Point in Pakistan

Irum Sanaullah, Amjad Hussain, Amna Chaudhry, Keith Case and Marcus Enoch

Abstract Technological advancements are continuously changing the human life. Like many other developments, autonomous vehicle system is attracting public interest and being widely discussed by all the stakeholders. Recent reports show that in future, autonomous vehicles or self-driving cars will be on roads in developed countries such as in UK and US. In this age of information technology, advancements made in developed countries not only move to the developing countries but also impact the opinions and lives of the people living in these countries. Therefore, there is a need to develop more effective strategies which can help the adaption of upcoming technologies in transport systems like autonomous vehicles. In this respect, user's perception becomes highly significant as this can help designers by providing them the information about real time issues and human observations. Up till now, no significant work has been carried out on exploring the user's perception about autonomous vehicles in developing countries like Pakistan. This study aims at capturing the user's view point about the use of autonomous vehicles which can provide relevant information on perceived benefits and chal-

I. Sanaullah · A. Chaudhry
Department of Transportation Engineering and Management,
University of Engineering and Technology, Lahore, Pakistan
e-mail: Irum.sanaullah@uet.edu.pk

A. Chaudhry
e-mail: aaminah.ch@gmail.com

A. Hussain (✉)
Department of Industrial and Manufacturing Engineering,
University of Engineering and Technology, Lahore, Pakistan
e-mail: chamjad@uet.edu.pk

K. Case
Wolfson School of Mechanical and Manufacturing Engineering,
Loughborough University, Loughborough, UK
e-mail: k.case@lboro.ac.uk

M. Enoch
Department of Civil and Building Engineering, Loughborough University,
Loughborough, UK
e-mail: m.p.enoch@lboro.ac.uk

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_47

lenges in user's perspective. An underlying objective is also to compare this perspective with developed countries like US, UK, and Australia etc. Findings of the study will help in assessing user's perceptions in terms of challenges, the level of awareness and understanding about autonomous vehicles. It will lead to shaping up the strategies to address the needs of users so that more viable and equally acceptable technological interventions can be made.

Keywords Autonomous vehicle · User's perception · Human driven vehicles

1 Introduction

This paper explores user's perception about the use of autonomous or self-driving vehicles in Pakistan. The focus of the study is to analyze the factors of awareness, reliability, safety, convenience and comfort of the autonomous vehicle system in the user's point of view.

2 Background Literature

In the last century, automotive sector has made advancements to improve the cost, comfort and safety of vehicles [1]. Autonomous Transport system is one of these developments defined as the unmanned system, which provides self- directed and independent transfer of people and goods without human involvement. This system includes the controlling of driverless vehicles, movement of intelligent vehicles and independent navigation [1]. Autonomous vehicle is known as 'driverless car', which can plan and drive by itself by detecting the painted lines or a magnetic monorail implanted in the road. It is also named as autopilot or automated guided vehicle (AGV) [2].

The autonomous vehicle consists of Visual Guidance System (VGS), Vehicle Control System (VCS), Vehicle Piloting System (VPS) and Robust Communications System (RCS) [3]. Studies suggest that advantages of Autonomous Transport Systems include increase in road and parking capacity, improved safety, efficient traffic flow, and decrease in traffic congestion and fuel economy. Autonomous vehicles are small in size, which can travel on roads with the smaller gaps between them and can increase the road capacity up to 30 % [2, 4]. The capacity on highways can increase by 3–8 times, if the autonomous vehicles travel in platoons. It is also possible for vehicles to drive bumper to bumper at full speed if the spacing is eventually reduced with the perfect connected device [5]. Driverless cars can drop off the passengers and can park themselves far away where the space is available and can return when they are required to pick up the passengers. In this way an autonomous transportation system can eliminate the problem of parking scarcity [6].

As human error cause 90 % of the traffic accidents, the use of autonomous vehicles can significantly reduce the number of crashes. Folsom (2011) carried out a comparison of accident rates of motor vehicles to those of autonomous vehicle (commuter rail) system in the city of Lille in France [7]. The study concluded that motor vehicles are 33 times as risky as autonomous trains on a reserved path. The Vancouver autonomous Sky Train System in Canada gave the same accident rate facts as the autonomous vehicle system in Lille gave in the comparison with Light Rail Transit (LRT) or Rapid Rail Transit (RRT) systems. The application of autonomous system can reduce the number of accidents due to its increased reliability and precision as compared to human drivers and therefore it is an effective solution for drunk driving. Also with the use of autonomous vehicles, the number of automated-related casualties can be decreased to halve (1.2 million) per year overall the world [1, 8].

In the start of autonomous system application, the autonomous vehicles will drive in combination with human driven vehicles. In this case autonomous vehicles will follow the traffic laws and human drivers will have the option to follow the laws or not. This can create the state of confusion and problem [6]. However with the fully autonomous system the vehicles will travel at design speed on the main line, and will only change the speed on exit or entry ramps. The vehicles in system will continue to drive at full speed if the system is saturated and new vehicles will not enter the system. The interchanges will be assisted by parking buffers to provide the space for merging while vehicle changes their routes [7]. The vehicles with consistent speed on roads will result in the efficient traffic flow and less traffic jams.

Drivers need gap between vehicles while driving for perception and reaction time to apply breaks in emergency situation. In autonomous vehicles, due to central computer control system the following vehicle will know about the leading vehicle that when it will decelerate, take turn and accelerate before it takes the action. The perception time will be few milliseconds and therefore it can decrease the safe following distance between vehicles. The computer-controlled vehicles would drive close together on highways due to the awareness and prompt response; hence traffic congestion can be reduced [7].

Wu et al. [9] carried out a study on fuel economy optimization system in human driven and autonomous vehicles. They concluded that the application of optimal model in autonomous vehicles will be more helpful to increase the fuel economy as compared to the human driven vehicles.

With all the advantages of autonomous vehicles mentioned above, there are perspective challenges and issues which need to be addressed.

Firstly it is necessary to make drivers to accept that car can have a drive independently. Naturally, it is difficult for a traditional driver to trust that car will apply the brakes when required in the case of accident. It might take years to convince the society to trust driverless car operating in rush hours traffic [9].

The social issues related to the autonomous system include:

- Will people trust using driverless car?
- Will policy makers allow the driverless cars on roads?

- Who will be responsible for mishaps and accidents?
- Will people have choice to drive by themselves when they want?

McCarthy [10, 11] states "The very nature of autonomous systems raises social, legal and ethical issues. People tend to be more accepting of a technology if they can choose whether or not to adopt it and have some control over its use".

The autonomous system is essentially outside the user's control of users; hence people's perception of the system can be negative. Therefore there is possibility of excluding the people who don't want to the part of system. Autonomous systems are required sometimes where humans might make bad decision, and therefore human dominance can be problematic. It leads to the query whether autonomous systems will be trusted more than humans in some situations. The slow provision of new features such as automatic parking and adaptive cruise control can make it easier for people to accept the autonomous system gradually (Stock, n.d.).

It is also important to convince the law makers to allow the autonomous cars on the road and to resolve issue of responsibility in the case of accidents.

DARPA urban challenge has proved that it is absolutely possible to imply fully autonomous cars to gain convenience, safety, and environmental benefits. However, the issues discussed above need to be solved by taking gradual steps to make people familiar with the system and gaining the acceptance by politicians and lawmakers.

3 Design of Survey Questionnaire and the Targeted Respondents

The survey questionnaire was prepared to get public opinion on the adaptation of self-driving vehicle technology in Pakistan. The factors that were considered for preparing the survey questionnaire included were mainly based on (i) Human Factors identified by the previous literature that impact the automation (ii) Roadway and Traffic Operational Factors such as safety and mobility.

The following factors mainly have been included for getting drivers' perception on the autonomous vehicle technology:

- Familiarity with the technology
- Interest
- Reliability
- Safety
- Stress
- Convenience
- Traffic Delays
- Roadway Safety
- Training Requirement

Additionally, questions were prepared for the drivers who have used some level of automation before such as ***Cruise Control System***. This was done to get an idea

Table 1 Demographics for 99 respondents

Demographics		Percent
Gender	Male Female	75.8 24.2
Age	20–30 years 30–40 years >40 years	56.6 27.3 12.1
Qualification	Graduate Technical Graduate others Doctorate	50.5 35.4 12.1
Driving experience	Less than 1 year 1–5 years 5–10 years >10 years	14.1 37.4 21.2 25.3

on the differences in opinions/perceptions that may exist between such users' and the other drivers.

In order to get more reasonable opinion on the aforementioned factors, mostly people from the technical background, holding either bachelor's degree or above, were targeted. People from non-technical background were also included in the sample for conducting the survey to get an idea on the differences in the perceptions. Different age groups of both male and female survey participants, that were considered, included: 20–30 years, 31–40 years, and more than 40 years as shown in the Table 1.

4 Results and Discussion

Most of the respondents were familiar with the autonomous vehicles as shown in the Table 2. The highest 88.9 % respondents had heard about the self-driving vehicles before, while 11.1 % respondents were unaware about it. As the respondents in this study were well educated and mostly (50.5 %) with the technical background of study, therefore it was expected they would be aware about the self-driving vehicles. The awareness level can be further explored particularly focusing on the group of people whose education level is lower than graduation.

Over all 74.7 % respondents expressed their interest to have self-driving vehicle and 25.2 % people were found not to be interested in having autonomous vehicles. It needs further investigation about the concerns of people who are not interested in owning the self-driving vehicle (Table 3).

Table 2 Awareness about self-driving vehicles

Frequency		Percent
No	11	11.1
Yes	88	88.9
Total	99	100.0

Table 3 Level of interest in having a self-driving vehicle

	Frequency	Percent	Cumulative percent
Very much interested	32	32.3	32.3
Interested	42	42.4	74.7
Not too much interested	22	22.2	97.0
Not interested at all	3	3.0	100.0
Total	99	100.0	

Table 4 Level of interest in having different levels of automation

	Frequency	Percent	Cumulative percent
No automation	5	5.2	5.2
Level 1 automation	8	8.2	13.4
Level 2 automation	26	26.8	40.2
Level 3 automation	33	34.0	74.2
Level 4 automation	25	25.8	100.0
Total	97	100.0	

When people were asked about the level of automation they wanted to have in their vehicles, about 34 % respondents expressed their interest to have level 3 automation in their vehicles following 26.8 and 25.8 % to have level 2 and level 4 respectively. It is represented in the Table 4. This depicts the psychology of user's that they want to have some sort of control (level 2 and 3) in their hands rather than totally relying on automation (such as in level 4). Similar kind of results were found in the study conducted in developed countries [12] where people were more concerned about using level 4 as compared to level 3.

Regarding the reliability of self-driving vehicle technology, 60 % respondents think that this technology would be moderately reliable and 16 % think it would be very reliable. While 13 % people expressed their views as an unreliable technology. The majority of the respondents (65.7 %) feel it would be safe to ride in the self-driving vehicles. While 34.4 % people disagree about feeling safe in self-driving vehicles. When people were asked "I may feel unsafe while riding in the self-driving vehicle due to possibility of system failure and collision". Most of the respondents (73.1 %) considered it unsafe in the case of system failure and accidents (Table 5).

Table 5 Perception about safety (system failure and collision)

	Frequency	Percent	Cumulative percent
Strongly disagree	4	4.1	4.1
Disagree	22	22.7	26.8
Agree	50	51.5	78.4
Strongly agree	21	21.6	100.0
Total	97	100.0	

Table 6 Perception about training requirements

	Frequency	Percent	Cumulative percent
Strongly disagree	3	3.1	3.1
Disagree	24	24.7	27.8
Agree	45	46.4	74.2
Strongly agree	25	25.8	100.0
Total	97	100.0	

Approximately 70 % people thought they need to monitor the roadway while driving in the self-driving vehicles though 28.3 % disagree with the concept of monitoring while sitting in the self-driving vehicle. Most of the people (59.6 %) expect that the self-driving vehicles will increase the road safety and decrease the road accidents. The higher percentage of respondents (67.7 %) is of view that self-driving vehicles could improve the traffic movement by minimizing the delays. When people were asked about feeling stressful while dealing with the self-driving technology, 53.6 % people disagree about this feeling. However, 44.5 % respondents express their concern about feeling stressful while interacting with the self-driving vehicles. Most of the people (72.2 %) think the intensive training will be needed to drive the self-driving vehicles (Table 6).

About the decrease in the driver's fatigue and work load, 83.9 % people agree it would happen in the case of self-driving vehicles. While small percentage of respondents (13.1 %) was not agreed with it. Similar kind of result was found when people were asked about the increase in travelling convenience in self-driving vehicles. Mostly (78.8 %) people think it will increase the convenience of travelling. When people were asked about the replacement of human drivers with the self-driving vehicles, interestingly 72.1 % people said that the self-driving vehicles can't replace the human drivers (Table 7).

Majority of the respondents (61.6 %) will not be comfortable riding a self-driving vehicle while interacting with human driven vehicles on road. The rest of the respondents (36.4 %) think it will be comfortable to come across with conventional vehicles on road. Approximately 84 % respondents think they will be more comfortable riding a self-driving vehicle on highways (Table 8).

If the perception about the above factors is compared with the cruise control users, majority of the users of cruise control (77.2 %) feel safe as compared to the

Table 7 Replacement of human drivers

	Frequency	Percent	Cumulative percent
Strongly disagree	3	3.1	3.1
Disagree	24	24.7	27.8
Agree	50	51.5	79.4
Strongly agree	20	20.6	100.0
Total	97	100.0	

Table 8 Comfort while using self-driving vehicle

	Frequency	Valid percent	Cumulative percent
Strongly disagree	1	1.0	1.0
Disagree	13	13.4	14.4
Agree	58	59.8	74.2
Strongly agree	25	25.8	100.0
Total	97	100.0	

non-user of cruise control (65.7 %). In a similar way higher percentage of cruise control users (77.3 %) said they don't feel stressful while using it. Also 90.9 % people said their workload is reduced and 86.3 % found it very convenient to use the cruise system.

Similar kind of opinion was found about the usage of automation for long distances. Overall 86.4 % respondents said they prefer to use cruise control system for long routes and highways.

5 Conclusion

Almost everybody has understanding about the latest vehicle technology (self-driving or autonomous vehicles) and people are interested in opting it. Majority is convinced about the benefits like increase in comfort, safety, convenience and reduction in delays and fatigue. However, training about the use of such system has been taken as the necessary requirement before the actual drive. Similarly, people have fear about the response of system in collision or system failure case. Cruise control users are relatively more satisfied and have more confidence, that can be because of the fact that they are familiar with the use of technology and have positive experience. It suggests that user's knowledge and training about self-driving vehicle system can overcome their concerns and their reliability could also be enhanced once they start using the technology.

References

1. Richard, W., Silberg, G.: Self-driving cars: the next revolution—KPMG and CAR, p. 2. (Center for Automotive research). http://www.kpmg.com/Ca/en/IssuesAndInsights/ArticlesPublications/Documents/selfdrivingcars-next-revolution.pdf (2012) (Consulted on 7 April 2014)
2. O'Toole, R.: Gridlock: why we're stuck in traffic and what to do about it. Cato Institute (2009)
3. Ibañez-Guzmán, J., Jian, X., Malcolm, A., Gong, Z., Chan, C.W. Tay, A.: Autonomous armoured logistics carrier for natural environments. In: Proceedings 2004 IEEE/RSJ International Conference on IEEE, Intelligent Robots and Systems (IROS), pp. 473

4. Le Vine, S., Zolfaghari, A., Polak, J.: Autonomous cars: The tension between occupant experience and intersection capacity. Transp. Res. C Emerg. Technol. **52**, 1–14 (2015)
5. Shladover, S., Ioannou, P.: Reasons for operating AHS vehicles in platoons. Autom. Highw. Syst, pp. 11–27 (1997)
6. Forrest, A., Konca, M.: Autonomous Cars and Society. Department of Social Science and Policy Studies, Worcester Polytechnic Institute, Worcester (2007)
7. Folsom, T.C.: Social ramifications of autonomous urban land vehicles. IEEE International Symposium on Technology and Society, Chicago (2011)
8. Cowen, T.: Can i see your license, registration and C.P.U.? Available at: http://www.nytimes.com/2011/05/29/business/economy/29view.html 28 May 2011. The New York Time [online]
9. Wu, C., Zhao, G., Ou, B.: A fuel economy optimization system with applications in vehicles with human drivers and autonomous vehicles. Transp. Res. D Transport Environ. **16**(7), 515–524 (2011)
10. Stock, K.: Autonomous driving. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.128.3935&rep=rep1&type=pdf. Accessed 3 March 2016
11. McCarthy, N.: Autonomous systems: social, legal and ethical issues. Transportation Research Board (2009)
12. Schoettle, B., Sivak, M.: A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia (2014)

A New Method and Results for Analyzing Decision-Making Processes in Automated Driving on Highways

Eugen Altendorf, Constanze Schreck and Frank Flemisch

Abstract While automated driving and advanced drivers' assistant systems (ADAS) become increasingly widespread, the human machine interaction for these technologies gains in importance. In today's traffic, some vehicles are capable of driving partially, conditionally or highly automated, at least in certain traffic situations, such as driving on developed highways. Nevertheless, these technologically advanced systems are not the only participants in traffic. With the interplay of more or less technologically advanced vehicles and humans on bikes and on foot, complex situations can arise that exceed the capabilities of an automated system and requires human cognition as a part of the solution. Although ADAS and automation solutions take this into account and try to compensate for the resulting effects, encounters with ambiguous situations can emerge. Furthermore, automation systems heavily rely on sensors and are therefore vulnerable to ambient conditions and situations that might limit the performance of the used sensor technology. For this reason, the (human) driver is still required for supervising the situation and often also as a fallback level in the case the technical system reaches or exceeds its performance restrictions. Guiding a vehicle, such as a car with partial or conditional automation, entails a different kind of driver vehicle interaction and cooperation between driver and automation as the one that is needed in the case of manual driving. For analyzing the decision making process of a human-machine-system with such an advanced automation during a typical driving situation like an take-over situation on a highway, a study addressing partially and conditionally/highly automated driving was conducted. The experiment with 30 participants consisted of three rounds with varying conditions in the driving simulator. During and after each round, participants were asked to answer several questions. For this purpose, a questionnaire has been developed to measure the relevant dimensions of the investigated driving situation. These were perceived utility, perceived time consumption, perceived safety, user satisfaction, perceived usability, and perceived

E. Altendorf (✉) · C. Schreck
RWTH Aachen University, Templergraben 55, 52062 Aachen, Germany
e-mail: e.altendorf@iaw.rwth-aachen.de

F. Flemisch
Fraunhofer FKIE, Fraunhoferstraße 20, 53343 Wachtberg, Germany

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_48

dominance (control over the vehicle guidance). The evaluation of the driving experiment shows that the level of automation as well as the volume of traffic have a significant effect on the decision-making behavior and on the individual perception when driving on a highway. This means that during automated driving, humans perceive and judge the driving situation differently. As a consequence, they tend to use the remaining decision authority for other purposes than when driving manually.

Keywords Automated driving · Advanced driver assistance systems (ADAS) · Human cognition · Decision-making · Human-machine interaction

1 Automated Driving and Cooperative Vehicle Guidance and Control

Over the last two decades, developments in advanced driver assistance systems (ADAS) and automated driving make the idea of a partially self-driving vehicle not only a research topic, but a reality on public roads. Today, several vehicle manufacturers offer cars on the market that are capable of driving automated on certain types of roads such as highways. Besides of the technical challenges of building and improving state of the art vehicle automation systems, another crucial challenge in this field becomes increasingly relevant. Since both, a human driver and a technical system, are steering the same vehicle, a multitude of complex technical functions have to be integrated in such a way that the human driver intuitively understands them. Considering that the level of automation is switched depending on the circumstances such as the state of the technical system or the environment, maintaining the driver's mode awareness is utmost important for keeping the driving situation safe. Furthermore, automation systems heavily rely on sensors and are therefore vulnerable to ambient conditions and situations that might limit the performance of the used sensor technology, e.g. the weather or unexpected objects. As a result, in many areas and situations the human driver is still needed. By designing automation systems that address such questions, the barriers between assistance systems and sophisticated automation solutions become blurred. Accordingly, complementing degrees of assistance and automation have been defined (e.g. [1–3]). For this reason, both cognitive compatibility and trust become paramount as they describe the way a human can be involved within the loop of the automation and with the automation or assistance itself (cf. [4, 5]). The development of systems that are cognitively compatible and trustworthy from a user's point of view can be supported by considering the user early in the development process [6]. These are the issues addressed by the concept of cooperative guidance and control, which is generic in describing the general degrees of freedom in the cooperation between the human and the automation as different levels of vehicle guidance and control. The H-Mode can be seen as a specific implementation of this concept. The H-Mode is

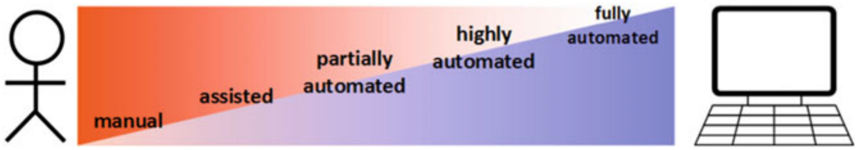


Fig. 1 Assistance and automation scale (based on [7])

based on the H-Metaphor, a design metaphor similar to the desktop metaphor for PCs [7]. It holistically describes the system behavior and the interaction and cooperation between partially as well as highly automated vehicles and the human driver. The H-metaphor draws its inspiration from the biological archetype of rider-horse or driver-carriage, where the driver can take the horse on the tight or the loose rein. A more extensive discussion of cooperative vehicle guidance and control can be found in [8, 9] describes the H-Mode in more detail (Fig. 1).

2 Guidance, Control, Responsibility and Decisions

Donges [10] introduces a three layer model of how human drivers conduct vehicles. Combinded with general models of human performance [11] and on the interaction with automated technical systems [4, 10] apply models in the context of vehicle guidance and control. Donges [10] introduces a three-layered model for the sub-tasks that occur during driving. According to this model, there are the levels of stabilization, guidance, and navigation. Technically speaking, the second level, guidance, can integrated as anticipatory control in the system [12]. An enhancement of the Donges-model can be found in Loeper et al. (2008), who add two sublayers within the level of guidance, i.e. the maneuver and the trajectory level. Thereby, they achieve a four-layered model of vehicle guidance and control, consisting of state control, trajectory planning, maneuver planning, and navigation. In this context, maneuvers can be understood as spatially and temporally related processes. In order to create a system with inner compatibility, the automation needs to follow the cognitive design laid out by this four-layered model for human drivers and needs to be constructed of similar levels.

In this paper, we mainly deal with the partially/highly-automated modes of cooperative guidance and control. In this driving mode, the vehicle takes over the tasks of stabilization and trajectory planning, but the human driver still decides about which maneuvers to be executed. The division of tasks is thus sequential along the levels that occur during driving with the automation taking care of the lower levels of trajectory planning and control, and the human focusing on the higher levels of maneuver planning and navigation.

Since vehicle guidance and control is executed on several different levels by different entities, complex decision making processes emerge. A very common situation would have the technical system taking over all aspects of vehicle guidance on the control level, while human driver and technical system plan and act together on the guidance level. Hereby, the human driver could decide about the maneuvers, such as a lane change, whereas the sensors of the vehicle analyse the environment and make suggestions to the human driver based on their scans. In this or any comparable situation, the decision making process is influenced by the evaluation of the technical system as well as by the human judgement. For supporting efficient and target-oriented decision making processes, the evaluations of the technical system should be communicated in a cognitively compatible way. Along the same lines, the basic characteristic, which is used for assessing driving situations, should be compatible between both partners, the human and the machine. In order to achieve this, a common and comparable goal and value system is very helpful, in which the technical system should map the human evaluations. To be able to systematically analyze these attitudes of human drivers in relevant driving situations, we have developed a questionnaire, which addresses the dimensions utility, time consumption, safety, user satisfaction, usability, and dominance. In addition to this, we use of one-dimensional scales for rating driving situations with respect to the dimensions use, safety, and time without interrupting the driving experience.

3 Measuring Acceptance and Intention

In 2015, a study using the driving simulator of IAW RWTH Aachen was conducted (n = 35). The average age has been estimated at nearly 29 years (SD = 8752). The simulation included a truck as a vehicle steered by an active sidestick. The study contained three phases. The first treatment was an almost manual ride with minor assistance systems such as emergency braking (Manual Mode with an additional light assistance, which prevented to drive on the verge of the motorway). The second and third treatment consisted in each way of a ride on the same track using the more extensive Loose-Rein (partially automated) automation level. The traffic situation varied between treatments two and three in a way, that each participant took a ride in Low Traffic as well as High Traffic in Loose Rein Mode.

Each phase was followed by an online questionnaire, where, among others e.g. socio-demographics, six different scales consisting of several items have been investigated on a seven-point Likert-scale. Hereby, 1 is the lowest rating, whereas 7 is the highest approval rating [13]. Furthermore, during the Loose Rein treatments the participants were asked questions about driving situation they have just experienced, more precisely overtaking manoeuvres. They were to evaluate their experiences on the dimensions of total satisfaction, perceived timing of the overtaking process as well as perceived safety. The evaluation was carried out on basis of a scale from 0 to 100 % whereby in any case 100 % means the best evaluation and 0 % the worst. All in all the aim of these assessments is to measure effects on acceptance in decision making processes what can be derived from correlations in between the different scales raised. In this chapter the starting point and final design of the online questionnaire and the interview are presented as well as first results concerning reliability of the different scales.

3.1 The Online Questionnaire: Starting Point

The central objective of the present study was to measure acceptance and intention of the driver towards automated driving and cooperative vehicle guidance and control in decision making processes. For this purpose, extensive research on previous studies concerning ADAS and automated driving was performed. Furthermore, studies regarding consumer products in a broader sense (e.g. software, hardware, cell phones, websites) with focus on usability and acceptability aspects as well as perceived ease of use were included. Out of this pool of different investigations similarities relating to the measured dimensions were extracted. The outcome were six scales consisting of different numbers of items depending on the findings of the research on preceding studies. The investigated scales are

- Perceived Utility
- Perceived Time Consumption
- Perceived Safety
- User Satisfaction
- Perceived Usability and
- Perceived Dominance (Control over the Vehicle Guidance).

The origin of each item is presented in the chapter below.

3.2 The Online Questionnaire in Detail

The questionnaire being used covers six scales consisting of 41 items which were all requested by a consistent 7-point Likert scale (“I strongly disagree” to “I strongly agree”). Some of the 9 original questionnaires were (loosely) translated from English to German. If the items were originally formulated as semantic differentials, they were modified into questions that can be answered by a Likert-Scale.

Furthermore, at the beginning of each experiment, sociodemographic data were collected (age, gender, if the participant is holding a driver license and for how long, if the participant owns a car, how often the car is used, the annual mileage and the usage of driver assistance systems) (Table 1).

Perceived Utility. The measurement of perceived utility includes 13 items, adopted from several previous studies.

Table 1 Average of Cronbach Alphas of the 6 scales measured

Scale	Number of items	Cronbach α (average)
Perceived utility	13	0.925
Perceived time consumption	4	0.722
Perceived safety	7	0.835
Perceived user satisfaction	3	0.852
Perceived usability	7	0.897
Perceived dominance	6	0.743

The questionnaire by Van der Laan et al. [14] was originally designed for the measurement of acceptance of new technologies, such as in-vehicle systems or advanced telematics. Van der Laan et al. have observed a high degree of correlation of the scale (Cronbach's α between 0.73 and 0.94). The items were tested in six different experiments (four in a driving simulator, two on the street), and they were found to be highly reliable. For our questionnaire, we use four items of the questionnaire by Van der Laan et al. [14]. The following items were used (the italic words in brackets show the actually chosen judgement):

- "My judgements of the system (…) are pleasant/unpleasant" (*pleasant*);
- "My judgements of the system (…) are useful/useless" (*useless*);
- "My judgements of the system (…) are assisting/worthless" (*I think there is a real added value with respect to conventional vehicles*);
- "My judgements of the system (…) are effective/superfluous" (*effective*).

The Technology Acceptance Model (TAM) by Davis [15] is a well-known model for user acceptance, similar to the Unified Theory of Acceptance and Use of Technology (UTAUT) by Venkatesh and Bala [16], Osswald et al. [17] have refined these two models to be better applicable to the automotive sector by introducing the Car Technology Acceptance Model (CTAM). Thereby the following 8 factors were created and measured on a 7-point Likert Scale (0 = not at all applicable, 6 = totally applicable): Performance expectancy, Effort expectancy, Social influence, Facilitating conditions, Self-Efficacy, Anxiety, Perceived Safety and Attitudes Towards Using Technology. The items Performance expectancy and Perceived safety were assumed for the present study. The two items taken from [17] were lightly modified in a way that they refer to the vehicle as a whole and not only to the assistant system. The final items are

- "The vehicle increases my driving performance" and
- "Driving this vehicle is very comfortable."

The Technology Acceptance Model (TAM) by Davis [15] was also used by Lund [18] to measure the Usability of consumer products. 6 of 30 items were used in the present study for the scales of perceived usability and perceived time consumption. Thereby five items were adopted from [18]. We specified the question to refer to the vehicle in the situation the participant has just experienced, so we replaced “it” by “the vehicle”. The following items were used in the present investigation:

- “The vehicle does everything I would expect it to do”;
- “The vehicle meets my needs”;
- “The vehicle makes the things I want to accomplish easier to get done”;
- “The vehicle is useful”;
- “The vehicle helps me be more productive”.

Two items, which were developed by Trommer and Höltl [19], were used to measure perceived time consumption, which is in the coming paragraph, and perceived utility. The original and standardized questionnaire (n = 5807) was generated for the project eCoMove, which was a study in 11 European countries concerning a navigation wizard (assistant) to reduce CO_2 emission. The following item was taken from [19] and also changed in the sense, that it refers to the vehicle as a whole.:

- “The vehicle restricts my freedom while driving”

The last item of the perceived utility scale was loosely based on an item by Ussat [20]. For measuring the acceptance of an assistant system or rather of a ubiquitous system, [20] has utilized a questionnaire by Rothensee [21]. The UbiTAM allows to collect data concerning the impact on driving performance and the quality of the given task, in this case a driver assistance systems for the personalized choice of destination. One item was adopted for the scale of perceived utility in the present investigation which was originally developed by Davis [15]:

- “I am satisfied with the vehicle guidance”.

The total reliability of the scale of perceived utility is excellent (average: $\alpha = 0.925$).

Perceived Time Consumption. The Perceived Time Consumption scale measures the subjective evaluation of timescales in the situation the participant has just experienced. For this purpose, four items were used. The first is adopted from Trommer and Höltl [19] and rotated so that “The system saves travel time” was modified into:

- “Driving with this vehicle extends travel time”.

The second item, originally by Osswald et al. [17], was adapted to the background of investigation. According to this “Using the system enables me to accomplish my goals more quickly”, was changed into

- “The vehicle enables me to reach my final destination more quickly”.

Third, an item by Lund [18] (originally: “It saves me time when I use it”) was adapted:

- “Driving with this vehicle saves time”.

Finally we formulated an item on our own to measure implicitly the actual perceived time consumption in the situation the driver has just experienced:

- "The vehicle helps me to get through the traffic in time".

The total reliability of this scale is acceptable (average: $\alpha = 0.722$).

Perceived Safety. The third scale "perceived safety" was measured by eight items, whereby six of them were adopted by Osswald et al. [17]. Five of them were originally created to measure "perceived safety". The items were:

- "I believe that using the system is dangerous";
- "Using the system requires increased attention";
- "The system distracts me from driving";
- "I feel safe while using the system";
- "Using the system decreases the accident risk".

The last one was originally used to measure "performance expectancy" but covers also security aspects:

- "If I would use the system I will reach my destination safely".

In a pilot project, [22] have developed an online questionnaire to identify diffusion, publicity, utilization, and perception concerning driver assistant systems as well as experiences and expectations of elderly drivers with driver assistant systems. Among other things, the subjective feeling of safety and the influence of the assistant systems were determined. One item regarding perceived safety was adopted. They have compared the perceived safety of different assistance systems with their questionnaire. For each driver assistance system a subjective assessment of the driving behavior was requested. In its original version, the question was "Do you think that the driver assistance system has an influence on how carefully you are driving? Please give an evaluation for each system." The response options were: "I rather drive more carefully/I drive as carefully as always/I drive rather less carefully/I don't know". A modification of this item was made:

- "I rather drove more carefully with this vehicle."

In a simulator study by Comte [23], two systems of speed check (Fixed system and Dynamic System) were compared to an "advisory system" and a "baseline control" (which means "no system"). Besides the semantic differentials by Van der Laan et al. [14] and the NASA-RTLX by Hart and Staveland [24], which were used in Comte's investigation, an Acceptability Questionnaire was developed to measure the attitude of the participants concerning the assistant system. Finally one item was used:

- "In your opinion, would this vehicle make people commit less offences?"

The total reliability of the scale of perceived safety is good (average: $\alpha = 0.835$), when leaving out the item "I rather drove more carefully with this vehicle" (otherwise $\alpha = 0.768$).

User Satisfaction. To measure the scale of user satisfaction, again three semantic differentials by Van der Laan et al. [14] were reformulated into questions which can be answered on Likert scales:

- "My judgements of the system (…) are undesirable/desirable", was turned into "Driving with this vehicle is desirable".
- "My judgements of the system (…) are nice/annoying", was changed into "Driving with this vehicle is nice".
- "My judgements of the system (…) irritating/likeable" was modified into "Driving with this vehicle is irritating."

The total reliability of the scale of User Satisfaction is good (average: $\alpha = 0.852$).

Perceived Usability. To evaluate the scale of perceived usability seven items were taken from Brooke [25] who developed the System Usability Scale (SUS) to gain a comprehensive view on the subjective evaluation of usability in the context of software systems. Brooke has chosen the final 10 items out of a pool of initial 50 items, which showed a high degree of correlation. 7 of the final 10 items were utilized in the present investigation. ("The system" was always replaced by "the vehicle".) The items being used are:

- "I found the vehicle very cumbersome to use";
- "I would imagine that most people would learn to drive with this vehicle very quickly";
- "I thought there was too much inconsistency in this vehicle";
- "I found the various functions in this vehicle were well integrated";
- "I thought the vehicle was easy to use";
- "I found the vehicle unnecessarily complex";
- "I think I would like to use this vehicle frequently".

The total reliability of the scale of Perceived Usability is good (average: $\alpha = 0.897$).

Perceived Dominance. For the last scale "Perceived dominance" six semantic differentials by Mehrabian and Russell [26] were transformed into statements that can be handled with Likert scales. Mehrabian and Russell [26] have created six semantic differentials which are introduced as follows: "Please rate the feelings in the previous situation with the following adjective pairs. Some of the pairs might seem unusual, but you will probably feel more one way than the other. So, for each pair, put a check mark close to the adjective which you believe to describe your feelings better. The more appropriate that adjective seems, the closer you put your check mark to it." The original differentials were: "controlling/controlled; influential/influenced; important/awed; in control/cared-for; dominant/submissive; autonomous/guided". On this basis, the following items were constructed:

- "The vehicle controls me a lot";
- "The vehicle has influenced me a lot";

- “I have dominated while driving with this vehicle”;
- “I felt influential while driving with this vehicle”;
- “I had control while driving with this vehicle”;
- “The vehicle guided me while driving”.

The total reliability of the scale of perceived safety is acceptable (average: $\alpha = 0.743$).

3.3 The Interview During Ride

The background of this interview is the RSME developed by Zijlstra and van Doorn [27] even though in an extensively modified way. The rating does not concern mental effort, which is the aim of RSME, but subjective evaluations of the situation just experienced. Also the participants were not to put a mark on the scale on their own, but hat to specify a value which was noted by the investigator, because the ride should not be interrupted. The justification of the engaged interview concentrates on the following point:

“The importance of subjective experiences goes beyond the issue of subjective ratings. The phenomenological experiences of human operators affect subsequent behavior and thus also their performance and physiological responses in a given situation. If operators consider the workload of a task is excessive, they may behave as though they are overloaded, even though in all objectivity the task demand is low.” [28].

Transferred to the situation of the study presented in this paper: Even if it looks like an overtaking maneuver went very well (e.g. no crash, no delay of traffic), the assessment can be different for the driver. Safety and satisfaction are highly subjective estimations so one cannot say that 100 % of satisfaction means the same for all participants, but it is possible to compare the estimations of one single participant, also with her evaluation by the online questionnaire to make reliable statements.

The participants were informed to be interviewed about their subjective evaluation of various overtaking maneuvers. They were asked three questions directly after initiating to overtake another vehicle and had to answer them on a scale from 0 to 100 % (all conceivable steps between 0 and 100 % were possible). Besides, the meaning of these questions were explained to each participant. The questions were:

- “All in all, how satisfied are you with this driving situation?”;
- “How satisfied are you with the timing of this driving situation?” (That means: “Did it go fast enough? By your feeling, were you thwarted in an uncomfortable way?”) and at last
- “How safe did you feel within this driving situation?”.

A value of 100 %, in any case, corresponds to the best evaluation (100 % of total satisfaction, 100 % of satisfaction with the timing, and 100 % of safety.). Each participant had to complete 20 overtaking maneuvers, but they were not informed about this amount.

4 Conclusion and Outlook

In this paper, we have presented how a questionnaire for experiments in the field of automated vehicles can be structured. We have identified six scales that can help the investigation of the driving experience and the analysis of the underlying decision making processes when guiding an automated vehicle through highway traffic. The scales were perceived utility, perceived time consumption, perceived safety, user satisfaction, perceived usability, and perceived dominance. Using the experimental data, we could show that the reliability of the scales, as measured by the Cronbach's alpha is acceptable to outstanding in all cases. By complementing the resulting comprehensive online questionnaire with repeated ad hoc questions, a more detailed picture of the user experience could be obtained. This can be helpful when looking into possible user acceptance of the systems, i.e. vehicles, studied.

The questionnaire is meant to serve as a basis for further experiments in the field. Future research can refine the proposed scales and add new scales, such as one concerned with ecological issues, monetary values, or the added-value of leisure time during the ride.

Acknowledgments The research conducted was partly funded by the Deutsche Forschungsgemeinschaft (DFG) within the project "Arbitrierung von kooperativer Bewegung von hochautomatisierten Mensch-Maschine-Systemen" and partly by RWTH Aachen University.

References

1. Gasser, T.M., Arzt, C., Ayoubi, M., Bartels, A., B€urkle, L., Eier, J., Flemisch, F., Häcker, D., Hesse, T., Huber, W., Lotz, C., Maurer, M., Ruth-Schumacher, S., Schwarz, J., Vogt, W.: Rechtsfolgen zunehmender Fahrzeugautomatisierung (Legal consequences of increasing vehicle automation) – Gemeinsamer Schlussbericht der Projektgruppe Bundesanstalt für Straßenwesen (BASt) (2012)
2. NHTSA: preliminary statement of policy concerning automated vehicles. National Highway Traffic Safety Administration (2013)
3. SAE J 3016: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Society of Automotive Engineers (2014)
4. Löper, C., Kelsch, J., Flemisch, F.O.: Kooperative, manöverbasierte Automation und Arbitrierung als Bausteine fuer hochautomatisiertes Fahren (Cooperative, maneuver based automation and arbitration spices for highly automated driving). In: Gesamtzentrum für Verkehr Braunschweig (ed) Automatisierungs-, Assistenzsysteme und eingebettete Systeme für Transportmittel. GZVB, Braunschweig, pp. 215–237 (2008)

5. Vanholme, B.: Highly automated driving on highways based on legal safety. (Doctoral thesis), University of Evry-Val-d'Essonne (2012)
6. Flemisch, F.O., Kelsch, J., Löper, C., Schieben, A., Schindler, J.: Automation spectrum, inner/outer compatibility and other potentially useful human factors concepts for assistance and automation. In: de Waard, D., Flemisch, F.O., Lorenz, B., Oberheid, H., Brookhuis, K.A. (eds.) Human Factors for Assistance and Automation, pp. 1–16. Maastricht, Shaker (2008)
7. Flemisch, F.O., Adams, C.A., Conway, S.R., Goodrich, K.H., Palmer, M.T., Schutte, M.C.: The H-metaphor as a guideline for vehicle automation and interaction. NASA Langley, Report No. NASA/TM-2003-212672 (2003)
8. Flemisch, F.O., Bengler, K., Bubb, H., Winner, H., Bruder, R.: Towards cooperative guidance and control of highly automated vehicles: HMode and conduct-by-wire. Ergonomics **57**(3), 343–360 (2014)
9. Altendorf, E., Baltzer, M., Heesen, M., Kienle, M., Weißgerber, T., Flemisch, F.: H-Mode: a haptic multimodal interaction concept for cooperative guidance and control of partially and highly automated vehicles. In: Winner, H., et al. (eds.) Handbook of Driver Assistance Systems. Springer International Publishing, Switzerland (2015)
10. Donges, E.: Aspekte der Aktiven Sicherheit bei der Führung von Personenkraftwagen (Aspects of active safety when driving a passenger vehicle). Automob. Ind. **27**, 183–190 (1982)
11. Rasmussen, J.: Skills, rules and knowledge; signals, signs, and symbols, and other distinctions inhuman performance models. IEEE Trans. Syst. Man Cybern. **3**, 257–266 (1983)
12. Donges, E.: Driver behavior models. In: Handbook of Driver Assistance Systems: Basic Information, Components and Systems for Active Safety and Comfort, pp. 19–33 (2016)
13. Altendorf, E., Borowski, D., Ramaekers, J., Flemisch, F.: Assistenz versus Kontrolle beim hochautomatisierten Fahren – eine Akzeptanzanalyse. 62. GfA-Frühjahrskongress, 2.03.-4.03.2016, Aachen (2016)
14. 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 Emerg. Technol. **5**, 1–11 (1997)
15. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, pp. 319–340 (1989)
16. Venkatesh, V., Bala, H.: Technology acceptance model 3 and a research agenda on interventions. Decis. Sci. **39**(2), 273–315 (2008)
17. 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, pp. 51–58 (2012)
18. Lund, A.M.: Measuring usability with the USE questionnaire. Usability Interface **8**(2) (2001)
19. Trommer, S., Höltl, A.: Perceived usefulness of eco-driving assistance systems in Europe. IET Intel. Transport Syst. **6**(2), 145–152 (2012)
20. Ussat, C.: Personalisierte Optionsauswahl im Fahrzeuginformationssystem. (Personalized option selection in vehicle information systems) (Doctoral thesis). Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II (2012)
21. Rothensee, M.: Psychological Determinants of the Acceptance of Future Ubiquitous Computing Applications. Kovač (2010)
22. Schmitt, K.-U., Jordan, B.: (AGU Zürich): Pilotprojekt: Ältere Fahrzeug-lenkerInnen und Fahrassistenzsysteme (Pilot project: elderly drivers and driver assistant systems) (2013)
23. Comte, S.L.: Evaluation of in-car speed limiters: simulator study. Managing Speeds of Traffic on European Roads: Working Paper, vol. 3, p. 1 (1998)
24. 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)

25. Brooke, J.: SUS—a quick and dirty usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)
26. Mehrabian, A., Russell, J.A.: An Approach to Environmental Psychology. MIT Press, Cambridge (1974)
27. Zijlstra, F.R.H., van Doorn, L.: The Construction of a Subjective Effort Scale, Report. Delft University of Technology, Department Social Sciences & Philosophy (1985)
28. Zijlstra, F.R.H.: Efficiency in work behaviour: a design approach for modern tools. Delft University of Technology, TU Delft (1993)

A Human-Machine Interface for Cooperative Highly Automated Driving

Frederik Naujoks, Yannick Forster, Katharina Wiedemann and Alexandra Neukum

Abstract Cooperative perception of the traffic environment will enable Highly Automated Driving (HAD) functions to provide timelier and more complex Take-Over Requests (TOR) than it is possible with vehicle-localized perception alone. Furthermore, cooperative perception will extend automated vehicles' capability of performing tactic and strategic maneuvers independently of any driver intervention (e.g., avoiding of obstacles). In this paper, resulting challenges to the design of the Human-Machine Interface (HMI) are discussed and a prototypical HMI is presented. The prototype is evaluated by experts from the field of cognitive ergonomics in a small-scale simulator study.

Keywords Automated driving · Human-systems integration · Human-machine interface

1 Introduction

Highly automated driving (HAD) functions are expected to be ready for marked introduction in the near future. By fusing vehicle-localized environment perception with information provided by other road users or infrastructure, so-called *cooperative*

F. Naujoks (✉) · A. Neukum
Center for Traffic Sciences, University of Würzburg, Röntgenring 11, 97070 Würzburg, Germany
e-mail: naujoks@psychologie.uni-wuerzburg.de

A. Neukum
e-mail: neukum@psychologie.uni-wuerzburg.de

Y. Forster · K. Wiedemann
Würburg Institute for Traffic Sciences, Robert-Bosch-Straße 4, 97209 Veitshöchheim, Germany
e-mail: forster@wivw.de

K. Wiedemann
e-mail: wiedemann@wivw.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_49

perception [1, 2], the capabilities of these systems can be greatly enhanced. For example, it may be possible to provide drivers with advance information about upcoming system limits to improve driving performance during take-over situations [3]. At the same time, *cooperative perception* will enable highly automated vehicles to handle some driving situations independently of any driver intervention by reacting *strategically* (e.g., lane change due to upcoming lane merge) or *tactically* (e.g., avoiding of obstacles on the road) to the situation. Thus, the HAD function eventually becomes an autonomous agent as it will perform driving maneuvers such as lane changes (e.g., because of merging lanes) or adaptations of the host vehicle's speed (e.g., because of speed limits) independently. Humans tend to attribute rationality and intentionality to such autonomous agents which may become "team members" [4]. Interacting with autonomous agents places new demands on the human operator such as understanding of the system's current behavior as well as its intentions [5, 6]. In view of these challenges, the development and evaluation of a HMI for cooperative HAD will be presented in this paper. The visual component of the HMI, that could be displayed on an in-vehicle display, is the focus of the paper. The HMI design is targeted on the basis of a shared control framework between driver and automation. This should put into effect automation benefits such as fast and accurate response in take-over situations and prevent problems such as mode errors or out-of-the-loop unfamiliarity at the same time [7]. The HMI was implemented in a driving simulator. Evaluation of the HMI (e.g., usability, system understanding, etc.) by a sample of $N = 6$ experts in the field of cognitive ergonomics, as well as future directions of the HMI design will be presented.

2 HMI Considerations for Cooperative Highly Automated Driving

Highly Automated Driving is characterized by a system that takes over both longitudinal and lateral control of the vehicle. In contrast to Partial Automation that is already commercially available by several automobile manufacturers, there is no longer a need to continuously monitor the system's status [8, 9]. The drivers' only responsibility is to be available in case the system's functional limits are exceeded (e.g., sensor degradations, unclear lane markings in work zones, missing lane markings, etc.).

2.1 Enhancing Take-Over Requests Through Cooperative Perception

Two different generations of automated vehicle functions may be distinguished on the basis of the sensor framework used to monitor the traffic environment. The "first generation" of automated driving functions is based on on-board sensor technology such as camera or radar. If safe vehicle guidance can no longer be ensured, a Take-Over Request (TOR) is presented immediately with the goal of getting the driver to take over manual vehicle control as fast as possible [10–13]. On the other hand, HAD of the "second generation" [14] will not only rely on information provided by on-board sensors, but will additionally rely on *cooperative perception* technology. For example, the exact location of a construction site that the automation cannot handle by itself could be calculated from information provided by other road users that have already passed it. TORs can thus be presented well in advance, providing the driver with more time to get back into the loop while the system is still active. This lengthened time window may be extremely beneficial to vehicle safety and comfort. However, there is little research on how to design the Human-Machine Interface so that drivers can ideally benefit from data provided by *cooperative perception*. The proposed HMI was developed on the basis of available studies that provided information about the HMI for automated vehicle functions [3, 13, 15, 16] as well as studies on the design of driver warning and information systems based on *cooperative perception* technology [17, 18].

Figure 1 shows the proposed HMI. During normal operation (i.e., vehicle guidance is executed by the automated driving function), the following information is provided to the driver:

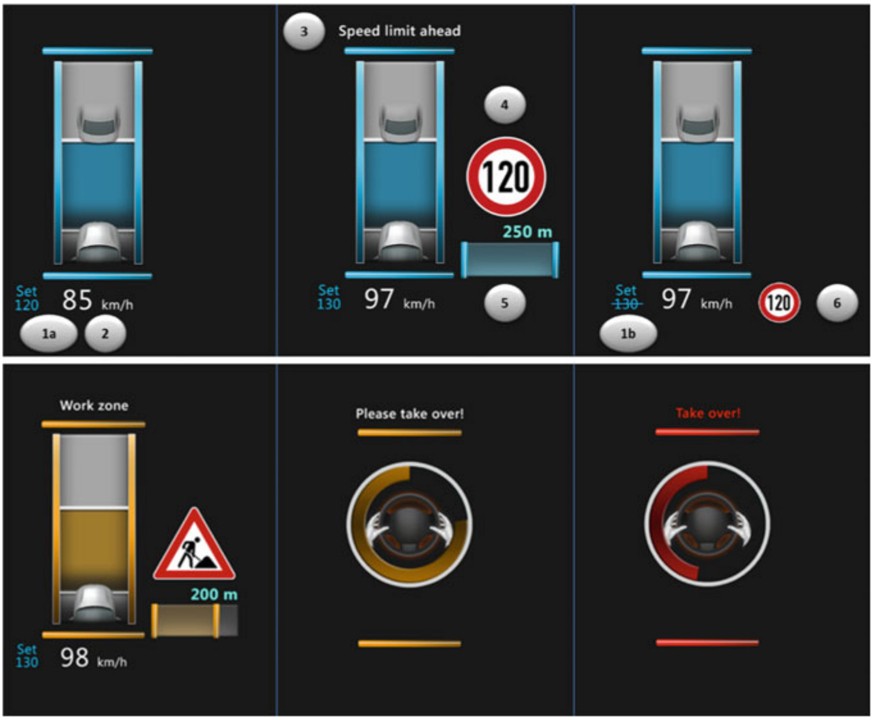


Fig. 1 Prototypical HMI for driving situations during normal operating state (*upper part*) and take-over situations (*lower part*)

- *Main state*: In the center of the HMI (Fig. 1, top left), it is displayed that lateral guidance is carried out by the HAD function by changing the color of the lane symbols from grey to blue. A blue rectangle in front of the host vehicle shows that longitudinal vehicle control is also carried out by the HAD function. The length of the rectangle shows the set distance to vehicles ahead. It is also displayed whether a lead vehicle is recognized by depicting the lead vehicle (or not). In sum, this part of the proposed HMI resembles that of existing HMI solutions for ACC with additional steering assistance [13].
- *Velocity*: Set speed (1a) and current speed (2) are displayed below the main state (Fig. 1, top left). Both driver and system may change the set speed. If the driver changes it, the new set speed is shown. If the system lowers it (e.g., because of a speed limit), this is depicted by crossing out the set speed (1b) until the speed limitation is cancelled.
- *Traffic events*: If a traffic event, such as an upcoming speed limit or road curvature, requires speed adaptation, this is displayed to the driver in advance by a message box on top of the HMI (3), a symbolic representation of the traffic event (4) and the distance to the traffic event (5, Fig. 1, top middle). If the adaptation will be carried out by the automated driving function, this is displayed by all HMI elements remaining in blue color. Subsequently, the resulting speed limitation is displayed until the situation has passed (1b and 6, Fig. 1, top right).

If, however, the upcoming traffic event cannot be handled by the automated driving function, a stepwise *take-over process* is initiated. First, the traffic situation is announced to the driver (Fig. 1, bottom left). By implementing situation announcements prior to the TOR, driver performance during take-over situations can be improved [3]. The situation announcement consists of several HMI elements. The type of traffic event is shown via a text box as well as a symbolic representation of the traffic event as research on cooperative warning systems has demonstrated the usefulness of specific warnings [17, 19]. The remaining distance until the traffic event is reached is also displayed on the basis of prior research on cooperative warning systems [18]. During the approach, a non-imminent indication to take over manual control is displayed (so-called Soft TOR [13], Fig. 1, bottom middle) when there is still a comfortable time budget to take over manual driving, followed by an imminent TOR (so-called Hard TOR [13], Fig. 1, bottom right) in case the driver does not react in time to the first TOR. The aim of this graded take-over procedure is to provide enough time and sufficient information about the upcoming system limit to the driver, so that she/he can (re)direct her/his attention to the driving task and get ready to drive manually [3].

2.2 Displaying Strategic and Tactic Maneuvers

Through a *cooperative perception* approach, it is possible that the automated driving system will be enabled to carry out certain driving maneuvers independently. The highly automated vehicle might thus be able to handle some driving situations independently of any driver intervention by reacting *strategically* (e.g., lane change because of upcoming lane merge) or *tactically* (e.g., avoiding of an obstacle on the road) to driving events. From a human factors perspective, this creates new challenges to the HMI design. Enhancing the capabilities of the automated vehicle may result in an information need as the driver might feel uncomfortable or even endangered when automated maneuvers are executed without communicating that no manual intervention is needed. This could result in decreased trust in the automated vehicle, as the *process* involved in the automated driving function cannot be completely understood [20, 21]. Distrust in the automated vehicle function could bring drivers to cancel the respective automated driving maneuvers.

The proposed HMI for displaying such maneuvers is shown in Fig. 2. When approaching the maneuver, the situation is announced to the driver (Fig. 2, left part). In order to unambiguously communicate that no manual intervention is needed, the same blue colors as in normal operating state are used in depicting the main state. Again, the type of traffic event [3, 17] and the remaining distance [18] are announced to the driver. Depending on the type of event, this announcement may be presented rather shortly before the situation is reached (in case of *tactical* maneuvers) or well in advance (in chase of *strategic* maneuvers). Before the maneuver is executed, the type of maneuver is shown, such as in-lane avoiding (upper part of Fig. 2, middle) or lane change (lower part of Fig. 2, middle), together

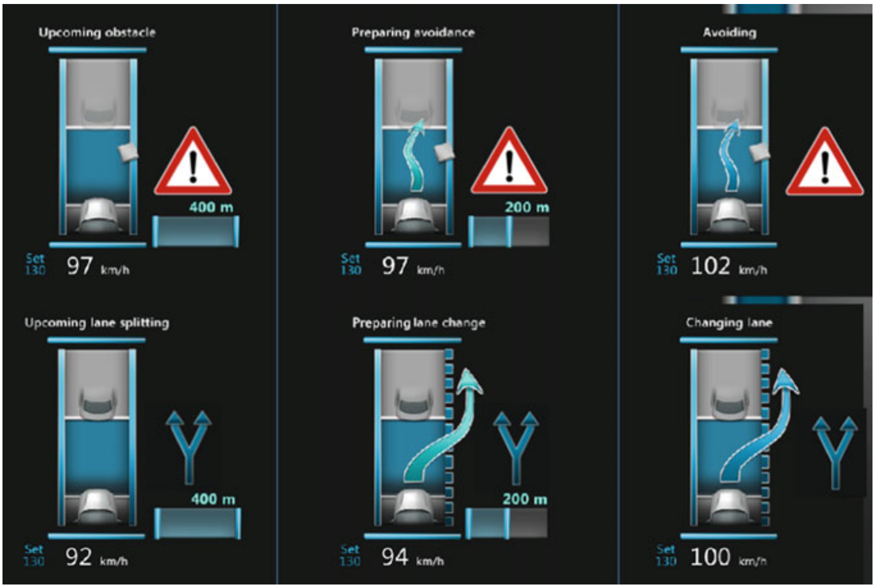


Fig. 2 Prototypical HMI for tactic (*upper part*) and strategic (*lower part*) driving maneuvers

with a notification via a message box that the automation is planning the maneuver. Subsequently, the execution of the maneuver is communicated via a text box (Fig. 2, right).

3 Pilot Study

3.1 Method

Study Setup. A low-fidelity driving simulator was used in the pilot study. The driving simulator consisted of three TFT-monitors displaying the environment and one TFT-monitor displaying the HMI. Participants were seated on a chair right in front of the desk with an attached gaming steering wheel and pedals underneath. The driving simulation software SILAB was used to simulate the traffic environment and the HAD function. The study setup did not include any audio output, as the primary aim of the study was to assess the visual HMI elements.

Experimental Design. Participants drove in highly automated mode on a three lane highway with moderate traffic density. The entire drive lasted approximately 17 min and covered 7 driving events that required more or less interaction with the HAD function:

- *Take-over situations*: In three cases a transition to manual driving was necessary. In the first situation, lane markings were completely missing on a road section (situation "missing lane markings"). In the second situation lane markings were not clearly visible (situation "unclear lane markings"). The third situation consisted of a construction site (situation "construction site").
- *Automated maneuvers*: In four cases, the HAD function was able to handle the traffic event independently of any driver intervention. These scenarios consisted of a lane change due to splitting lanes (situation "splitting lanes"), adaptation of the host vehicle's speed due to a speed limit (situation "speed limit"), adaptation of the host vehicle's speed due to high road curvature (situation "curvature") and avoiding an obstacle on the road (situation "avoiding").

Sample. The sample consisted of six experts from the field of cognitive ergonomics (3 male). All participants were researchers working at the Würzburg Institute for Traffic Sciences (WIVW).

Dependent Variables and Instructions. Participants were instructed to complete the simulator drive using the HAD function. They were instructed on how to activate and deactivate the system, but they were not given any advance information about the HMI. The following dependent measures were collected during the pilot study (see Table 1):

- *Usability*: Participants were asked to rate the comprehensibility of the HMI on a scale ranging from 0 ("not at all") to 15 ("fully comprehensible") [22] after each of the test situations. Having finished the test drive, they were also asked to

Table 1 Dependent measures

Evaluation criterion	Description of dependent measures	Unit/range
Usability	Comprehensibility of HMI (*How comprehensible was the displayed information?*)	Rating [0…15]
	System usability scale (SUS)	Rating [0…100]
	Improvement suggestions	Frequency and type of improvement suggestion
Understanding of HMI elements	Participants were asked to explain the most important HMI elements (*What is the meaning of the blue color coding? When is a lead vehicle displayed? What is the meaning of the distance bar? What is the meaning of the blue lane markings?*)	Frequency of correct answers
Interaction behavior	Take-over situation: Stage of take-over process in which drivers take over manual control (announcement, Soft TOR or Hard TOR) System maneuvers: Frequency of unnecessary deactivations	Frequency of observations

complete the System Usability Scale (SUS) [23] to get an estimate of the system's overall usability. During the drive, participants were told to verbally comment on the visual HMI elements and to communicate any misunderstandings or improvement suggestions to the experimenter. The experimenter noted down the comments.

- *Understanding of HMI elements*: After the drive, participants were also asked to explain the most important HMI elements (i.e., color coding, lead vehicle recognition, distance bar, state of lateral guidance) to the experimenter to assess system understanding. The HMI was shown on a handout and participants were asked to explain the HMI elements. Here, it was of interest whether participants were able to explain the HMI elements correctly.
- *Interaction behavior*: Interactions with the HAD function were observed by the experimenter during the drive. With regard to situations including TORs, the stage of the take-over process (announcement, Soft TOR or Hard TOR), in which drivers would take over manual driving, was of main interest. In case of system-initiated maneuvers, the frequency of unnecessary system deactivations was assessed.

3.2 Results

Usability. The HMI was evaluated positively on the standardized measures (i.e., comprehensibility rating and SUS). The drivers' comprehensibility ratings are shown in Fig. 3. From the figure, it is evident that most drivers rated the comprehensibility of the HMI to be "high" to "very high". SUS scores were calculated as described in [23]. These results revealed SUS-scores in the upper quartile of the

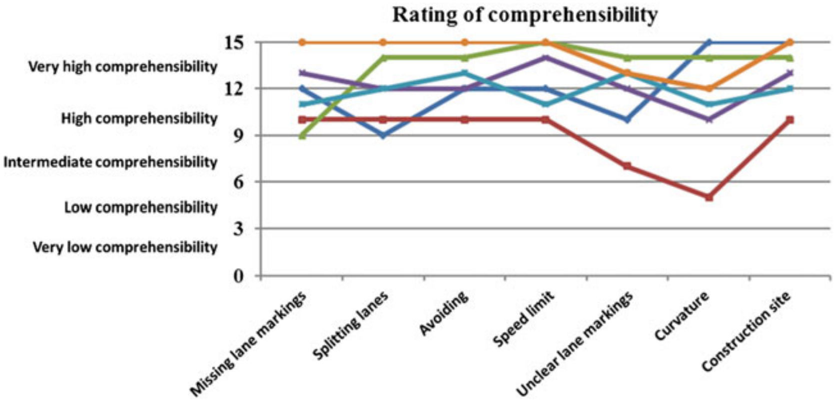


Fig. 3 Ratings of comprehensibility gathered after each situation (*Question* "How comprehensible was the displayed information?"). Every line represents one participant

reference distribution, indicating "good" to "excellent" usability of the HMI (M = 82.92, SD = 7.83). The analysis of the verbal comments gathered during the automated drive revealed the following main improvement suggestions:

- *Distance indication*: Three experts suggested to use a vertical distance bar instead of the horizontal one (see Figs. 1 and 2); one expert additionally suggested to avoid inconsistency of the vertical bar used for distance indications and the depiction of the remaining time to take over manual driving during take-over situations (which was presented by a more or less filled circle, see Fig. 1).
- *Avoiding clutter*: Two experts suggested deleting the horizontal bars pertaining to the HMI main state as they do not carry any additional information.
- *Creating redundancy*: One expert suggested to provide explicit information about the upcoming take-over situations (i.e., that taking over manual driving will be necessary) during the announcement stage in addition to sole color coding. One expert suggested using additional speech output, emphasizing the information provided in the message box.

Understanding of HMI Elements. Coding of responsibility through color, lead vehicle recognition and distance indication were understood by all participants. However, two drivers stated that the blue lane markings would indicate that lane markings are detected correctly by the HAD function, whereas they were supposed to also indicate that lateral control is automated.

Interaction Behavior. Out of 18 total take-over reactions, the drivers took over manual control in the announcement phase in two cases. In most cases, the drivers took over manual control during the Soft TOR (14 cases). Only in two cases, they waited until the Hard TOR was presented. Regarding automated driving maneuvers, only one unnecessary deactivation was recorded, which took place in the "avoiding" situation.

4 Summary and Future Directions

A prototypical HMI for cooperative Highly Automated Driving was proposed, acting on the assumption that *cooperative perception* of the driving environment will extend the capabilities of HAD functions and thereby creating new challenges to the design of the HMI. The proposed HMI was specifically designed to enhance driver-vehicle interactions in two ways. First, communicating advance information about upcoming system limits (such as missing lane markings or construction sites) to the driver might greatly enhance the safety of transitions to manual driving. However, there is little research on how to display such information to the driver. Second, *cooperative perception* will enable HAD functions to carry out *strategic* and *tactic* driving maneuvers such as lane changes, or adaptations of the host vehicle's speed and position, independently of driver input. Consequently, suitable HMI strategies have to be developed so that the driver can understand and trust the HAD functions in order to avoid unnecessary deactivation of the automation.

The HMI was designed on the basis of available studies on interfaces for automated driving. The HMI was implemented in a driving simulator and a small-scale expert study was conducted. The results can be summarized as follows. In general, the usability of the HAD function was rated as "good" to "excellent" on the SUS and the comprehensibility of the HMI was mostly rated as "high" to "very high". However, participants also made suggestions on how to further improve the HMI, such as decluttering the display. From observations of driver-vehicle interactions during the test situations, it became evident that unnecessary deactivations during automated maneuvers occurred very rarely. Regarding transitions to manual driving, it appears that the drivers used the advance information provided by the HMI, given that they mostly took over control of the vehicle in early stages of the take-over process.

Taken together, the results demonstrate that the proposed interface is a successful prototypical approach to the HMI challenges of cooperative HAD that were the focal point of this study. However, the results should be validated with a larger sample of naive participants and in more realistic test environments than the low-fidelity simulator used in this study. Furthermore, several HMI considerations were not taken into account in this study that might be crucial for successful human-machine interaction. For example, the *integration of auditory or haptic HMI elements* may be of great importance as drivers are expected to be occupied with non-driving related tasks during HAD [24–26]. Consequently, it may be necessary to direct the driver's attention to relevant information provided by the HMI. Suitable *timing* of the different messages (such as announcement of take-over situations, soft TOR and hard TOR) is possibly the most important prerequisite of safe and comfortable transitions between highly automated and manual driving. However, the perfect timing is yet to be determined. In conclusion, it has to be emphasized that the proposed HMI represents a promising approach to promote safe and efficient cooperation between drivers and HAD functions, but that there is still a need for significantly more research in this area.

Acknowledgments This work partly 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. We would like to thank Samantha Fritzsch for her support with the graphical HMI design and Martin Grein for the implementation of the HAD function in the driving simulation. Both Samantha Fritzsch and Martin Grein are with WIVW.

References

1. Naujoks, F., Grattenthaler, H., Neukum, A., Weidl, G., Petrich, D.: Effectiveness of advisory warnings based on cooperative perception. IET Intell. Transp. Syst. **9**, 606–617 (2015)
2. Rauch, A., Klanner, F., Rasshofer, R., Dietmayer, K.: Car2x-based perception in a high-level fusion architecture for cooperative perception systems. In: Intelligent Vehicles Symposium (IV), pp. 270–275. IEEE Press, New York (2012)
3. Wiedemann, K., Schömig, N., Mai, C., Naujoks, F., Neukum, A.: Drivers' monitoring behaviour and interaction with non-driving related tasks during driving with different automation levels. In: 6th International Conference on Applied Human Factors and Ergonomics, Las Vegas, USA (2015)
4. Bruemmer, D., Few, D., Boring, R., Marble, J., Walton, M., Nielsen, C.: Shared understanding for collaborative control. IEEE Trans. Syst. Man Cybern. A Syst. Hum. **35**, 494–504 (2005)
5. Woods, D., Tittle, J., Feil, M., Roesler, A.: Envisioning human-robot coordination in future operations. IEEE Trans. Syst. Man Cybern. C Appl. Rev. **34**, 210–218 (2004)
6. Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L., Nass, C.: Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. IJIDeM **9**, 269–275 (2015)
7. Saffarian, M., De Winter, J.C.F., Happee, R.: Automated driving: human-factors issues and design solutions. Proc. Hum. Factors Ergon. Soc. Annu. Meet. **56**, 2296–2300 (2012)
8. NHTSA: Preliminary Statement of Policy Concerning Automated Vehicles. National Highway Traffic Safety Administration, Washington, DC (2013)
9. Gasser, T., Westhoff, D.: BASt-study: definitions of automation and legal issues in Germany. Workshop on the Future of Road Vehicle Automation. Irvine, USA (2012)
10. 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.) 5th International Conference on Applied Human Factors and Ergonomics, vol. 7, pp. 2099–2106. AHFE Conference, Krakau (2014)
11. 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)
12. 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. Factors Ergon. Soc. Annu. Meet. **58**, 2063–2067 (2014)
13. 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? Trans. Res. F Traffic Psychol Behav **35**, 185–198 (2015)
14. Lüke, S.: Ko-HAF—cooperative highly automated driving. In: Autonomous Vehicles Test and Development Symposium, Stuttgart (2016)
15. Banks, V.A., Stanton, N.A.: Keep the driver in control: automating automobiles of the future. Appl. Ergon. **53**, 389–395 (2016)
16. Manca, L., De Winter, J., Happee, R.: Visual displays for automated driving: a survey. Automotive UI 2015—Workshop on Adaptive Ambient In-Vehicle Displays and Interactions (WAADI), Nottingham (2015)

17. Naujoks, F., Neukum, A.: Specificity and timing of advisory warnings based on cooperative perception. In: Butz, A., Koch, M., Schlichter, J. (eds.) Mensch & Computer 2014—Workshopband, pp. 229–238. De Gruyter Oldenbourg, Berlin (2014)
18. Totzke, I., Naujoks, F., Mühlbacher, D., Krüger, H.-P.: Precision of congestion warnings: do drivers really need warnings with precise information about the congestion tail´s position? In: Waard, D., Jamson, A.H., Barnard, Y., Carsten, O.M.Y. (eds.) Human Factors of Systems and Technology, pp. 235–248. Shaker Publishing, Maastricht (2012)
19. Totzke, I., Naujoks, F.: Der Fahrer im Zentrum eines vernetzten Strassenverkehrssystems - Herausforderungen und Chancen aus Sicht der Verkehrspsychologie [The driver and connected driving—challenges and chances from the perspective of traffic psychology]. Strassenverkehrstechnik **58**, 684–689 (2014)
20. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**, 50–80 (2004)
21. Lees, M.N., Lee, J.D.: The influence of distraction and driving context on driver response to imperfect collision warning systems. Ergonomics **50**, 1264–1286 (2007)
22. Naujoks, F., Neukum, A.: Timing of in-vehicle advisory warnings based on cooperative perception. In: Waard D.D., Brookhuis K., Wiczorek R., Nocera F.D., Brouwer R., Barham P., Weikert C., Kluge A., Gerbino W., Toffetti A. (eds.) Proceedings of the Human Factors and Ergonomics Society Europe Chapter Annual Meeting, pp. 1–13 (2014)
23. Brooke, J.: SUS-A quick and dirty usability scale. In: Jordam P.W., Thomas B., McClelland L. M., Weerdemeester B. (eds.) Usability Evaluation in Industry, pp. 189–195. Tayler & Francis, London (1996)
24. Politis, I., Brewster, S., Pollick, F.: Language-based multimodal displays for the handover of control in autonomous cars. Automotive UI 2015, Nottingham (2015)
25. Griffiths, P.G., Gillespie, R.B.: Sharing control between humans and automation using haptic interface: primary and secondary task performance benefits. Hum. Factors **47**, 574–590 (2005)
26. Petermeijer, S.M., de Winter, J.C., Bengler, K.J.: Vibrotactile displays: a survey with a view on highly automated driving. IEEE Trans. Intell. Transp. **17**, 897–907 (2015)

Part VII
Road and Rail—Logistics and Passengers

Child Restraint Seat Installation Errors and a Simple Mechanical Device to Mitigate Misuse

Xinqi Chen and William Altenhof

Abstract This research focuses on the development of a child restraint system installation-aid device (CRSIAD) for the purpose of mitigating child safety seat misuse in terms of installation. A geometric study was performed base on surveying dimensions of currently existing child safety seat products. Material property experiments were conducted in the lab to develop an anisotropic wood material model for CRSIAD in order to virtually evaluate device stress levels. Finite element analysis (FEA) of both the material model and CRSIAD were performed in comparison with lab test data to validate structural performance. The CRSIAD was then fabricated and finalized after multiple design iterations for geometry and components based on in-car testing. User satisfaction survey and professional review by certified CRS installation personnel were completed to ensure the value of CRSIAD as well as provide feedback on future improvements. The CRSIAD was believed to be an important contribution towards the improvement of child safety in vehicles.

Keyword Child restraint system installation aid device · Child safety seat

1 Introduction

LATCH (Lower Anchors and Tethers for Children) became standard equipment, by law, for all passenger vehicles on September 1st, 2002, which was enforced by NHTSA (Federal Motor Vehicle Safety Standard 225) [1]. This safety feature allows Child Restraint Seats (CRSs) to be properly fitted into the car backseat firmly so that CRS will perform at its best during a car accident, which drastically reduces the chances of injuries and fatalities of the child occupant.

The use of LATCH started with a good concept but formed a new direction of misuse [2]. LATCH did not increase the safety for children in car accident due to its

X. Chen · W. Altenhof (✉)
Department of Mechanical, Automotive, and Materials Engineering,
University of Windsor, Windsor, ON N9B 3P4, Canada
e-mail: altenh@uwindsor.ca

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_50

unfortunate misuse. The leading cause of CRS misuse is loose installation, which is caused by human error due to both lack of understanding and physical difficulties. The latest experimental survey, completed between November 2013 and May 2014 at Oregon Health and Science University Hospital, has revealed a disturbing amount of CRS misuse among 267 participating families. Ninety three percent of installations have contained at least one critical error, out of which forty three percent involved loose installation [3] of either the upper or lower LATCH tethers.

CRS misuse is not only caused by education and lack of understanding but also the physical effort needed from the parents and caregivers who perform the installation. An important experiment was conducted in 2006 to specifically study the physical effort needed to complete the CRS installation. Electrogoniometers were fitted on participants' bodies to measure their body joint extension and radial deviation [4]. Additionally, muscle activation was also assessed in the study. It was found out that some muscles had peak levels of 100 % of maximum activation during seat installation and tethering. Peak levels were observed to be over 90 % of maximum in 18 of 30 muscles. Results showed that shoulders were forced into extreme postures during all 3 tasks (install seat, secure tether and place child). Even for large size vehicles, the necessary workspace to for CRS installation is tight and requires the installer to bend down and sit on his/her knees to complete installation.

Correct CRS installation is very physically demanding. Maximum effort of some muscles could result in inadequate securing of the CRS into the vehicle. The seat installation typically requires high efforts and awkward postures for tightening the upper and lower LATCH straps. These demanding tasks result in loose installation being a frequent misuse characteristic. Correspondingly, the need for a supplemental device, to reduce human effort during installation, could play a significant role in reducing the CRS install errors and thus mitigate misuse and increase child safety in vehicles.

2 Research Scope

Specific to loose LATCH tether connections, CRS misuse is a result of both human error and inability to provide suitable preloads to the tethers. Additionally, the constraints associated with the CRS LATCH implementation and limited working space within a vehicle interior also contribute towards installation errors. Correct installation of LATCH tethers is vital for a CRS to operate safely in real-world driving situations. Tightening LATCH straps in the vehicle is a challenging task due to space limitation and physical effort needed [4].

In this research, computer aided engineering (CAE), manufacturing, and physical testing was conducted to develop a Child Restraint System Installation Aid Device (CRSIAD) for the purpose of mitigating CRS installation errors. The main focus of this research can be summarized in the following points:

1. To study and build a material model that is suitable to be used for developing the CRSIAD. Standardized material mechanical characterization is needed to assess the performance of lightweight wood materials, for data needed in the generation of an anisotropic material model. The developed material model will be validated to the experimental data using a suitable quantitative means of model validation.
2. To design a CRSIAD mechanism that will preload the CRS into the car seat, thus reducing the physical effort for installing CRS and potentially mitigating misuse.
3. To construct the CRSIAD geometry based on given CRS geometry and space envelop developed from the surveying of existing CRS products among major suppliers in the market. Following the creation of such CAD data a full FEA model of the CRSIAD will be developed and also incorporate the material model developed in item (1) above. Simulation of the CRS installation loading conditions for both forward-facing and rearward-facing configurations will also be considered. The appropriate geometry design iterations will be performed, using FEA, until the loading stress level is within the yield limit of the material and weight is quasi-optimized.
4. To fabricate the CRSIAD according to the CAE efforts identified in item (3) above. Following the manufacturing, the device will be tested in a number of popular production vehicles. Design iterations will be completed to ensure appropriate performance and ease of use.
5. To obtain user satisfaction input for further design enhancement of the CRSIAD. A comparison of CRS installation ease, with and without the use of the CRSIAD, will also be considered.

3 Material Model Development

A suitable material model is vital for performing FEA analysis for CRSIAD. Plywood is an upgraded version of conventional sheet wood and is widely used in household and construction applications. It presents good strength, weight, hardness and low cost. Compared to conventional wood sheets, its strength is consistent across all directions of the panel. Its multiple (odd number) plies prevent warping and provide improved rigidity compared to other conventional wood sheets [5].

An anisotropic material model was used in modeling the plywood and its performance under loading conditions. Given the nature of the problem to be suited, LS-DYNA was selected as the finite element solver for CAE studies. A material model applicable to wood, namely, *MAT-143 (MAT_WOOD), was chosen to simulate the plywood for its ability to model orthotropic wood materials. This model is also applicable for all varieties of wood when appropriate material parameters are selected [6].

To acquire these parameters, good quality plywood sheets were first purchased for material testing, namely a -Baltic Birch 4 foot by 8 foot 3/4-inch plywood sheet

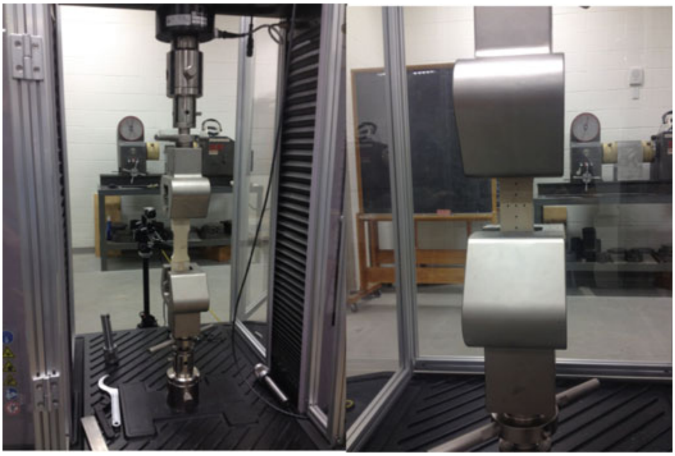


Fig. 1 MTS machine testing for tensile test in the parallel orientation (*left*) and perpendicular orientation (*right*)

from Miller's Millwork & Hardware, Windsor, Ontario. All test specimens of this material were conducted according to specific wood testing ASTM standards [7].

Six primary tests were completed for different orientations of the plywood sheet, namely, -tensile and compression tests in the fibre parallel direction, tensile and compression tests perpendicular to the fibre directions, and shear test in parallel and perpendicular to the fibre directions. Three-point bending test specimens were constructed to validate the material model. Test setups for tensile in parallel and compression orientation are shown in Fig. 1.

Once all sets of load versus deflection and strain data were collected, they were converted to stress versus strain data according to ASTM standards [7]. With material stress data acquired, critical parameters such as Young's Modulus, Poisson's ratio, yield strength and ultimate strength were used to construct the material model. The material model was then used to simulate all the test events in the lab to match the raw data in order to ensure adequate predication performance of the material model. Finally, simulation results of three-point bending test were compared with raw test data to complete the validation process.

Equation (1) was used to calculate validation metrics, where V is the validation metric, L is the crosshead displacement, L_1, L_2 are the initial and final values of crosshead displacement., $R_{Exp}(L)$ is the experimental load value, and $R_{Theory}(L)$ is simulation load value [8].

$$V = 1 - \frac{1}{L_2 - L_1} \int_{L_1}^{L_2} \tanh\left(\left|\frac{R_{Exp}(L) - R_{Theory}(L)}{R_{Theory}(L)}\right|\right) dL \tag{1}$$

Accumulated error has been calculated by Eq. (2).

Table 1 Maximum stress level from simulation for both forward-facing and rearward-facing configurations

	Forward-facing Stress at max loading (MPa)	Rearward-facing stress at max loading (MPa)	Yield strength (MPa)
Tensile parallel to fibre direction	3.576	3.368	17.638
Compression parallel to fibre direction	3.135	1.553	29.818
Tensile perpendicular to fibre direction	0.218	0.279	0.327
Compression perpendicular to fibre direction	1.537	1.556	11.546
Shear parallel to fibre direction	0.952	1.152	5.490
Shear perpendicular to fibre direction	0.813	0.931	3.137

$$E(Error) = \frac{1}{L_2 - L_1} \int_{L_1}^{L_2} \left| \frac{R_{Exp}(L) - R_{Theory}(L)}{R_{Theory}(L)} \right| dL \qquad (2)$$

Two conditions associated with the amount of deformation were used to calculate the validation metrics and accumulated error; validation metrics for the entire range of data (from test start to material failure) and validation metrics prior to material yielding. For our specific application, the latter criterion is more appropriate since deformation prior to yielding is expected during service. The validation metric was calculated to be 0.974 for the entire range with accumulated error of 0.053. Values prior to yielding were computed to be 0.975 and 0.052 for validation metric and accumulated error, for the range prior to yielding, respectively. Material yield strength data is available in Table 1.

4 CRSIAD Design and Function

4.1 Dimensional Envelop Definition

The dimensional envelop of the CRSIAD is highly dependent on the dimensions of the CRS. To define such a dimensional envelop, surveying the geometry of over 30 existing CRS products was conducted for measuring existing CRSs from major suppliers in the market. Figure 2 shows two main aspects of the dimensional envelop—widths (inner and outer) and side profile limits. For inner width, the

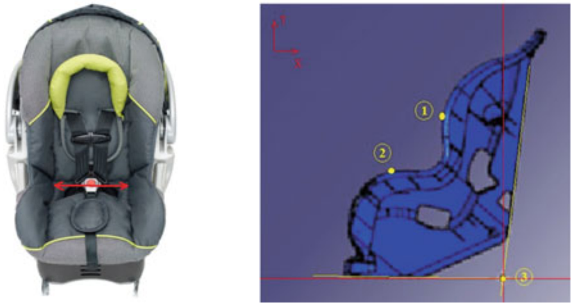


Fig. 2 Model name Baby Trend Flex Loc Infant Seat (*red double arrow* shows narrowest inner width) (left) and CRS side profile coordinate system with critical Point 1, 2 and 3 (*right*)

narrowest value on the seating area (Fig. 2 left) was measured and recorded. The smallest surveyed value for inner width was 215.9 mm with average value of 252.49 mm of all surveyed units. The greatest surveyed value for outer width was 510.54 mm with average value of 469.20 mm. The seating width of the CRSIAD has to be smaller than the smallest surveyed inner width value and therefore determined to be 190.5 mm.

Due to the variations among the CRS models in the market, critical points were selected to limit the side profile of the CRSIAD (Fig. 2 right). This is a typical forward facing configuration when the CRS is installed in the car. Point 3 is the virtual lower anchor points in the vehicle seat. Typically this point is approximately located at the intersection of the two yellow lines in Fig. 2 (right). Point 3 was conveniently defined as the origin in this configuration in the x/y coordinate axes. Point 2 is the highest point of the bottom side wings of the CRS and Point 1 is the left-most point of the bolster side wings in the side view. Both Point 1 and Point 2 were considered as the peak points of the two wing sections. Their coordinates were defined based on the origin (Point 3). These coordinates were recorded and only one coordinate for Point 1 and 2 were selected; Point 1 having the smallest x-value (negative) and Point 2 with greatest y-value. These two x and y coordinates were used as boundaries in this plane for the CRSIAD.

4.2 *Geometry Design of the CRSIAD*

Main Frame Shaping. The contact section at the base and back of the CRSIAD needed to be matched closely to existing CRS geometry to avoid any alignment or load concentration issues.

Existing CAD data of a CRS geometry was used to create the bottom section of the CRSIAD so that the bottom profile was appropriately aligned between the two entities. The matched contact surfacing allows the load to be distributed evenly with minimum to zero point-loading, ensures stability during the preloading process.

Main Loading Mechanism. A ratchet strap assembly was used as the main loading mechanism due to its robustness, compact size and availability. When properly used, the ratchet strap can appropriately preload with minimal rotation, motion and effort, which satisfies the need of a reasonably short amount of loading time. The loading points of the ratchet webbing must be located out of the boundary lines from dimensional envelop to avoid any contact between itself and the CRS body.

Loading Structure. With a ratchet strap loading assembly, the need for a loading beam was identified to permit the strap webbing to run across the CRSIAD in order to appropriately preload, on both sides of the installation aid, the CRS within the vehicle seat. The width of the loading beam has to be wider than the greatest surveyed CRS outer width value (Sect. 4.1) and determined to be 550 mm.

Seatbelt D-rings were used for minimizing strap friction in the system and act as swivel point on the loading beam where the webbing can pass through. The loading beam design is shown in Fig. 3.

CRSIAD Restraint System. The ratchet strap assembly ends are needed to be restrained to appropriately preloading the CRS into the vehicle seat. The existing car seatbelt system was found to be useful locations where the ends of the webbing could be buckled or clipped into existing buckles of the vehicle. To prevent relative motion between the belts on the same seatbelt, a H-clip (also known as a stop clip) can be implemented in the installation.

With all components mentioned above combined, a concept restraint system was designed to serve the loading structure to the CRSIAD (Fig. 4).

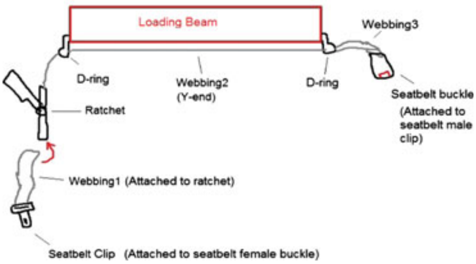


Fig. 3 Loading beam design

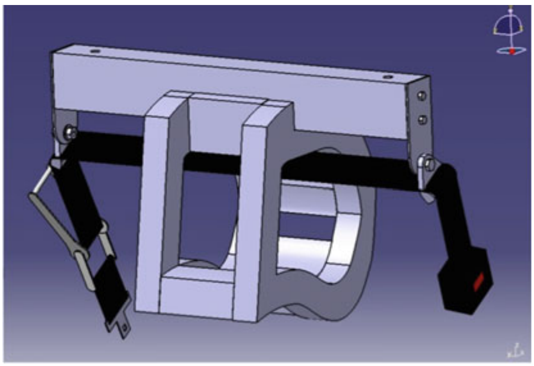


Fig. 4 Complete concept CAD model of the CRSIAD

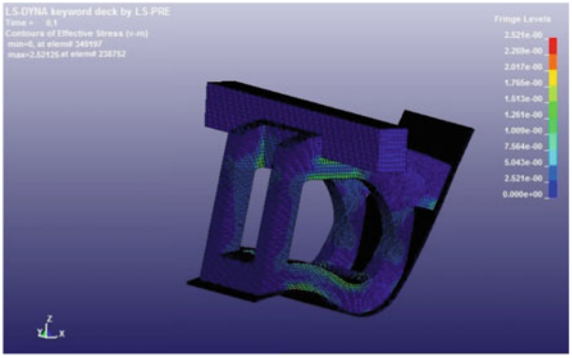


Fig. 5 Simulated FEA model of CRSIAD under forward-facing configuration under maximum loading at 1.8 kN (units for Time: second; Stress: MPa)

5 CRSIAD Virtual Testing

An FEA model of the CRSIAD was constructed and simulated in LS-DYNA with the validated material model. Assembly modeling was simplified to three parts, namely, (1) loading beam without attachments, (2) main frame and (3) CRS seating base. Both forward-facing and rearward-facing configurations were simulated under maximum loading at 1.8 kN (400lbf) to ensure the effective stress is within the material yield strength. With design iterations to address geometry issues, the final FEA model is shown in Fig. 5 for the forward-facing loading configuration. Maximum stress levels of both forward and rearward configurations are presented in Table 1 in appendix section (Sect. 10).

6 CRSIAD Fabrication

6.1 Main Frame

Part A from Fig. 6 was made by first gluing two sheets of given plywood together to achieve the desired thickness. The glued sheet was then cut by computer numeric control (CNC) machine. Two pieces of Part A were constructed, one on each side of

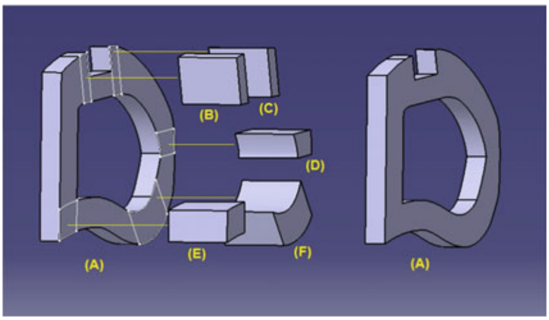


Fig. 6 Assembly diagram for CRSIAD main frame

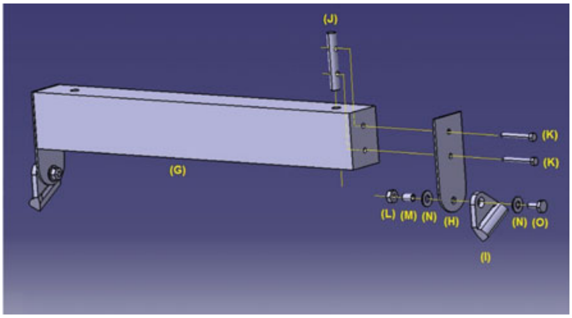


Fig. 7 Loading beam assembly

the device. Parts B and C were cut and trimmed from one sheet of plywood. Part E, F, and G were all made in similar fashion as Part A (glued and cut).

Parts A to F were then glued together with wood glue as shown in Fig. 6. Each connection point was also reinforced by wood nail to ensure strength.

6.2 Loading Beam

The main beam structure was made of three sheets of plywood glued together to achieve its desired thickness. The wood orientation has the parallel direction of the sheets in the vertical direction to match the FEA model strength predictions.

Aluminum plates, having a thickness of 0.125 inch (Part H in Fig. 7) were cut with a CNC machine. Barrel nuts and attaching bolts are also incorporated into the design as presented in Fig. 7. Barrel nuts were simply placed into the vertical holes of the beam and then fastened by the bolts to sandwich the aluminum plate in place. Sleeve (M) was cut from a section of tube to match the inner diameter of D-ring eye. Bolts (K) were obtained according to the size of Sleeve (M). Bolts were hand-torqued with spring-locking washers applied in-between. Free movement of D-rings for rotation was checked and additional grinding was done to the sleeve insert to ensure smoothness of D-ring rotation.

6.3 Loading Mechanism

The ratchet assembly and webbing were obtained from hardware stores. Loading capacity of the strap is 2.2 kN where as ratchet is rated at 6.7 kN respectively.

The design of the system requires the end attachments of the webbing to be connected to the existing car seatbelt system. Compatibility of end clips and buckles in the CRSIAD to production vehicles had to be ensured. Due to the variety of the clips and buckles being used on production vehicles across different

Fig. 8 Successfully preloaded CRS with the aid of CRSIAD in test vehicle

manufacturers, a survey was done to find out all the types of clips and buckles existing on the cars in the market.

Over 30 attainable vehicles were surveyed for this matter from different sources. Through this compatibility examination, two particular clips and buckles were proven to work all the surveyed vehicles and therefore chosen as part of CRSIAD ratchet webbing assembly.

6.4 Complete Assembly

The complete CRSIAD assembly is presented in Fig. 8 and illustrates the device engaging both the CRS and automobile.

7 CRSIAD Testing and Critical Review from Certified CRS Installers

User feedback from certified CRS installers was an important aspect of the assessment of the CRSIAD and was needed to verify its functionality as well as improve its design. With approval of the local Ontario Provincial Police (OPP) Constable S. Coulter (based at Windsor, ON), the CRSIAD was allowed to be brought into a child seat clinic session on February 27th 2016 to be reviewed by OPP officers and other trained and certified CRS installation experts.

The device was tested and demonstrated on a convertible CRS (Safety 1st - Alphaomega 65) in a 2007 Ford Freestyle. A forward-facing configuration installation for the CRS was performed, as shown in Fig. 8. The CRSIAD was used to preload the CRS into the second-row car seat of the vehicle followed by completion of CRS installation. Installation was performed by one of the co-authors of this manuscript while observed by OPP officers. The CRS was successfully installed with the aid of CRSIAD.

After observing the installation, OPP constable S. Bertoni (based at Harrow, ON) was generally satisfied with the performance of CRSIAD in terms of its functionality. The fact that CRSIAD was loading the CRS towards the lower anchor direction was appreciated and noted as the most effective pre-loading direction for CRS. Officer Bertoni also believed this device was easy to use. He suggested further reduction of the material as he believed some areas of the structure where excessive for the loading condition associated with installation. Additionally, he commented that the size of the CRSIAD could be more compact. Furthermore, he suggested that the ratchet device would be best if it was located on top of the loading beam so that it would be easier to operate.

OPP Auxiliary S. Brazil, who also complimented a trial CRS installation with the CRSIAD device, had provided positive feedback on the CRSIAD's functionality as well as easiness to operate. He had also suggested that the size of the device could be reduced and indicated that—perhaps a different type of tightening mechanism could be adapted to shorten the time necessary for pre-loading. He indicated that the device has the potential to be used at car seat clinic events where time is important.

OPP Constable S. Coulter supported the use of the CRSIAD. She suggested a load-measuring device to be fitted in the ratchet assembly so that users of the CRSIAD would know when to stop ratcheting. This would also be a feature to prevent any potential damage to the CRS or CRSIAD itself.

Certified child seat installer Mr. K Czubernat (FEA Simulation Engineer from FCA Canada, Windsor, Ontario) had successfully performed installation of a CRS with aid of the CRSIAD in his personal vehicle (2005 Saturn Vue). Easiness of using the device was also indicated after completing the installation. An important afterthought was also provided; any excessive preloading force may prevent users themselves from removing the CRS at later time when necessary. This comments further justified the need for some form of load measuring device to provide user feedback to ensure that the LATCH webbing is not tightened.

OPP officers and engineers generally appreciated the use of the CRSIAD to help installing CRSs and ensuring a suitable load within LATCH webbing. They all believed that the CRSIAD project is a good start in an effort to mitigating CRS misuse as well as reducing physical effort of installing a CRS into a passenger car.

8 Summary

This research effort has resulted in the design, manufacturing and successful implementation of a CRS installation aid. The use of the CRSIAD has reduced physical effort needed to install CRS. The ratcheting mechanism was identified to be quick and effective which will meet most users' requirement for saving time during CRS installation. The CRSIAD is also compact and lightweight that favours many users. Improvements have been suggested by professionals and certified CRS installers, which may see future implementation in subsequent design iterations. The CRSIAD is a successful starting point on the path towards reducing CRS installation errors and improving child safety in vehicles.

Appendix: Material Yield Strength Data

see Table 1.

References

1. FMVSS 225: Child restraint anchorage systems, 1999. Federal motor vehicle safety standards, Standard No. 225; U.S. Department of Transportation, Federal Motor Carrier Safety Administration, Washington, D.C., USA
2. Decina, L.E., Lococo, K.H., Doyle, C.T.: Child restraint use survey: LATCH use and misuse, DOT HS 810 679. TransAnalytics, LLC, Kulpsville, PA, USA (2006)
3. Hoffman B, MD, FAAP, Carlson K, MS PhD, Lindsay DIckey MPH, Adrienne Gallardo, MS, Marianne Bridwell-Chapman BA: Unsafe from the start: critical misuse of car safety seats for newborns at initial hospital discharge. OHSU Doernbecher Children's Hospital, Portland, OR, OHSU, Portland, OR, USA (2014)
4. Brown SHM, Grondin DE, Potvin JR: Strength limitations to proper child safety seat installation: Implications for child safety. Department of Kinesiology, University of Waterloo, Department of Kinesiology, McMaster University, Department of Kinesiology, University of Windsor, Windsor, Ontario, Canada (2007)
5. Handbook of Finnish Plywood: Finnish forest industries fedaration. Lahti, Finland (2002)
6. Murray, Y.D.: Manual for LS-DYNA wood material model 143, FHWA-HRT-04-097, Research, Development, and Technology Turner-Fairbank Highway Research Center, 6300 Georgetown Pike. Mclean, VA, USA (2007)
7. ASTM: Standard test methods for small clear specimens of timber. pp 143–194. ASTM International, West Conshohocken, PA, USA (2009)
8. Oberkampf WL, Trucano TG: Verification and validation in computational fluid dynamics. Prog. Aerosp. Sci. **38** (2008)

Experimentation System for Determining Bus Routes for Customers of Supermarket Chains

Artur Pohoski, Leszek Koszalka, Iwona Pozniak-Koszalka and Andrzej Kasprzak

Abstract It may be observed that supermarket chain companies invest in special buses to deliver customers directly to a given store. The companies expect a tool which allows to design the routes of buses in order to minimize costs (maximize profits). The objective of this paper is to present the created computer experimentation system (simulator) with the designed and implemented algorithms to determine the optimal bus route for customers. The bus route problem was divided into two stages. At the first stage, the three designed heuristic algorithms called Most Occupied, High Gain Neighbor, and Cut the Worst are responsible for selection (choosing) of the location of bus stops. At the second stage, the five algorithms allow to determine a route between previously chosen stops. These algorithms are based on approaches from so-called Artificial Intelligence area: Simulated Annealing (SA), Taboo Search (TS), Genetic Search (GS), and Ant Colony (AC) as well as on the simple ideas like Random Search (RS). In the paper, the results of the investigations made with the created experimentation system are presented, concerning an adjustment of the parameters of any algorithm, and a comparison of the algorithm's efficiency on both stages.

Keywords Bus route planning · Optimization · Algorithm · Computer system · Simulation · Experiment

A. Pohoski
Metegrity Inc., Edmonton, Canada
e-mail: artur.pohoski@gmail.com

L. Koszalka (✉) · I. Pozniak-Koszalka · A. Kasprzak
Department of Systems and Computer Networks,
Wroclaw University of Technology, Wroclaw, Poland
e-mail: leszek.koszalka@pwr.edu.pl

I. Pozniak-Koszalka
e-mail: iwona.pozniak-koszalka@pwr.edu.pl

A. Kasprzak
e-mail: andrzej.kasprzak@pwr.edu.pl

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_51

1 Introduction

Nowadays, many technical, economic, logistic issues and other problems, which we may encounter in daily practice, is stated in the form of optimization task which consists in finding the optimal solution—the best in terms of quality. However, there is a very time-consuming procedure when trying solving with the traditional methods. The alternative is using heuristic methods and creating algorithms for solving which are based on the artificial intelligence. Heuristic and meta-heuristic algorithms provide an opportunity to find an approximated solution in an acceptable period of time [1, 2].

The aim of this paper is to create the system, which could support the process of planning the bus route to delivering customers to the supermarket. The main criterion, which is taken into account, is a gain considered as the total tickets revenue (a price of a single ticket multiplied by the number of customers at travel) reduced by the cost of a route (cost per meter multiplied by a length of the route). The created experimentation system which contains a module with the implemented algorithms allows adjusting parameters for any particular algorithm and estimating the total cost (gain) of the route.

The rest of the paper is organized as follows. The considered problem is formulated in Sect. 2. Section 3 is focused on the description of the proposed algorithms for solving the problem at the first and the second stage. The implemented experimentation system is presented in Sect. 4. The results of the investigations concerning the adjustment of parameters of the algorithms, and the comparison of the efficiency of the algorithms are shown in Sect. 5. The final remarks appear in Sect. 6.

2 Problem Statement

The problem can be considered as a modified version of the traveling salesman problem (TSP) [3]. There are given distinguished points inside a town which can play a role of bus stops. Moreover, there are given the estimated numbers of customers positioned in these points. The route should connect a chosen starting point (usually the base station) and the final destination, i.e., the bus stop in the front of the supermarket. The connected points can form a graph with known weights of edges which are corresponded to the cost of traveling along with the edges (e.g., distances, cost of fuel, cost of driver's wages, time of traveling, etc.). The problem is to select some points (from the all considered points) for choosing bus stops in these locations and to find a path from the starting node to the destination node which has to ensure visiting all the selected bus stops (vertices of the weighted graph) such that to maximize the defined gain (profit) of the company. The problem is a NP-hard [2, 3] and may be interpreted as skipping some less profitable bus stops (which constitutes the set Z) what can cause necessity of changes in using bus by some of customers. We introduce a minimum percentage of service parameter denoted as L, i.e., a minimal percentage of all potential customers. The mathematical formulation of the problem is as follows:

Given: N—the set of possible bus stops, M^*_n—the number of the estimated customers awaited on the n-th potential bus stop (n = 1, 2, …, M), M—the total number of all customers, C—the ($n \times n$) matrix of the weights (costs), I—the starting bus stop, D—the destination bus stop.

To find: the route R from I to D connecting the selected v points ($v \leq N$).

Such that: the gain G(R) expressed by (1) is maximum.

$$G(R) = \max \sum_{i=0}^{n} \sum_{\substack{j=0 \\ i \neq j}}^{n} (n_i p_t - p_f c_{ij}) x_{ij} \tag{1}$$

where n_i is the number of people on the i-th stop, p_t is the ticket price, p_f is the fuel cost, c_{ij} is the cost (distance) between i-th stop and j-th stop; x_{ij} equals to 1, if the connection between i-th stop and j-th stop belongs to the resulting route R, elsewhere equals to 0.

Subject to the constraint: {[the sum of M^*_n for v selected points/M] 100 %} L.

3 Algorithms

We proposed to divide the bus route problem into two stages. At the first stage, the selection (choosing) of the location of bus stops is made with one of the *stops selection algorithms*. The resulting set of the selected bus stops is an input data at the second stage. At this stage the final solution is obtained using one of the *best route determining algorithms*.

3.1 Bus Stops Finding Algorithms

Three algorithms of the selection of bus stops were designed and implemented: Most Occupied (MO), Cut the Worst (CTW), and High Gain Neighbor (HGN).

Most Occupied (MO).

1: Selecting the bus stop with the most quantity of students.
2: Checking if the minimal percentage of the collected customers is reached, i.e., the inequality (*sum*/*M*)*100 > L) is satisfied: if yes go to Step 3, otherwise go to Step 1.
3: Set of bus stops Z.

Cut the Worst (CTW).

1: Determining the shortest route between all the stops (TSP). Stops with the smallest amount of awaiting customers are being removed
2: Checking a gain of the whole route and comparing with the previous value of a gain
3: If a gain (income) is profitable, it is saved, otherwise bus stop is restored.
4: Algorithm keeps analyzing the next stops till the minimal percentage of collected customer's threshold is reached or all of the stops are analyzed: if yes go to Step 5.
5: Set of bus stops Z.

High Gain Neighbor (HGN).

1: Analyzing the starting/initial point.
2: For every analyzed stop, three closest stops are searched.
3: Choosing the one bus stop of them—this with the biggest gain.
4: This bus stop becomes the next stop being analyzed.
5: All the time is checking if the minimal percentage of collected customers is reached: if yes go to Step 6.
6: Set of bus stops Z.

3.2 *Best Route Determining Algorithms*

The five metaheuristic algorithms were implemented: Simulated Annealing (SA), Taboo Search (TS), Random Search (RS), Ant Colony (AC), and Genetic Search (GS). More detailed information about ideas of SA, AC, and GS can be found in [4–6].

Simulated Annealing (SA)

1: At the beginning input solution processed with Greedy Algorithm is considered to be the best. Result is copied to arrays which store current and temporary solution.
2: Randomizing two elements of route which will be replaced in the permutation test.
3: Comparing the length of a new route with a current one. If the length is shorter, then a temporary solution is copied to the current one.

4: Checking if current solution is better than solution which has been till now. In this moment the best route becomes the current.
5: If the current solution is better than a test, then the mechanism which prevents finding the solution in a local minimum is activated. Worse solutions are accepted with the probability: $P(T, \pi', \pi) = e^{-\frac{f(\pi')-f(\pi)}{T}}$, where $f(\pi)$ is a length of the route before replacement, $f(\pi')$ is a length of the route after replacement.
6: After each algorithm iteration, T parameter decreases according to $T = \frac{T}{1 + \lambda T}$, i.e., the lower temperature is, the probability of solving is higher.
7: After reducing the temperature, the algorithm returns to Step 1 in which the elements are selected to replace.

Taboo Search (TS)

1: At the beginning input solution processed with Greedy Algorithm is considered to be the best. Result is copied to arrays which store current and temporary solution.
2: The list (array) is reset to zero and then eight-element array of the elements to replace is created.
3: Randomizing four pairs of elements for replacement if taboo value on list equals 0. 4: Checking which place replacement of randomized elements gives the best result (shortest route). After designating best local solution replace two chosen elements (considered as the best).
4: These two elements are initialized with values of the length of taboo.
5: Checking if a current route is shorter that the best route till now. If yes, actual solution becomes the best. Go to Step 2.

Random Search (RS)

1: At the beginning input solution processed with Greedy Algorithm is considered to be the best and copy to current solution.
2: All elements which are going to be substituted in current solution are randomized.
3: Checking if actual route is shorter than the best till now. If yes, the best solution is copied to actual. The whole process is repeated with the given number of iterations.

Ant Colony (AC)

1: Random placement of ants on the route—from the list of stops choosing random starting point ants.
2: Pathway ant—ant following the formula $Fp_{ij}(t) = \tau_{ij}(t)\eta_{ij}$—choosing the most optimal bus stop which maximizes F function, until all the stops will be visited, where
3: Pheromone evaporation—pheromone matrix is multiplied by the parameter called "elusiveness of pheromone" what is interpreted as a cause that the used paths disappear and the better path is found. Go to Step 2.

Genetic Search (GS)

1: Designating an initial pool of N chromosomes. The first individual of this pool is chosen with the Greedy Algorithm, rest by Random Search.
2: Sorting the pool. Finding K best K = N * k best individuals, where k is the parameter called the quantity of best individuals in the pool. Parental pool containing K best individuals is created.
3: Crossing operation (simple crossing).
4: Mutation operation for all descendants. It is not check whether the result is better after making a change.
5: Self-improvement functions for all descendants. If solution turns out to be better, it is replaced with the current one.
6: Designating the best individual of generated pool of descendants. Turning back to the point of "choosing the parental pool" with a new population. Go to Step 2.

4 Experimentation System

To make an opportunity for finding the optimal routes, and making investigation concerning adjusting the algorithms and testing their properties—the input-output experimentation system was designed and implemented in C# environment. In preparing the system some ideas presented in [7, 8] were utilized.

- *Input parameters*:
  - U1—A vector of the possible bus stops.
  - U2—A vector of the numbers of customers at any of the possible bus stop.
  - U3—A matrix of the costs (distances) between possible bus stops.

- *Output/input parameters*:
  - Q1—A vector of the selected bus stops.
  - Q2—A matrix of the costs (distances) between the selected bus stops.
- *Output parameter*:
  - R—The result.

The main window of the program can be visualized on the map (Fig. 2)—the founded by an exemplary algorithm connection between bus stops is drawn. The simulator can create also a report of each investigation with the detailed results (Fig. 1).

Two averaging methods were used: the exponential smoothing and the moving average [9]. They helped to designate optimal values of parameters, because they significantly improved legibility of a chart.

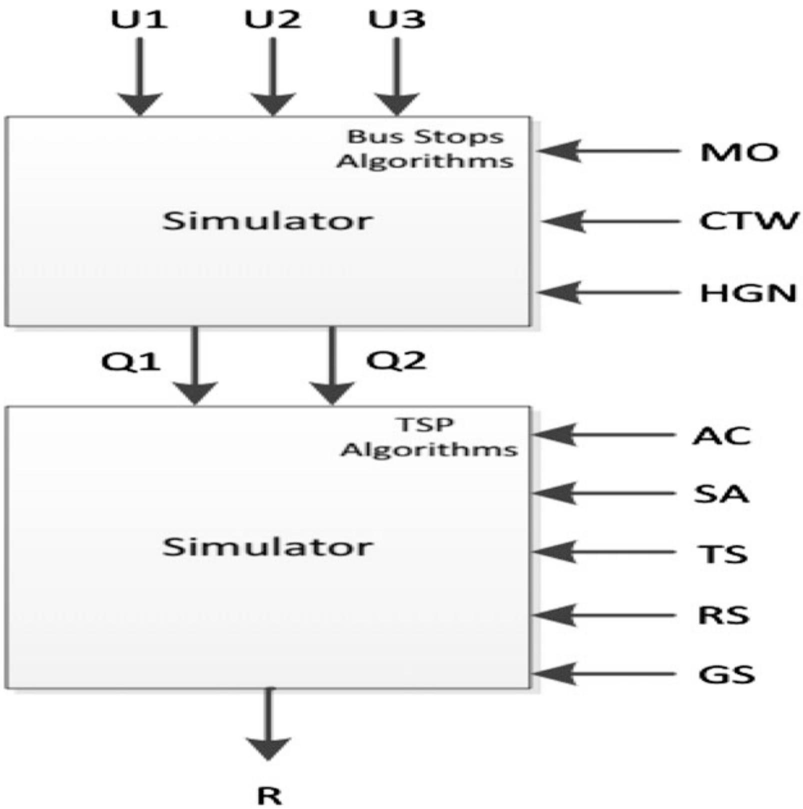


Fig. 1 Block-diagram of the input–output two stage experimentation system

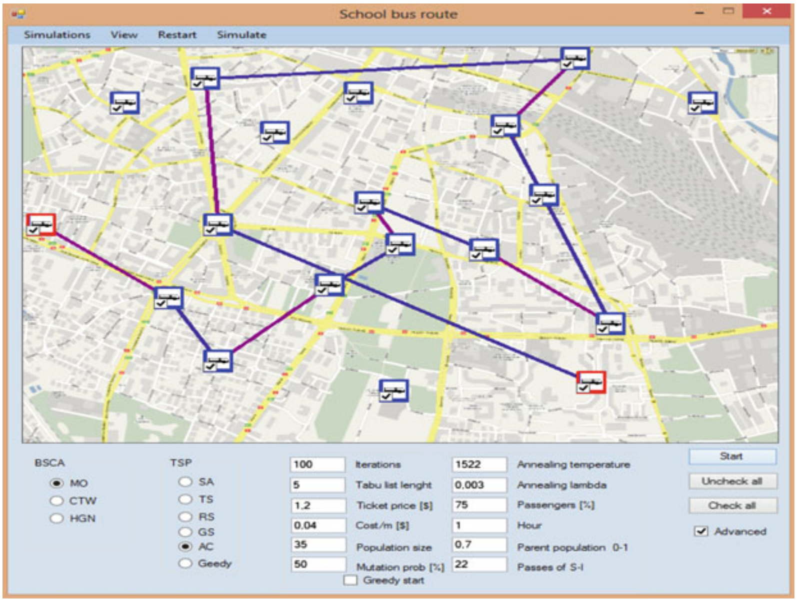


Fig. 2 An example of a screen view given by the visualization module of the system

5 Investigation

The created algorithms were implemented as modules of the experimentation system. All experiments were performed on the PC class computer with AMD E-450 (1.65 GHz dual core) processor and 4 GB of RAM. The system allows for two-level experiments. At the first level, the objective was to find the best values of inner parameters of the algorithms. At the second level, the algorithms (with the best parameters determined at the first level) compete—the objective was to find the most efficient algorithms for planning bus routes. The algorithms considered at the second level were the combinations of bus stops finding algorithms and best route determining algorithms. The research idea is partially based on [10]. The results of four complex experiments are presented in the consecutive sub-sections.

5.1 Experiment 1. Parameters of Best Route Determining Algorithms

For all tests carried out with these algorithms, the input data (the set of bus stops) was found using Most Occupied algorithm. For all tested algorithms, the single experiments were carried out 100 times for 500 iterations of any algorithm, except GS for which were taken 50 times for 300 iterations of the algorithm in view of its

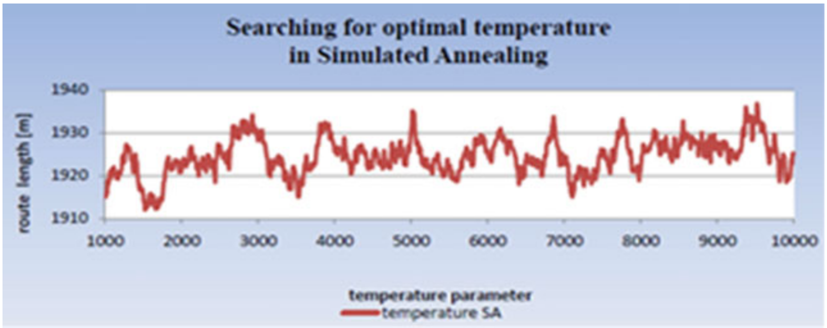


Fig. 3 Temperature parameter in SA algorithm

significantly longer working time. The experiments were performed with a big precision, e.g., for the initial temperature (for SA algorithm) were carried out in the range of <0.001; 1> with step of 0.001, but for the temperature in the range of <1000; 10000> with the step of 9 (see Fig. 3).

The obtained results (the averaged values were analyzed) allows for selecting the following values of the parameters: for SA: temperature T = 1522, lambda = 0.003; for TS: length of the list = 5; for AC: the number of ants = 3, the number of pheromone = 1.41, elusiveness of pheromone 0.47; for GS: initial population size = 35, parental pool size = 0.7, mutation probability = 50 %, the number of self-improvements = 22.

5.2 Experiment 2. Parameters of Bus Stops Finding Algorithms

This experiment concerned the bus stops finding algorithms, including MO, CTW, and HGN. Studies were carried out in a range of <1; 1000> iterations. For each *best route determining algorithm* there were made tests for all kinds of *bus stops determining algorithms*. The obtained results allowed to compare them and to draw a conclusion about the most optimal solutions in a given conditions. An example for the AC algorithm is presented in Fig. 4. The gains obtained for SA and TS were not so high as for AC while RS gave the worst results.

5.3 Experiment 3. Percentage of Collected Customers

This experiment consisted on checking dependence of the minimal percentage *L* of the collected passengers (customers) and the obtained gain. The tests were made for all bus stop selection (finding) algorithms and all best route determining algorithms.

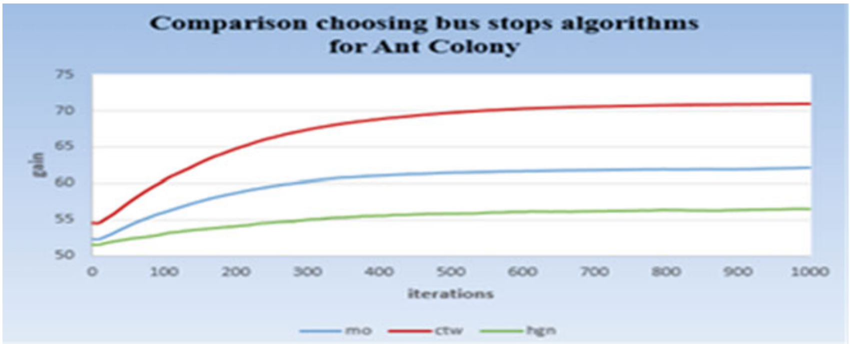


Fig. 4 The gain function in relation to the number of iterations for three bus stops finding algorithms connected to AC algorithm

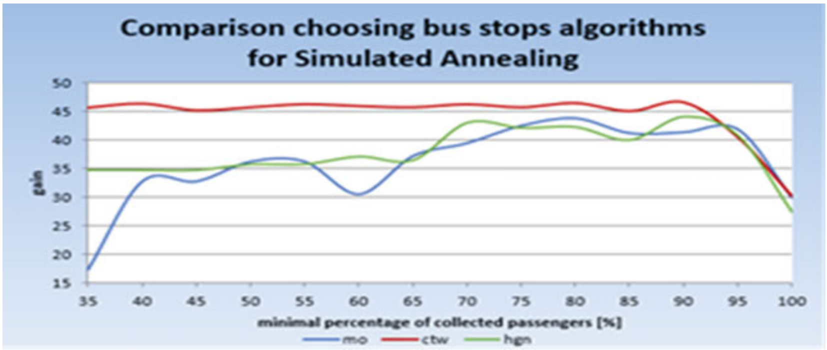


Fig. 5 Gain related to parameter *L* and bus stop finding algorithms for GS algorithm

The simulation were made for various parameter L which was tested in the range of <35; 100> with the step of 5 %. The exemplary results obtained for GS are in Fig. 5.

It may be observed that the most stable bus stop finding algorithm with the highest gain is CTW.

5.4 Experiment 4. Comparison of Compositions of Algorithms

The goal of this experiment was studying the properties of the combinations of the algorithms. Firstly, for four path (bus route) determining algorithms, the best bus stops finding algorithm were chosen. Such a selection was made basing on the best received value of the gain. After several preliminary experiments the four

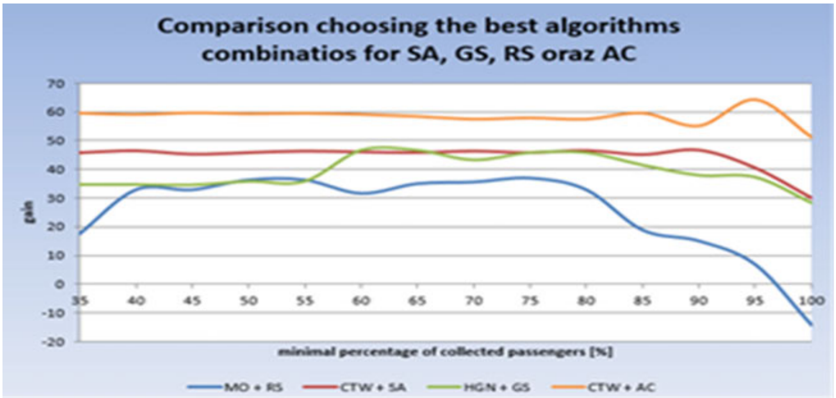


Fig. 6 Comparison of the combinations of the algorithms

combinations were recognized as the most interested. They were the following: MO + RS, CTW + SA, HGN + GS, CTW + AC.

The results of comparison between these four combinations are presented in Fig. 6. For 35 %: the combination MO + RS produced the gain of 18, HGN + GS produced the gain of 35, CTW + SA produced the gain of 46, and CTW + AC produced the gain of 60. The only better gain with the value of 64 was given for the last combination for L = 95.

6 Conclusion

The designed experimentation system with implemented algorithms which solve the considered modified travelling salesman problem reveals the difficulty in solving optimization problem and their complexity. On the other hand, the system allows noticing a few interesting regularities.

The most efficient of the implemented algorithms was Ant Colony (AC). In almost every case it gave the best result, when overall quality was based on the reached gain. The SA can be treated as the second efficient algorithm, especially when input is formed by greedy algorithm. The results for SA are better in relation to others when a bigger number of bus stops are contained into the considered route (caused by the greater requirements for minimal percentage of the served passengers). Conversely is in case of RS. This algorithm reaches good results for relatively small number of bus stops and relatively big number of random solutions. Sometimes, it is even better than TS, however, for instances with bigger number of bus stops, the results significantly differs from the desired values.

Another interesting observation, which indicates on an important regularity, is a significant fall of the gain for L greater than 95 %. The reason for this phenomenon is an increased number of the chosen bus stops, which is related to overcoming longer route by buses. Therefore, both economic and time aspect application of bus stops choice algorithms is fully justified. When applying Greedy Algorithm at the input of the best route determining algorithms (what was done in our implementations), we may observe a high improvement of results. It concerns especially TS algorithm.

Interesting results are also given by bus stops finding algorithms. For MO, the usually received gain is minimal and the difference is bigger when more bus stops can be ignored. In case of CTW, even when not many stops are skipped, a gain is maximal. For HGN, the obtained gain is much lower than that given by CTW, even in the case when many bus stops were ignored. If it is important to maximize the gain only, the best solution is to use CTW; however, if the important requirement is time limit, it seems to be better to use HGN.

The experimentation system was tested as a good tool for the practical users. The system (simulator) is also serving as a tool for lab training and projects preparing for the MSc students in Wroclaw University of Technology.

Acknowledgment This work was supported by the statutory founds of the Dept. of Systems and Computer Networks, Wroclaw University of Technology, Poland.

References

1. Grossman, D.A., Frieder, O.: Algorithms and Heuristics. Springer, Berlin (2004)
2. Cormen, T.H., Leiserson, C.E., Rivert, R.L., Stein, C.: Introduction to Algorithms. MIOT (2009)
3. Zuhori, S.T: Traveling Salesman Problem. LAP (2012)
4. Bertsimas, D., Nohadani, O.: Robust optimization with simulated annealing. J. Glob. Optim. **48**(2), 323–334 (2010)
5. Dorigo, M., Stützle, T.: Ant Colony Optimization. MIT Press (2004)
6. Zaki, S.A., Mousa, A.A., Geneedi, H.M., Elmekawy, A.Y.: Efficient multiobjective genetic algorithm for solving transportation and transshipment problems. Appl. Math. **3**, 92–99 (2012)
7. Kakol, A., Pozniak-Koszalka, I., Koszalka, L,. Kasprzak, A., Burnham, K.J.: An Experimentation System for Testing Bee Behavior Based Algorithm to Solving a Transportation Problem. LNCS. vol. 6592, pp. 11–20. Springer, Berlin (2011)
8. Golonka, K., Koszalka, L., Pozniak-Koszalka, I., Kasprzak, A.: An experimentation system for bus route planning and testing metaheuristic algorithms. In: Proceedings to the Sixth International Conference on Systems. ICONS (2011)
9. Hyndman, R.J.: Forecasting with Exponential Smoothing. Springer, Berlin (2008)
10. Kubacki, J., Koszalka, L., Pozniak-Koszalka, I., Kasprzak, A.: Comparison of heuristic algorithms to solving mesh network path finding problem. In: Proceedings of Fourth Conference on Frontier of Computer Science and Technology, Shanghai (2009)

Carbon Nanotube Modified Asphalt Binders for Sustainable Roadways

Rafiqul A. Tarefder and Arif Zaman

Abstract Asphalt binders are usually modified using several modifiers such as polymers, acid, lime, anti-stripping chemicals, and carbon fibers in order to improve binder properties such as durability, adhesion, cohesion, and performance in service. Among the properties, adhesion is key to understanding moisture damage behavior of asphalt binder. Moisture damage in asphalt is yet an unsolved problem in pavement engineering. In this study, microscopic images of adhesion force in carbon nanotube modified asphalt binders is determined using an Atomic Force Microscope (AFM). Moisture damage predictions are made based the adhesion forces affected by the amount of single-wall and multi-wall carbon nanotubes. It is hoped that this study will be useful for using carbon nanotube and nanoscale testing in creating new and durable pavement materials for sustainable infrastructures.

Keywords Carbon nanotube · Polymer modified asphalt · Atomic force microscopy · Moisture damage

1 Introduction

About 90 % of paved roads in US are asphalt pavement. Asphalt can be thought to behave like a glue, which keeps rocks/aggregates together. However, under the action of roadway traffic and moisture, such glue is not strong enough to hold the aggregates.

R.A. Tarefder (✉)
Department of Civil Engineering, University of New Mexico,
Albuquerque, NM 87113, USA
e-mail: tarefder@unm.edu

A. Zaman
Department of Civil Engineering, University of Bahrain, Isa Town, Bahrain
e-mail: st100132@gmail.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_52

Often, styrene-butadiene (SB) and Styrene-butadiene-styrene (SBS) polymers are used at 3–5 % to strengthen the gluing capacity of asphalt [1]. In this study, in addition to polymer, a small percentage of Carbon nanotubes (CNTs), for their well-known high stiffness properties, are used and the glue or adhesion properties of the resulting asphalt binders are evaluated using an Atomic Force Microscopy (AFM) technique.

2 Sample Preparation and Testing

At first, polymer was mixed with base binders. Base asphalt, at ambient temperature, was cut out of a five gallon container with heated trowels and placed into a clean one gallon container, weighed and labeled. The cans of asphalt were heated to 192 °C on a hotplate. While on the hotplate the polymer SB or SBS was added to make 4 and 5 % (by weight) mixes for a total of six samples. Polymer was mixed into the asphalt and stirred with an electric mixer until it was melted and incorporated into the asphalt. After ten minutes of mixing at 192 °C the asphalt was then cooled, labeled and stored. The process included the reheating of the one gallon cans of binder asphalt and once the asphalt reached a workable viscosity, it was poured into a glass vial and labeled. This process was repeated to make samples to mix with the single-walled carbon nanotubes (SWNT) and multi-walled carbon nanotubes (MWNT). Three different percentages of carbon nanotubes, 0.5, 1 and 1.5 % were mixed with asphalt binders. A small amount of sample was poured on a 25 mm × 75 mm and 1 mm thick glass slide. Glass slides had cellophane tape applied so that only a centered 3–6 cm^2 was exposed on one side of the slide. A clean glass stir rod was used to transfer the nanotube/asphalt mixture from the glass vial to the exposed square of the glass slide. The edge of another clean glass slide was used to scrape and smooth the drop of nanotube/asphalt mixture over the exposed square. The asphalt slide was then left to cool and dry for 10 min. Once drying was complete, the cellophane tape was removed. The slide was labeled, and ready for AFM testing. Slides were conditioned wet and for aging [2]. In an AFM test, the surface of the asphalt sample on glass slide was probed with a sharp tip located at the free end of a cantilever, and cantilever deflection is measured using laser beam. By multiplying the deflection by the cantilever spring constant, the attractive or repulsive force acting on the cantilever tip is measured as a function of the distance between the tip and the surface. The magnitude of the adhesion force, which corresponds to the jump out from the potential minimum, is measured precisely in these experiments [2, 3]. A total of five tips namely, Si_3N_4, NH_3, OH, CH_3 and COOH, representing chemistry of an asphalt molecule, were used to measure the AFM forces.

3 Results and Discussions

3.1 Base Binder Modified with SWNT

Figure 1 show the adhesion forces of dry, aged and wet base binders modified with 0.5, 1.0 and 1.5 % single wall nanotubes (SWNT). The adhesion force of COOH is almost always highest among all the dry, aged and wet samples. This indicates the less availability of COOH functional group on sample surface. When the chemicals in tip and the sample are similar, the force required to separate the tip and the surface is small [3]. The availability of the chemical functionals on sample surface can be summarized as OH > Si_3N_4 > CH_3 > NH_3 > COOH. The adhesion forces are increasing fairly in aged and wet samples as compared to dry samples. When applying OH tips on dry, wet and aged samples we can see that 100 nN adhesion force is acting in all three percentages of SWNT samples. That means the OH type functional in asphalt does not change the adhesion force due to SWNT or environmental condition variation. The adhesion force using CH_3 functional is unchanged in dry and aged sample but is higher in wet sample. So water conditioning and temperature conditioning (i.e., T-283 method) has some effect on the samples when considering CH_3 functional. NH_3 has more variation in adhesion

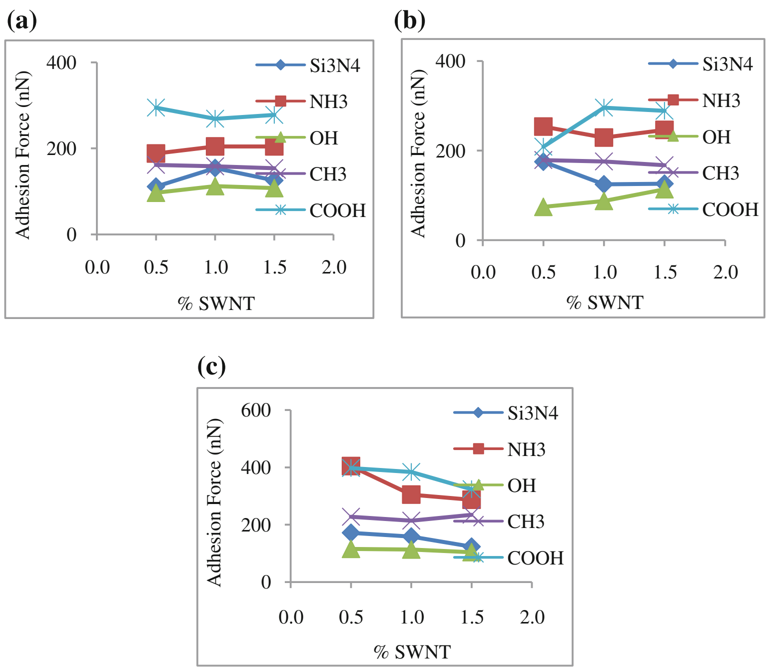


Fig. 1 SWNT modified base binders forces. **a** Dry. **b** Aged. **c** Wet

forces in aged and wet samples as compared to dry samples. The adhesion force in aged samples is almost 1.5 times in aged sample and almost 2 times in wet samples as compared to the dry samples using the NH_3 functional tips. Using the COOH tips the wet samples show the highest adhesion force as compared to dry and aged samples.

3.2 SB Polymer (4 %) Modified Binder with SWNT

Figure 2 show the adhesion forces of dry, aged and wet SB (4 % SB polymer) binders modified with 0.5, 1.0 and 1.5 % single wall nanotubes (SWNT). The OH modified tips show the lowest adhesion forces in all dry, wet and aged samples as compared to other tips. The SWNT modified SB 4 binder's adhesion forces are not affected much by the T-283 conditioning (wet) but affected by aging. Not much variation in adhesion force can be notices with the CH_3 functional group. But the adhesion forces in wet samples are higher than the dry and aged samples with the NH_3 tips. Similarly the adhesion forces of wet samples are higher than the dry and aged samples with the COOH tips.

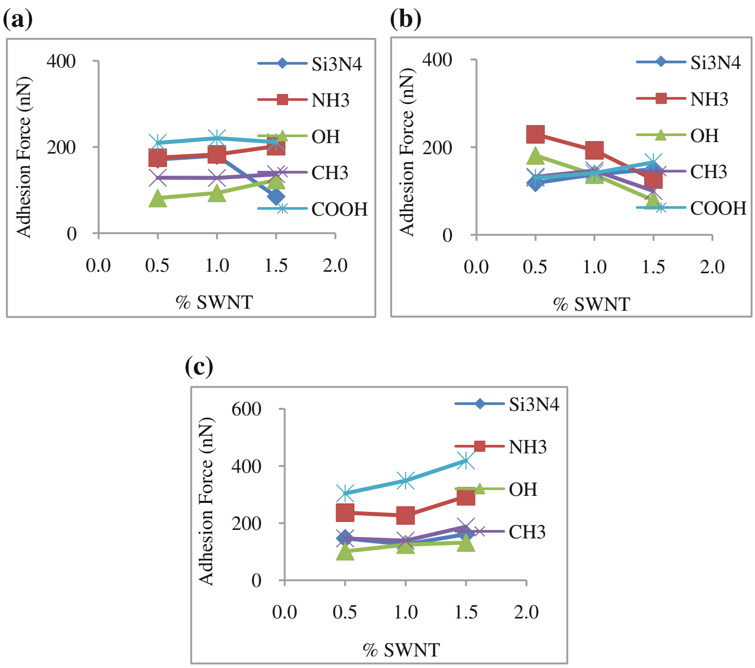


Fig. 2 SWNT modified SB 4 % binders forces. **a** Dry. **b** Aged. **c** Wet

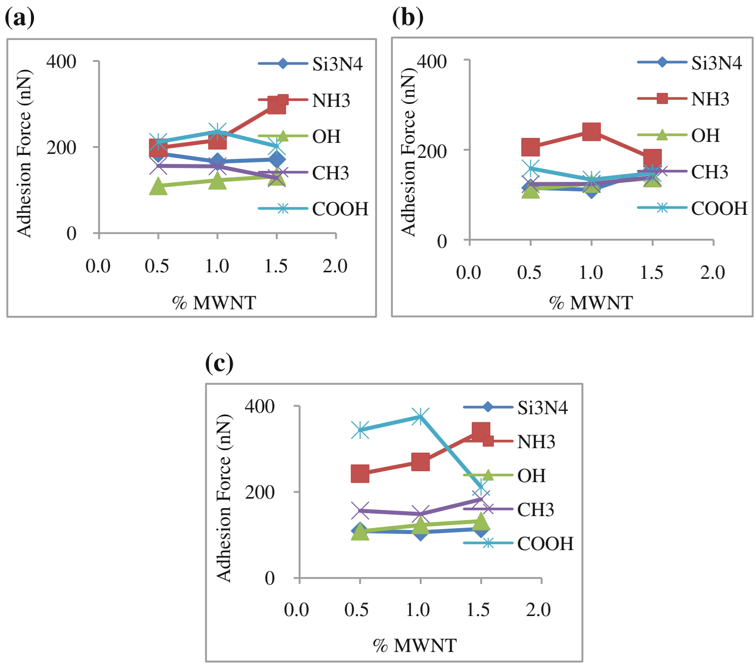


Fig. 3 MWNT modified SB 4 % binders forces. **a** Dry. **b** Aged. **c** Wet

3.3 SB Polymer (4 %) Modified Binder with MWNT

Figure 3 show the adhesion forces of dry, aged and wet SB 4 binders modified with 0.5, 1.0 and 1.5 % multi wall nanotubes (MWNT). The adhesion force due to OH functional is very much similar in all the dry, wet and aged samples. The adhesion force in aged samples is smaller than the dry and wet samples with CH_3 functional tips. With the NH_3 tips the wet samples adhesion force is the maximum as compared to dry and aged samples. The similar trend is also true for COOH tips.

3.4 SBS Polymer (4 %) Modified Binder with SWNT

Figure 4 show the adhesion forces of dry, aged and wet SBS 4 binders modified with 0.5, 1.0 and 1.5 % single wall nanotubes (SWNT). The adhesion forces are the maximum in wet samples as compared to dry and aged samples with the OH tips. The similar trend is also found with the CH_3 functional tips. The adhesion forces increased drastically in aged and wet samples as compared to dry samples with the NH_3 tips. It indicated the failure of SBS 4 and CNT modification of in terms of

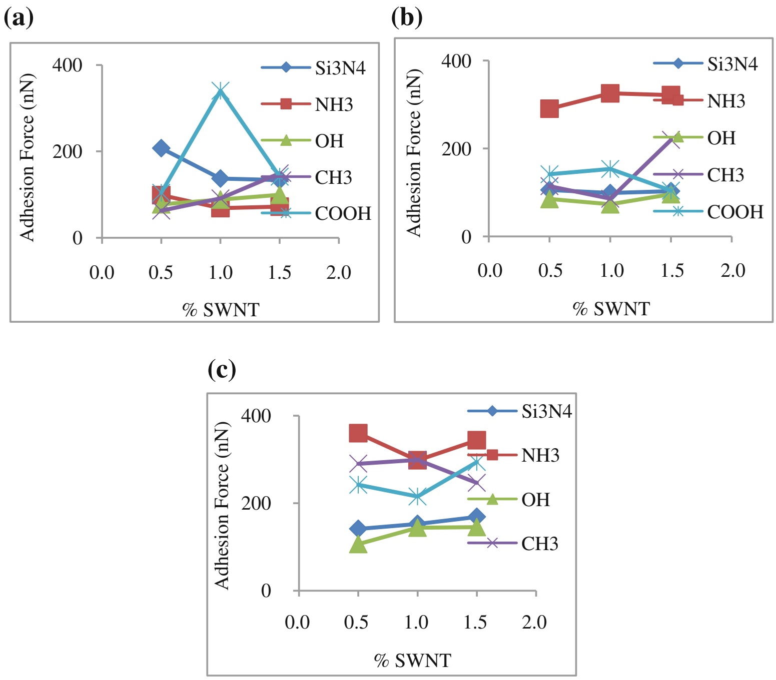


Fig. 4 SWNT modified SBS 4 % binders forces. **a** Dry. **b** Aged. **c** Wet

strengths when the NH_3 functional is concerned. The COOH related adhesion forces are decreased in both aged and wet samples.

3.5 *SBS Polymer (4 %) Modified Binder with MWNT*

Figure 5 show the adhesion forces of dry, aged and wet SBS 4 binders modified with 0.5, 1.0 and 1.5 % multi wall nanotubes (MWNT). The adhesion force in wet sample is little higher as compared to both dry and aged samples when dealing with the OH functional. The CH_3 related adhesion forces in wet samples are considerable higher in wet samples as compared to dry and aged samples. The adhesion forces of aged and wet samples are higher than that of dry samples with the NH_3 functional tips. The COOH related adhesion force is noticeable higher as compared to dry and aged samples.

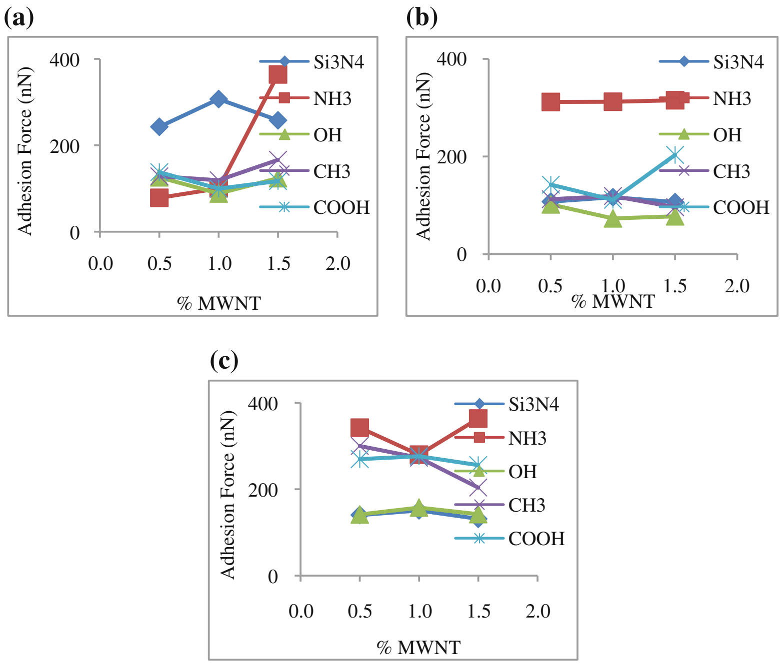


Fig. 5 MWNT modified SBS 4 % binders forces. **a** Dry. **b** Aged. **c** Wet

3.6 SB Polymer (5 %) Modified Binder with SWNT

Figure 6 show the adhesion forces of dry, aged and wet SB 5 binders modified with 0.5, 1.0 and 1.5 % single wall nanotubes (SWNT). The adhesion force due to OH functional is not much different in all dry and wet samples. Similar trend is also found for CH_3 tips. The adhesion forces in wet samples are higher than the adhesion forces in dry and aged samples with the NH_3 functional tips. The aged samples show lower amount of adhesion forces as compared to dry and wet samples with the COOH functional tips.

3.7 SB Polymer (5 %) Modified Binder with MWNT

Figure 7 show the adhesion forces of dry, aged and wet SB 5 binders modified with 0.5, 1.0 and 1.5 % multi wall nanotubes (MWNT). The aged binders show little less amount of adhesion force as compared to dry and wet samples with the OH tips. But the wet samples adhesion forces are higher as compared to dry and aged samples with the CH_3 tips. The adhesion forces in all dry, wet and aged samples are

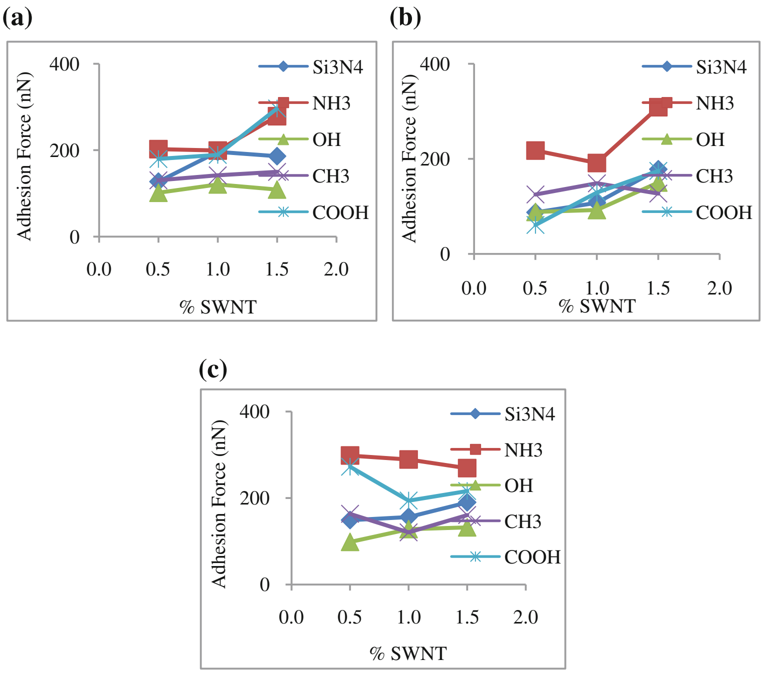


Fig. 6 SWNT modified SB 5 % binders forces. **a** Dry. **b** Aged. **c** Wet

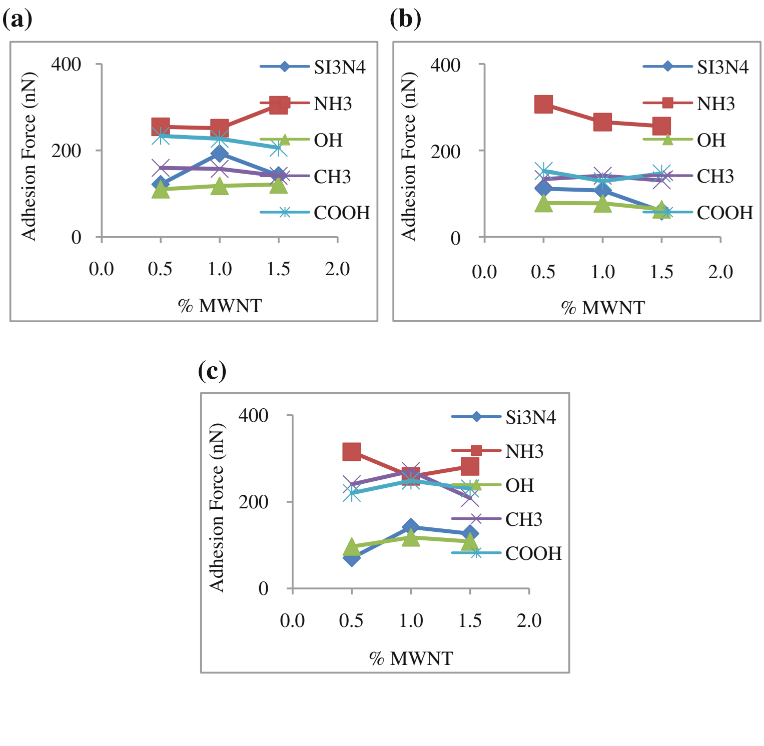


Fig. 7 MWNT modified SB 5 % binders forces. **a** Dry. **b** Aged. **c** Wet

kind of similar with the NH_3 tips. The aged sample's adhesion force is lower as compared to dry and wet samples with the COOH functional tips.

3.8 SBS Polymer (5 %) Modified Binder with SWNT

Figure 8 show the adhesion forces of dry, aged and wet SBS 5 binders modified with 0.5, 1.0 and 1.5 % single wall nanotubes (SWNT). The variation in adhesion forces in all dry, wet and aged samples is minute with the OH tips. The adhesion forces in wet samples are higher as compared to dry and aged samples with the CH_3 tips. The adhesion force in aged sample is smaller than that of dry and wet samples with NH_3 tips. The COOH functional tip shows higher value in 1.0 % SWNT aged sample as compared to dry and wet samples.

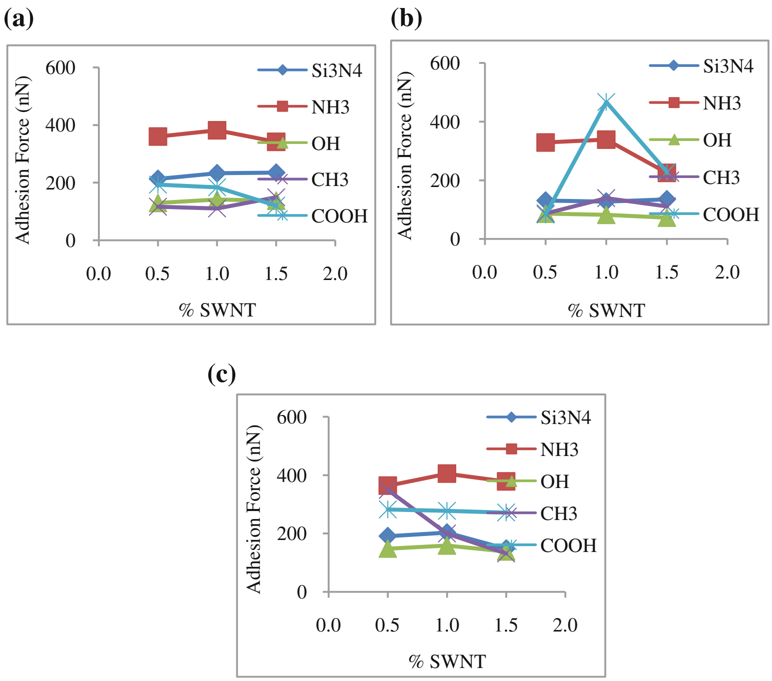


Fig. 8 SWNT modified SBS 5 % binders forces. **a** Dry. **b** Aged. **c** Wet

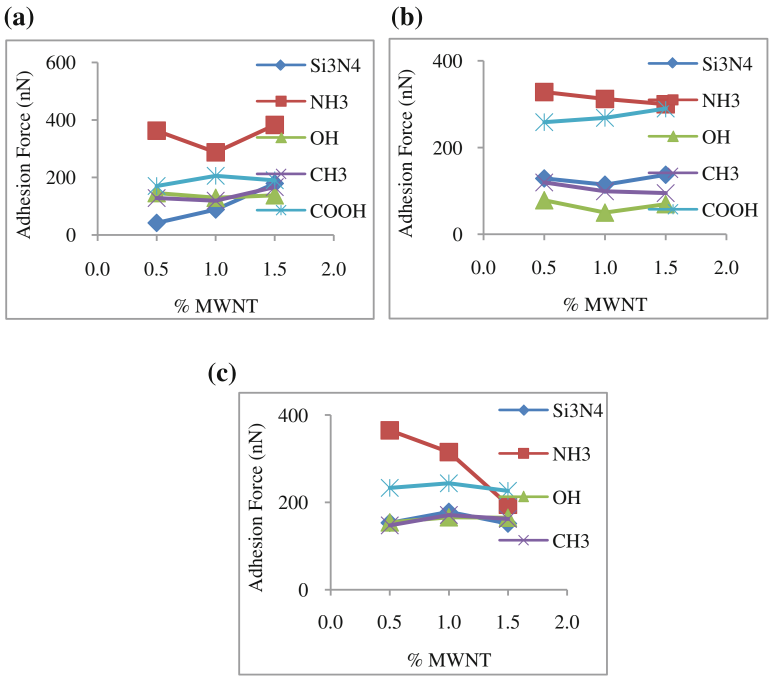


Fig. 9 MWNT modified SBS 5 % binders forces. **a** Dry. **b** Aged. **c** Wet

3.9 SBS Polymer (5 %) Modified Binder with MWNT

Figure 9 show the adhesion forces of dry, aged and wet SBS 5 binders modified with 0.5, 1.0 and 1.5 % multi wall nanotubes (MWNT). The adhesion force in aged sample is lower than that of dry and wet samples with the OH functional tips. Wet samples show higher adhesion forces as compared to dry and aged samples with the CH_3 functional tips. The aged and wet samples show lower amount of adhesion forces as compared to dry samples with the NH_3 tips. The adhesion forces in aged samples are lower than that of dry and wet samples with the COOH tips.

4 Conclusion

For the first time, a fundamental study of measuring adhesion forces in modified asphalt system is presented here. Wet sample has shown higher adhesion (AFM force) value than the dry sample, which means water weakens or softens the sample. In all cases, SWNT and MWNT modifications have decreased the adhesion value when compared to the adhesion value of base binder. There is no significant difference in adhesion between SWNT and MWNT modifications.

References

1. Tarefder, R.A., Yousefi, S.S.: Laboratory evaluation of moisture damage in Asphalt. Can. J. Civ. Eng. **39**(1), 104–115 (2012)
2. Tarefder, R.A., Arifuzzaman, M.: A study of moisture damage in plastomeric polymer modified asphalt binder using functionalized AFM tips. J. Syst. Cybern. Inf. **9**(5), 1–12 (2011)
3. Noy, A., Vezenov, D., Lieber, C.: Chemical Force Microscopy Nanoscale Probing of Fundamental Chemical Interactions. Handbook of Molecular Force Spectroscopy. Springer, US (2007)

Sensor Oriented Approach to Prevent Hyperthermia for Children in Car

Maysoon Abulkhair, Lujain Mulla, Amani Aldahiri, Hanin Alkhatabi, Hala Alonezi and Somia Razzaq

Abstract Leaving children unattended in a car for few moments; especially in hot atmosphere, can cause a catastrophic tragedy to occur. This paper presents HACC system to help in preventing tragic child death caused by hyperthermia using detection and control system inside car. HACC system has been accomplished by developing a phone application and a surveillance system connected together to monitor the temperature and the presence of a child inside the car. The system starts to measure temperature inside car via temperature sensor. At the same time, it checks constantly the presence of child inside car via motion sensor. When system detects the presence of a child and the temperature inside car reaches unsafe limit, it will alert the caregiver via smartphone application and allows him/her at the same time to take an action and open windows remotely. If there is no response from the caregiver, system itself reacts and windows will be opened automatically.

Keywords Hyperthermia · HACC · Surveillance system · Mobile application · Sensors

1 Introduction

Nowadays, cars are the most popular transportation that brings convenience and ease for people to perform their daily activities. However, many families have constant fear and suspicions about safety of their children inside cars. Many solutions were made to limit these dangers by preventing children from being injured or getting hurt. For example, using of best cars' safety equipment or cars safety application.

M. Abulkhair (✉) · L. Mulla · A. Aldahiri · H. Alkhatabi · H. Alonezi · S. Razzaq
Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, B.P. 42808, 21551- Girl Section, Jeddah, Saudi Arabia
e-mail: mabualkhair@kau.edu.sa

S. Razzaq
e-mail: srabdulrazzaq@kau.edu.sa

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_53

Table 1 Difference between temperature inside and outside vehicle

Temperature outside car (°F)	Temperature inside car (°F)
75	100
85	120
100	140

Parents around the world must do their intense researches on keeping their children safe inside car and shop for best children car's seats. But do they know, these seats are not enough for children safety? Do they know, there are many other dangers in and around a vehicle that could seriously harm or even kill their children?

Hyperthermia is one of the most common dangers that can affect children lives. It is considered the third cause of children deaths inside cars around the world. It occurs when body is not able to cool itself quickly enough and body temperature rises to dangerous levels. Young children are particularly at risk as their body heats up 3 to 5 times faster than an adult.

"Since 1998, until 2012 (14 years) more than 530 children across the United States have died from hyperthermia when unattended in a vehicle [1]." Imagine the number of children who die by the same phenomenon around the world. Preventing this dramatic event affecting children lives will be accomplished with a surveillance and detection system implemented inside car to monitor children and protect them from hyperthermia. Every year, many avoidable deaths occur due to vehicular hyperthermia and rising numbers are worrying child safety. As children are trapped or placed out of sight in back seat of vehicles, the incidence of deaths due to this phenomenon has increased.

For children less than 4 years, heatstroke occurs when their body temperature exceeds 104 °F. In addition, when it exceeds 107 °F is considered lethal as the body's cells are damaged and internal organs shut down. On average, studies have shown that in 10 min the temperature in a vehicle raises about 20 °F. Within an hour, the temperature jumps around 50 °F. So even if the temperature outside the vehicle does not feel warm enough to heat it; thus it is better not to leave a child unattended in the vehicle because it is not safe [2]. Table 1 shows the temperature difference between the vehicle inside a closed vehicle compared to the outside temperature.

2 HACC System Architecture

Hyperthermia Alarm for Children in Cars (HACC) is an interactive safety system provided to public with aim to monitor safety of children who are left behind inside closed, parked vehicle. It helps in preventing tragic child deaths in cars caused by hyperthermia by implementing detection and control system in car. HACC system consists of two parts; first car surveillance system which checks temperature

Fig. 1 HACC architecture

constantly and detects presence of a child inside a car. The second part is mobile application that enables user (parents or caregiver) to monitor child inside the car and take appropriate reaction to help child through smartphone application. Moreover, it alerts application's user when temperature inside the car reaches unsafe limit for a child to be inside it. When there is no response taken, system takes an action and opens car's windows automatically. At the same time, it starts an alarm to grab people attention around the car to get the required help for child. Obviously, connecting surveillance system to caregiver's smartphone application remotely will make system more acceptable and portable to users. Thus, safety inside the car will increase to achieve a safer environment for children. Figure 1 illustrates the general architecture for HACC.

3 Related Work

There are several models for monitoring car temperature to detect children from hyperthermia. For example, Car Seat Monitor (Cars-N-Kids) idea is to sense child's presence in child car seat inside car. This was done using a small sensor pad that is placed under the cushion of child car seat. The device is synchronized with the mobile phone application that is installed in the parent's or caregiver's smartphone. When a child is placed in child car seat, the associated weight activates the device. If the car comes to a stop for more than four seconds, the mobile phone vibrates and/or sends a text message to the smartphone application. The purpose of this device is to remind parent that child is present before he or she gets out of the car [3].

4 Evaluation

After studying the previous application, it showed that the device which is implemented in car to detect the child's presence only if he/she in child car seat otherwise it will not detect him/her. This device is synchronized with a smartphone application, which can be downloaded and installed into parent's smartphone. Obviously, since we need to alert parent or caregiver remotely, the mobile phone application is suitable solution for the project because it is the best way to ensure that it will be always with application's user. Also, it is easy to interact with and more practical. Moreover, in previous proposed application the device will sense child's presence only when child is in child car seat using weight sensor. While the goal of HACC system, is to check presence of child anywhere inside car using motion sensor. Unlike previous application, HACC system is designed to work in both situations either child is left unattended intentionally or inadvertently inside car. Instead of using text message notification, HACC uses an alarm sound notification, emitted from phone application to alert parent or caregiver about unattended child. Finally, HACC is designed as an interactive system to save children automatically from hyperthermia danger even if no action is taken from parent or caregiver. Table 1 illustrates a comparison between Cars-N-Kids Car Seat Monitor Application and Hyperthermia Alarm for Children in Car System Table 2.

5 Usability Testing

Usability testing is a technique used in user-centered interaction designed to evaluate a system or product by testing it on multiple different users. We conducted a usability testing for HACC system; the testing scenario was papered by the team members and test was taken by multiple users. HACC system usability testing team consists of four members and each one has assigned a role. The team members are: Lujain, Amani, Hanin and Hala.

Table 2 Comparison between HACC and Cars-N-Kids system

Criteria	Cars-N-Kids car seat monitor	HACC
Functionality	It detects child's weight on child car seat only	It detects any child's motion inside whole car
Conditions	It works when child left unattended inadvertently inside car	It works when child left unattended intentionally or inadvertently inside car
Compatibility	It works with child car seats only; it is not compatible with cars	It is compatible with all types and sizes of cars
Distance limit	It works around the car only	It works remotely
Response action	It does not allow parents to take any action to save child	It allows parents to take action remotely to save child. Even, it reacts and rescues child, if no response has been taken by parent
Language	It is available in English only	It is available in other languages such as Arabic

Lujain was assigned to be a facilitator who provides an overview to the participants about the application and defines usability and purpose of usability testing for HACC system. On other hand, Hanin was assigned to be a test observer who is a note taker and records participant's actions, behaviors and comments. Also she has observed participants during the test and identifies problems and procedural errors. In addition, Amani and Hala were assigned to respond to participant's questions and requests for assistance. They respond to their questions and suggestions for improvement after they finish the testing.

In this step we provide the volunteers of the usability testing participant with the purpose of HACC system and to know how much is the HACC smartphone application will be easy to use. Ideally testers should have some experience using technology but they do not need to be expert users. We need to make sure that application is friendly and easy to use.

The process for the usability testing starts with downloading the HACC application in the participants mobile (must be android) followed by performing several basic tasks on the application.

6 Usability Testing Analysis

We have tested the HACC application with five testers. After test has been completed, we observed that:

1. 80 % of the testers completed task easily while 20 % of them could not complete the assigned tasks such as responding to the alert notification by opening the widows and turning the alert off. This percentage shows that the application is very usable.

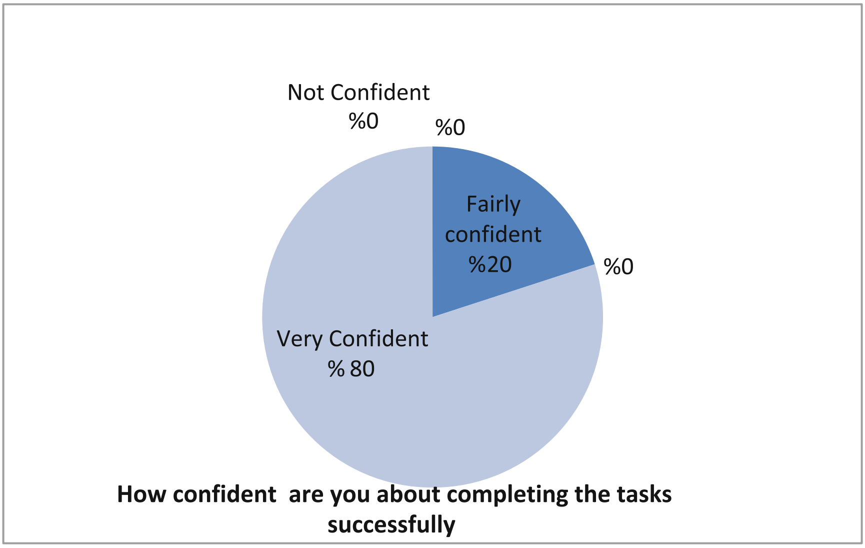


Fig. 2 The confident about completing the usability testing successfully

2. 80 % of participants say that they are very confident about their ability to complete the task, while 20 % were fairly confident. Figure 2 shows the percentage of confident in completing the usability testing.

7 Conclusion

We presented a system for detecting the unattended child who left intentionally or inadvertently inside car to prevent him/her from hyperthermia. Our system is developed based on multi-sensor for both motion and temperature sensors and smartphone application. In doing so, this system advanced the state of the art by improving the way for preventing children from hyperthermia and enhancing the communication features for more children protection. Additionally, the system itself reacts and open car's windows automatically when there is no response from child's caregiver. Also, the application provides an automatically immediate connection to the emergency center to get the necessary help for child.

References

1. Heat Stroke. http://www.kidsandcars.org/heatstroke.html. 10 Apr 2013
2. Helpful Tips to Keep Your Child Save. http://www.ocfs.state.ny.us/main/publications/Pub5036text.asp
3. Cars-N-Kids. http://carseatmonitor.com/

Relationships Between Years of Licensure and Driving Style Measured with a Short Simulator-Based Test (N = 650)

Joost de Winter and Jorrit Kuipers

Abstract Young and inexperienced drivers are over-involved in traffic violations and car crashes. There is a paucity of research on the use of driving simulators for assessing driving style. This study investigated the relationships between years of licensure and driving style measured with a short simulator-based test. At a motor show, 650 licensed drivers completed a 6.5-min driving style test and responded to a questionnaire about their on-road driving experience. The results showed that participants who had their driving license for a longer period adopted a less risky driving style and drove with slower speeds in the simulator. Furthermore, females and experienced drivers reported more simulator sickness than males and inexperienced drivers, respectively. The present results may be useful in the development of simulator-based driving tests.

Keywords Human factors · Driving simulator · Driving assessment

1 Introduction

On a yearly basis, road traffic crashes claim the lives of 1.25 million people and between 20 and 50 million people suffer non-fatal injuries [1]. Engineering innovations such as electronic stability control (ESC) and advanced emergency braking (AEB) have led to an impressive improvement in road safety [2]. However, the

J. de Winter (✉) · J. Kuipers
Department of BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands
e-mail: j.c.f.dewinter@tudelft.nl

J. Kuipers
e-mail: j.kuipers@greendino.nl

J. Kuipers
Green Dino BV, Bronland 12-G, 6708 WH Wageningen, The Netherlands

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_54

human remains the weakest link in driver-vehicle interaction, with human error being the primary cause of 90 % of road traffic crashes [3, 4]. Young and inexperienced drivers are overrepresented in crash statistics. An important reason is that young drivers commit a large amount of traffic violations and more often engage in risky driving behaviors than older drivers [5–7]. A driver's driving style, including violations, bad habits, and predisposition to risk, has a large effect on road safety [8].

On-road driving tests are currently the gold standard for deciding whether someone is competent to drive, both for young people obtaining their driver's license and for older people or patients whose driving needs reassessment. However, on-road driving tests suffer from issues of inter-rater reliability of driving examiners [9]. Moreover, even if examiners were always consistent and in agreement with all other examiners, there would still be limitations to the validity and fairness of on-road driving tests, due to random traffic and weather conditions as well as local differences in road infrastructure (e.g., [10]).

Psychometric tests, such as visual and psychomotor tests, are an alternative to on-road testing [8, 11–14]. A limitation of psychometric tests is that they are remote from what occurs on the road (e.g., [15]). One important aspect hard to capture in psychometric tests is the self-paced nature of driving. That is, drivers can adapt their driving style to cope with high task demands or to compensate for their own limitations. For example, although older persons are known to suffer from degraded processing speed and spatial skills [16, 17], they tend to drive more carefully and with wider safety margins than young drivers do [18, 19].

Driving simulators provide visual and auditory sensations that mimic real car driving while providing high controllability. With a simulator, hazardous driving scenarios can be offered without the physical risk of crashing [20]. Simulators are therefore reported to be "a bridge between laboratory-based artificial tasks and real-world driving that enables context selective testing" ([21, p. 24]). Another important quality of simulators is that they can measure the state of the car (e.g., speed, position) objectively. Although there is a growing body of evidence on the validity of driving simulators for predicting various types of on-road performance [5, 22–28], there is still little empirical knowledge on the suitability of driving simulators for testing the driving population. The aim of this research was to investigate the validity of a simulator-based driving style test, by measuring the extent to which the driving style in the simulator relates to drivers' years of licensure. An adjacent purpose of this research was to investigate possible bottlenecks in simulator-based testing, in particular regarding simulator sickness, which is an important topic in driving simulator development (e.g., [29–31]).

2 Methods

2.1 Simulators

Two identical portable driving simulators (Model Drive Master, Green Dino BV, 2009) were used to administer the simulator test. Each had a throttle, brake, and clutch pedal, steering wheel, horn, gear lever with five forward gears, and ignition key. Five LCD screens (resolution 1024 × 768 pixels) presented the virtual environment. A sixth screen presented the dashboard with two dials showing driving speed and engine rpm, respectively. A rear-view mirror and two side mirrors were integrated in the simulated image. The simulated field of view was 149° horizontally and 23° vertically.

The steering wheel provided force feedback derived from a combination of static and self-aligning torque through a torque engine connected to the steering shaft. Deceleration cues were simulated by a vibration unit on the steering shaft. Engine, wind, and tire sounds were audible on a speaker positioned in front of the driver. The simulated car was a mid-class vehicle with a mass of 1265 kg and a top speed of 180 km/h.

2.2 Driving Style Test

Visitors could voluntarily complete the driving style test free of charge at a demonstration stand of a driving simulator manufacturer at a motor show held in April 2009. Attended by 220,000 people, the motor show featured car manufacturers' stands with various demonstrations of innovative car products and driving simulators.

Only people with a driver's license were eligible to participate in the driving style test. The test consisted of three simulated drives and lasted about 10 min in total, including the time needed to enter and exit the simulator. The durations of the three drives were 75, 150, and 160 s, respectively.

Prior to the first drive the simulator presented a short welcome text, informing participants that the driving style test consisted of three drives after which they would be presented with their scores. The text also stated that the first drive allowed participants to get used to the simulator and that if they felt discomfort, they had to report that to the supervisor who would then alter the field of view. Following both the first and second drive, there was a short interval and the text on the screen reminded participants to report any discomfort. The supervisors could adjust the field of view by shutting down the outer two or outer four LCD screens. The simulator software provided no feedback or guidance during driving, except for spoken route instructions (recorded male voice) and a corresponding display of arrows at the bottom of the central screen.

The rationale behind the test was to expose participants to challenging driving scenarios within a short amount of time. Therefore, the test took place in a city environment with a high frequency of encounters with other road users. Furthermore, the driving style test was self-paced. That is, the test did not include scenarios where the participants would approach with fixed speed. Instead, the drivers had to choose their own pace, could violate speed limits, and had to make right-of-way decisions at intersections. Accordingly, the driving style test assessed a variety of behaviors, in particular driving speed and adherence to traffic rules.

Drive 1 took place in a city environment containing intersections without authorized priority. Curved road segments of about 200 m long separated the intersections. At each intersection, the voice instructed the participant to go straight ahead. The speed limit was 50 km/h. In Drives 2 and 3, the participant drove along a figure-of-eight route in a city environment with segregated cycle lanes on both sides of the road. Pre-programmed scenarios such as crossing cyclists and pedestrians were triggered on time-to-arrival of the participant. Traffic lights were not triggered according to the position of the participant, but changed color according to a fixed time pattern. The participant drove the route described in Table 1. In Drive 2, the participant's starting location was near the beginning of Item 1 (Table 1). In Drive 3, the participant started near the end of Item 7. In all three drives, the participant started with zero speed and with the engine off. The majority

Table 1 The route that participants had to drive in Drive 1 and Drive 2

Item	Description of environment
1	Straight ahead (275 m, speed limit 50 km/h) in a built-up area
2	Cross an intersection without authorized priority (35 m). This means that the 'right-of-way' rule (give way to the right) holds
3	Straight ahead (295 m, speed limit 50 km/h)
4	Turn right at a traffic-light-controlled intersection. A cyclist (travelling in the same direction as the participant) rides straight ahead and therefore has right of way
5	Leave a built-up area and drive through a mild 90-degree curve to the right (499 m, speed limit 80 km/h)
6	Turn right at an intersection without authorized priority. Two parallel cyclists (travelling in the same direction as the participant) ride straight ahead and therefore have right of way
7	Enter a built-up area, driving straight ahead (290 m, speed limit 50 km/h)
8	Cross an intersection without authorized priority (35 m). A scooter approaches at high speed from the left and (unexpectedly) does not give way to the participant
9	Straight ahead (42.5 m, speed limit 50 km/h), entering a residential area with three speed humps (192.5 m, speed limit 30 km/h). People are standing on the footpaths and on the street. A child crosses the street, right in front of the participant. After leaving the residential area, the participant drives along a further 55 m (speed limit 50 km/h)
10	Turn left at an intersection without authorized priority. Two parallel cyclists approach, that is travelling in the opposite direction to the participant. They ride straight ahead and therefore have right of way

(continued)

Table 1 (continued)

Item	Description of environment
11	Straight ahead (290 m, speed limit 50 km/h), passing a stationary bus on the opposite side of the road. A pedestrian standing beside the bus intends to cross the road, but pulls back
12	Turn left at an intersection without authorized priority. An approaching scooter drives straight ahead, having right of way
13	Straight ahead (285 m, speed limit 50 km/h)
14	Turn left at a traffic-light-controlled intersection. An approaching scooter turns right, that is, left from the participant's perspective

of participants completed less than 50 % of the figure-of-eight route per drive, meaning that most participants did not encounter the same scenario more than once. A degree of route overlap occurred for the fastest 5 % of the drivers. To enhance realism during the three drives, a small number of additional cars drove around, controlled by artificial intelligence.

2.3 Collected Data

The following data were collected per participant:

Driver Errors While Driving. A total of 36 errors were rated automatically. The errors were categorized into three types. In Type 1 errors (e.g., collisions, driving off the road), the number of occurrences was counted. In Type 2 errors (e.g., ignoring right of way, ignoring traffic lights), the score was calculated by dividing the number of recorded occurrences to the maximum possible number of occurrences. For example, if the participant failed to provide right of way in one out of four occasions, then the score for this error was 0.25 on a scale from 0 (*good*) to 1 (*poor*). Type 3 errors were the result of a continuous assessment on a scale from 0 (good) to 1 (poor). Examples of Type 3 errors are driving too fast, driving too slow, driving off-center, and driving too close to a car in front. Here, scores were proportional to the severity of the deviation from the norm and the total period. For example, *driving too fast* was possible on straight road segments only. A score was determined for each sampling instant (approximately 10 Hz) that the participant drove on a straight road segment. This score was 0 when the participant drove slower than a lower limit (equaling the speed limit of that road segment plus 5 %), 1 when the participant drove faster than an upper limit (30 % above the speed limit of that road segment), and between 0 and 1 when the speed was between the lower and upper limit. The score for driving too fast was calculated by averaging the score across all sampling instants.

After completing the simulator test, participants were shown their error scores on a computer outside the simulator. These scores were aggregated into various categories (safety, driving skill, risk avoidance, eco-friendly driving, vehicle control, speeding, traffic rules) and shown on a scale from 0 (*poor*) to 10 (*good*). A corresponding textual explanation was shown as well.

Measures Based on Driving Performance in the Simulator. For each completed road segment per participant, the simulator logged a number of variables, including start and end times, mean speed, standard deviation of speed, throttle position, lateral lane position, and maximum brake position. Using these segment variables, the following measures were estimated per drive: (1) mean speed (km/h), (2) maximum speed (km/h), (3) standard deviation of speed (km/h); a low value means that the participant drove at a constant speed, whereas a high value means that the participant drove with high fluctuations in speed, (4) standard deviation of lateral position (m); A low value means that the participant drove the segments at a constant distance to the center of the lane, and a high value means that the participant showed large fluctuations in lateral position, (5) standard deviation of throttle position (0–1); a low value means that the participant held the gas pedal in a constant position, whereas a high value means that the participant showed large fluctuations in pedal position, such as regularly providing full throttle and then releasing, and (6) maximum brake position (0–1); a high value means that the participant pressed the brake deeply during the drive, for example in response to a critical event, whereas a low value means that the participant did not brake hard during the drive.

The mean speed, standard deviation of speed, standard deviation of lateral position, and standard deviation of throttle position were calculated using the completion time of the road segment as a weighting factor. Maximum speed and maximum brake position were calculated using the recorded maximum of the road segments per drive. These performance measures were not shown to the participants.

Questionnaire Results. After completing the simulator test and observing their error scores, participants could fill in a single-page electronic questionnaire on a computer with touch screen. The heading of the questionnaire stated that the collected data would be used by the Delft University of Technology and treated as confidential. The questionnaire consisted of the following items (translated from Dutch): (1) Gender (possible answers: male or female), (2) How many years do you have your driving license? (any number could be typed in), (3) How many kilometers do you drive per year? (any number could be typed in), (4) Have you ever caused damage? (possible answers: yes or no). The questionnaire also included the following statements: (5) The driving simulator is a good alternative to a driving lesson on the road (Alternative), (6) My driving style corresponds to the driving style provided by the test (Agreement), (7) The next time I get into a car I will pay more attention to my driving style (Attention), (8) I suffered from car sickness (Sickness). Participants were asked to indicate the extent to which they agreed with each of these statements, using a five-point Likert scale with anchors at 1 (*disagree*) and 5 (*agree*). The questionnaire was completed by pressing the 'Submit Questionnaire' button.

It was decided to omit Question 4 from the analyses, because it asked whether the participant had *ever* caused a crash. Drivers who have their license for a longer period (Question 2) reported a higher number of crashes, giving the false impression that these drivers were less safe than younger drivers were. A positive correlation between crashes and years of licensure is self-evident because the longer you have been a driver, the greater the chance of *ever* having caused an accident.

2.4 Statistical Analyses

Descriptive statistics were calculated, and a Pearson correlation matrix amongst the simulator measures and questionnaire responses was constructed. The number of years of licensure and annual mileage were converted to ranks.

3 Results

Eight hundred and twenty-six participants (747 men, 79 women) completed the simulator test and 650 of them (598 men, 52 women) completed all items of the questionnaire, a response rate of 79 % (80 % for men, 66 % for women). Due to a data storage error, no questionnaire data were available from the second day of the motor show. Because 46 of the 826 participants completed the test on the second day of the 12-day motor show, the true response rate was therefore estimated at 83 % (i.e., 100 % * 650/(826 − 46) = 83 %).

Because the eigenvalues of the correlation matrix among the 36 error scores indicated strong unidimensionality (the first three eigenvalues were 9.84, 2.52, and 1.96, respectively), the 36 z-transformed scores were reduced into one total risk score using the item-total correlations as weights. The total risk score therefore represents a weighted average of all recorded errors. High item-total correlations for speed-related errors confirmed that the total risk score is a measure of driving style rather than driving skill.

Table 2 shows the means, standard deviations, and correlations of the obtained data. Participants who had their license for a larger number of years had lower risk scores and lower speeds in the simulator. The correlation between maximum speed and years of licensure was −0.33 (variable 2 and 10), and the correlation between mean speed and years of licensure was −0.30 (variable 2 and 9) as Fig. 1 illustrates.

Participants with a lower total risk score as well as participants with more years of licensure were more likely to agree that their error scores corresponded with their actual driving style (higher Agreement score). The corresponding correlations were −0.20 (variable 5 and 8) and 0.12 (variable 2 and 5), respectively.

Women and experienced drivers (both in terms of years of licensure and annual km) reported higher ratings of simulator sickness than males and inexperienced drivers. The relationship between years of licensure and simulator sickness is illustrated in Fig. 2 (r = 0.31; variable 2 and 7). Sixty-five percent of the participants disagreed (i.e., score = 1) with the statement ‘I suffered from car sickness’,

Table 2 Means, standard deviations, and correlations of the dependent variables ($N = 650$)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Gender (1 = man, 2 = woman)	1.08	0.27													
2. Licensure (years)	12.6	12.7	−0.08												
3. Annual km	24,709	27,129	−0.22	0.46											
4. Alternative (1 = disagree, 5 = agree)	3.20	1.33	−0.10	0.11	0.06										
5. Agreement (1 = disagree, 5 = agree)	2.83	1.37	−0.03	0.12	0.05	0.21									
6. Attention (1 = disagree, 5 = agree)	2.64	1.33	−0.02	−0.12	−0.12	0.23	0.09								
7. Sickness (1 = disagree, 5 = agree)	1.90	1.36	0.09	0.31	0.10	−0.06	0.03	−0.07							
8. Total risk score	0	1	0.06	−0.25	−0.04	−0.09	−0.20	0.00	−0.11						
9. Mean speed (km/h)	30.89	7.17	0.01	−0.30	0.01	−0.15	−0.10	−0.01	−0.06	0.66					
10. Maximum speed (km/h)	64.23	17.12	0.04	−0.33	−0.08	−0.13	−0.19	0.01	−0.07	0.88	0.81				
11. *SD* speed (km/h)	11.35	3.19	0.06	−0.27	−0.09	−0.09	−0.19	0.00	−0.09	0.83	0.57	0.84			
12. *SD* lateral position (m)	0.19	0.11	0.08	−0.05	0.02	−0.04	−0.11	0.01	−0.08	0.57	0.28	0.46	0.40		
13. *SD* throttle (0 – 1)	0.14	0.06	0.04	−0.30	−0.11	−0.06	−0.20	0.05	−0.11	0.82	0.52	0.80	0.82	0.41	
14. Maximum brake (0 – 1)	0.41	0.17	−0.03	−0.17	0.00	−0.01	−0.07	0.00	−0.06	0.30	0.28	0.33	0.37	0.01	0.39

Note $p < 0.05$ for $|r| \geq 0.08$, $p < 0.001$ for $|r| \geq 0.13$. Variables 8–14 are averaged over the three drives. The widths of the bars linearly correspond to the correlation coefficient

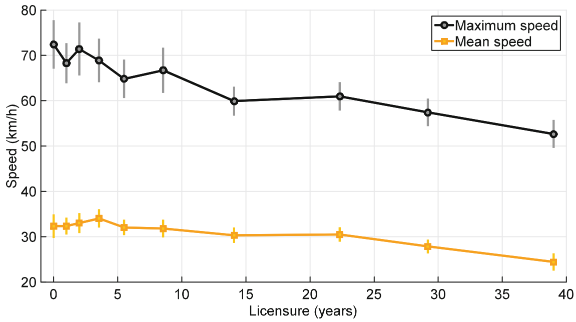


Fig. 1 Maximum and mean speed (averaged over the three drives) as a function of licensure. The participants ($N = 650$) were sorted on years of licensure in 10 groups of approximately equal size. The horizontal axis shows the mean years of licensure and the vertical axis shows the averaged speed per group. The vertical lines are 95 % confidence intervals

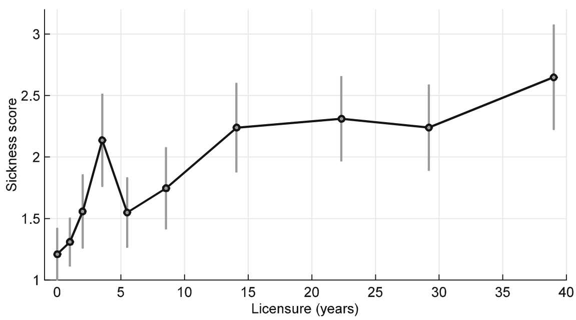


Fig. 2 Simulator sickness score (1 = disagree, 5 = agree) as a function of licensure. The participants ($N = 650$) were sorted on years of licensure in 10 groups of approximately equal size. The horizontal axis shows the mean years of licensure and the vertical axis shows the averaged sickness score per group. The vertical lines are 95 % confidence intervals

whereas 6 % agreed (i.e., score = 5). Of those who agreed, 18 % were female, while among those who disagreed, only 7 % were female. The simulator records revealed that there were 15 persons (13 men, 2 women) who did not complete the driving style test, for undocumented reasons. Three of them filled in the questionnaire, and all three disagreed with the statement ‘I suffered from car sickness’ (score = 1). This finding suggests that, for these participants, simulator sickness was not the cause of dropout.

4 Discussion

This study investigated how years of licensure and yearly mileage relate to driving style as measured with a short simulator-based test. The results showed that drivers who had their driving license for a longer period (thus generally also older drivers) drove more slowly and with a less risky driving style. This correlation is consistent with on-road driving data showing that young drivers are more likely to engage in risky behaviors than older drivers [7]. This study adds to a body of evidence showing that simulators are able to measure a person’s driving style in a valid manner. For example, it has been found that age [32], self-reported driving style [33], and self-reported violations [5] are predictive of driving speeds in a simulator.

The use of portable simulators at a motor show allowed us to obtain a large sample size (and see [34], who tested 624 people in a simulator using a similar approach). Although our sample size was large, the present sample is not representative of the driving population, because the participants were motor show visitors; males were highly overrepresented. The unusual sample may also explain why no statistically significant gender differences in driving behavior were observed (Table 2), whereas large gender differences have been observed among learner drivers during simulation-based driver training [35]. It requires further research to investigate whether the present results are generalizable to other driving populations.

With a drive time of 6.5 min per participant, the driving style test was of short duration. According to McGehee et al. [36], experienced drivers need at least 6 min to get used to a simulator and learn how to steer the virtual car stably. A longer test is not necessarily desirable, however. Allen et al. [22] showed that the participants’ (N = 488) first simulator drive was predictive of future on-road crashes; subsequent drives had no significant predictive value. A possible explanation for this finding is that individual differences in driving style may be more valid during the first simulator encounter, because the situation is new for participants and mental workload is high. In long drives, drivers may have adapted to the simulator experience, and make few errors and experience low mental workload.

It is unknown which psychological mechanisms may have caused the negative correlation between years of licensure and driving speed. Possibly, young drivers were more inclined to try out the simulator and drive as fast as possible. Another possible explanation is that the driving test provided a measure of confidence with computers (rather than actual information on how a participant drives in real traffic), and see [37–40] for the effect of age and experience on driving performance in a simulator. The motor show was not a quiet testing environment, and possibly this contributed to increased workload and distraction. It remains to be investigated how drivers would drive in a simulator-based test in a formal context, such as the hypothetical situation where a simulator test is part of the driver's license test. Most likely, drivers would then be more inclined to adhere to traffic rules, and driver behavior would be more homogenous.

People with a higher total risk score were less likely to agree with their error scores. Possibly, the simulator gave a too strict assessment for some drivers, and these drivers were therefore correct in disagreeing with their scores. On the other hand, these results may be interpreted in light of the fact that most drivers believe themselves to be better than the average driver [41]. Our study also found that the longer license-holders tended to agree more with their error scores. This may be a direct consequence of the fact that the error scores were better for the longer license-holders. This finding is also in line with work by Finn and Bragg [42] who showed that young drivers (incorrectly) perceived their chances of being involved in an accident as lower than those of their peers and older males, whereas older drivers (correctly) indicated that their chances of having an accident were comparable to those of their peers and less than those of young male drivers. The present results may point to a problem in teaching people how to drive safely, namely that those drivers who adopt the most deviant driving styles are the least willing to agree that they drive badly.

An important advantage of simulator-based testing as compared to on-road testing is the possibility to present identical scenarios to all individuals. However, car driving is a self-paced and it is crucial to assess voluntary aspects such as speeding and decision-making, for example regarding route choice. This raises the question of how to find a satisfactory balance between standardization and voluntariness of the driving task. In the present driving test, the simulator did not intervene if a participant did not follow the route instructions. If the participant drove in the wrong direction, he or she could continue driving but missed the remainder of the pre-programmed scenarios. A commonly used option is to let a simulator intervene automatically and put the car back on track with zero speed. Both aspects have disadvantages; in the first, not all participants encounter the same scenarios; in the second, the participant is interrupted and important performance measures, such as the mean speed, are distorted. A possible answer to this dilemma may be to use scripts, so that virtual agents adapt their behavior to the participant [43]. For example, if a driver makes a wrong turn, the situation can adapt so that the driver will still encounter the planned events without the need to stop and reset the vehicle [44].

The results showed that women were more susceptible to simulator sickness than men were (see also [45, 46]). The results also showed that driving experience correlated with simulator sickness and that sickness ratings systematically increased with years of licensure (Fig. 2). This systematic increase suggests that besides experience, age also contributes to motion sickness (see [47], for review on the effects of experience and age). Simulators are widely used for young driver training [48], and although simulators are not yet used for formal driving examination, this may be a logical next step. However, the present findings point to some reservations in this regard, because elevated incidences of simulator sickness should be anticipated when testing experienced (older) drivers. Severson et al. [49] proposed an interesting solution to simulator sickness in a test using a PC-based simulator with a relatively simple monitor, steering wheel, and pedals. The simulation offered an abstract visual environment in which participants had to make go/no-go decisions when approaching opening and closing gates. Severson et al.'s test was able to distinguish between healthy older drivers and older drivers with a brain disease, and there was no reported incidence of simulator sickness. Other studies have also shown that driver training and assessment can be effective with PC-based systems [25, 31].

Driving is a primary mode of transport, and driving privileges and independent mobility are critically important for most people. The introduction of new driver-testing tools is likely to meet with great skepticism and simulators will probably be accepted only when they can be shown to yield valid results. This study showed that years of licensure substantially correlates with driving speed and a total risk score in the simulator. This result corresponds to on-road-data showing that young drivers are more likely to be involved in speed-related crashes than older drivers (e.g., [7]) and to research showing that young drivers report more violations and higher sensation seeking scores than older drivers [50, 51]. However, this study provided just a first step in studying the potential usefulness of simulators for driver testing. More research will be required with alternative safety criteria, with other populations, and in other testing contexts (e.g., driver testing center instead of a motor show). Furthermore, this study highlighted challenges that must be anticipated when developing future simulator-based tests, namely issues of standardization and simulator sickness.

References

1. World Health Organization: Global status report on road safety. http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/ (2015)
2. Farmer, C.M.: Effects of electronic stability control: an update. Traffic Inj. Prev. **7**, 319–324 (2006)
3. Evans, L.: The dominant role of driver behavior in traffic safety. Am. J. Public Health **86**, 784–785 (1996)
4. Treat, J.R., Tumbas, N.S., McDonald, S.T., Shinar, D., Hume, R.D., Mayer, R.E., et al.: Tri-level study of the causes of traffic accidents: executive summary. Report DOTHS034353579TAC(5), National Highway Traffic Safety Administration, Washington, D.C. (1979)
5. De Winter, J.C.F.: Predicting self-reported violations among novice licensed drivers using pre-license simulator measures. Accid. Anal. Prev. **52**, 71–79 (2013)
6. Lee, J.D.: Technology and teen drivers. J. Saf. Res. **38**, 203–213 (2007)

7. Organisation for Economic Co-operation and Development [OECD]: Young Drivers: The Road to Safety. OECD, Paris, France (2006)
8. 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)
9. Baughan, C.J., Simpson, H.: Driving tests—test reliability, consistency of candidates performance and other issues. In: Rothengatter, T., Huguenin, R.D. (eds.) Traffic and Transport Psychology, pp. 411–420. Elsevier, London (2004)
10. De Winter, J.C.F., Kovácsová, N.: How science informs engineering, education, and enforcement: a message for driving instructors. In: Fisher, D., Caird, J., Horrey, B., Trick, L. (eds.) Handbook of Teen and Novice Drivers. CRC Press, Taylor & Francis Group, Boca Raton, FL (in press)
11. Arthur, W., Barret, G.V., Alexander, R.A.: Prediction of vehicular accident involvement: a meta-analysis. Hum. Perform. **4**, 89–105 (1991)
12. Heikkilä, V.-M.: Relationship of laboratory and on-road tests for driving-school students and experienced drivers. Percept. Motor Skills **90**, 227–235 (2000)
13. McKenna, F.P.: Accident proneness: a conceptual analysis. Accid. Anal. Prev. **15**, 65–71 (1983)
14. Ranney, T.A.: Models of driving behavior: a review of their evolution. Accid. Anal. Prev. **26**, 733–750 (1994)
15. Medeiros, F.A., Weinreb, R.N., Boer, E.R., Rosen, P.N.: Driving simulation as a performance-based test of visual impairment in Glaucoma. J. Glaucoma **21**, 221–227 (2012)
16. Salthouse, T.A.: Selective review of cognitive aging. J. Int. Neuropsychol. Soc. **16**, 754–760 (2010)
17. Sekuler, B., Bennett, P.J., Mamelak, M.: Effects of aging on the useful field of view. Exp. Aging Res. **26**, 103–120 (2000)
18. Brar, S.S., Rickard, D.P.: Teen and senior drivers. Report CAL-DMV-RSS-13-240, California Department of Motor Vehicles, Sacramento, CA (2013)
19. Evans, L.: Traffic Safety. Science Serving Society, Bloomfield Hills, MI (2004)
20. Underwood, G., Crundall, D., Chapman, P.: Driving simulator validation with hazard perception. Transport. Res. F **14**, 435–446 (2011)
21. Hancock, P.A., Caird, J.K., White, H.G.: The use of driving simulation for the assessment, training, and testing of older drivers. Report HFRL NIA 90-01, Human Factors Research Laboratory, Minneapolis, MN (1990)
22. Allen, R.W., Park, G.D., Cook, M.L., Fiorentino, D.: Training and assessment of novice drivers. In: Driving Simulation Conference (DSC) 2009 Europe, Monte Carlo, Monaco, pp. 91–102 (2009)
23. Bédard, M.B., Parkkari, M., Weaver, B., Riendeau, J., Dahlquist, M.: Assessment of driving performance using a simulator protocol: validity and reproducibility. Am. J. Occup. Ther. **64**, 336–340 (2010)
24. Hoffman, L., McDowd, J.M.: Simulator driving performance predicts accident reports five years later. Psychol. Aging **25**, 741–745 (2010)
25. Lee, H.-C., Lee, A.H., Cameron, D., Li-Tsang, C.: Using a driving simulator to identify older drivers at inflated risk of motor vehicle crashes. J. Saf. Res. **34**, 453–459 (2003)
26. Lew, H.L., Poole, J.H., Lee, E.H., Jaffe, D.L., Huang, H.-C., Brodd, E.: Predictive validity of driving-simulator assessments following traumatic brain injury: a preliminary study. Brain Inj. **19**, 177–188 (2005)
27. Stanton, N.A., Young, M.S., Walker, G.H., Turner, H., Randle, S.: Automating the driver's control tasks. Int. J. Cogn. Ergon. **5**, 221–236 (2001)
28. Uhr, M.B.F., Felix, D., Williams, B.J., Krueger, H.: Transfer of training in an advanced driving simulator: comparison between real world environment and simulation in a manoeuvering driving task. In: Driving Simulation Conference North America, Dearborn, MI (2003)

29. Dziuda, Ł., Biernacki, M.P., Baran, P.M., Truszczyński, O.E.: The effects of simulated fog and motion on simulator sickness in a driving simulator and the duration of after-effects. Appl. Ergon. **45**, 406–412 (2014)
30. Gálvez-García, G.: A comparison of techniques to mitigate simulator adaptation syndrome. Ergonomics **58**, 1365–1371 (2015)
31. Kuipers, J.: Multi variable strategy reduces symptoms of simulator sickness. In: Spink, A.J., Van den Broek, E.L., Loijens, L.W.S., Woloszynowska-Fraser, M., Noldus, L.P.J.J. (eds.) Proceedings of Measuring Behavior, pp. 239–243. Wageningen, The Netherlands (2014)
32. Allen, R.W., Park, G.D., Cook, M.L., Fiorentino, D.: A simulator for assessing older driver skills. Adv. Transport. Stud. 2007 Spec. Issue 71–80 (2007)
33. Hoedemaeker, M., Brookhuis, K.A.: Behavioural adaptation to driving with an adaptive cruise control (ACC). Transport. Res. F **1**, 95–106 (1998)
34. Maycock, G., Brocklebank, P.J., Hall, R.D.: Road layout design standards and driver behavior. TRL Report 332, Transport Research Laboratory, Crowthorne (1998)
35. De Winter, J.C.F., Wieringa, P.A., Kuipers, J., Mulder, J.A., Mulder, M.: Violations and errors during simulation-based driver training. Ergonomics **50**, 138–158 (2007)
36. McGehee, D.V., Lee, J.D., Rizzo, M., Dawson, J., Bateman, K.: Quantitative analysis of steering adaptation on a high performance fixed-base driving simulator. Transport. Res. F **7**, 181–196 (2004)
37. 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**, 763–771 (2009)
38. Konstantopoulos, P., Chapman, P., Crundall, D.: Driver's visual attention as a function of driving experience and visibility. using a driving simulator to explore drivers' eye movements in day, night and rain driving. Accid. Anal. Prev. **42**, 827–834 (2010)
39. Ni, R., Kang, J.J., Andersen, G.J.: Age-related declines in car following performance under simulated fog conditions. Accid. Anal. Prev. **42**, 818–826 (2010)
40. Shanmugaratnam, S., Kass, S.J., Arruda, J.E.: Age differences in cognitive and psychomotor abilities and simulated driving. Accid. Anal. Prev. **42**, 802–808 (2010)
41. Sundström, A.: Self-assessment of driving skill—a review from a measurement perspective. Transport. Res. F **11**, 1–9 (2008)
42. Finn, P., Bragg, B.W.: Perception of the risk of an accident by young and older drivers. Accid. Anal. Prev. **18**, 289–298 (1986)
43. Wassink, I., Van Dijk, B., Zwiers, J., Nijholt, A., Kuipers, J., Brugman, A.: In the truman show: generating dynamic scenarios in a driving simulator. IEEE Intell. Syst. **21**, 28–32 (2006)
44. Park, G.D., Allen, R.W., Rosenthal, T.J.: Flexible and real-time scenario building for experimental driving simulation studies. In: 3rd Workshop on Multimodal Interfaces for Automotive Applications, International Conference on Intelligent User Interfaces, Palo Alto, CA, 1–4 (2011)
45. Allen, R.W., Park, G.D., Cook, M.L., Rosenthal, T.J., Fiorentino, D., Viirre, E.: Novice driver training results and experience with a PC based simulator. In: Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Park City, UT, 165–170 (2003)
46. Matas, N.A., Nettelbeck, T., Burns, N.R.: Dropout during a driving simulator study: a survival analysis. J. Saf. Res. **55**, 159–169 (2015)
47. Johnson, D.M.: Introduction to and review of simulator sickness research. Research report 1832, U.S. Army Research Institute, Fort Rucker, AL (2005)
48. Pollatsek, A., Vlakveld, W., Kappé, B., Pradhan, A.K., Fisher, D.L.: Driving simulators as training and evaluation tools: novice drivers. In: Fisher, D.L., Rizzo, M., Caird, J.K., Lee, J.D. (eds.) Handbook of Driving Simulation for Engineering, Medicine and Psychology, pp. 30-1–30-18. CRC Press, Boca Raton, FL (2011)

49. Severson, J., Rizzo, M., Wagner, J., Cremer, J., Best, A.F., Severson, M.A.: Driving simulation: how low can you go? In: Driving Simulation Conference 2007, Iowa City, USA (2007)
50. Martinussen, L.M., Møller, M., Prato, C.G.: Assessing the relationship between the driver behavior questionnaire and the driver skill inventory: revealing sub-groups of drivers. Transport. Res. F **26**, 82–91 (2014)
51. Zuckerman, M., Eysenck, S., Eysenck, H.J.: Sensation seeking in England and America: cross-cultural, age, and sex comparisons. J. Consult. Clin. Psychol. **46**, 139–146 (1978)

The Usefulness of Augmenting Reality on Vehicle Head-up Display

Kyongho Kim and Yoonsook Hwang

Abstract Driving task requires cognitive processes to sense traffic environment, understand context and situation, and decide how to control vehicle. This cognitive workload is mainly caused by processing visual information acquired while driving. This visual cognitive workload is one of the key factors for driving safety with visual distraction. To reduce these visual workload and distraction, augmented reality technology is applied to head-up display (AR-HUD). However, AR-HUD can confuse driver by overlaying graphical object onto real traffic object, degrade driver performance, and as a result increase risk of accidence. To confirm whether the AR-HUD is useful enhancing driving safety or not, we conducted experiments on driver's cognitive response behavior under daytime and nighttime conditions. In this paper, we describe the AR-HUD system we have been developing and experiments and results.

Keywords Augmented reality · Head up display · Usefulness

1 Introduction

Motor vehicle accidents result usually from a complex interaction between the driver, vehicle and environmental factors. Analyses of traffic accidents indicate that human factors are a sole or a contributory factor in approximately 90 % of road traffic accidents [1, 2].

Driving task requires cognitive processes to sense traffic environment, understand context and situation, and decide how to control vehicle. This cognitive workload is mainly caused by processing visual information acquired while driving.

K. Kim (✉) · Y. Hwang
Electronics and Telecommunications Research Institute, Daejeon, Korea
e-mail: kkh@etri.re.kr

Y. Hwang
e-mail: hanulai403@etri.re.kr

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_55

This visual cognitive workload is one of the key factors for driving safety with visual distraction.

To decrease drivers' visual distraction head-up display (HUD) technology has been developed and launched in a lot of commercial vehicles. It was reported that the usage of HUD reduces drivers' distractions and improves driving safety, and is estimated to contribute lower vehicle crashes up to 25 % [3]. And the drivers had a faster response and they were more consistently controlled speed when using a HUD compared with a head-down display (HDD) [4].

To reduce visual cognitive workload in addition to visual distraction, augmented reality technology is applied to head-up display (AR-HUD). The augmented reality can enhance the traffic situation awareness of road vehicles, pedestrians, obstacles, and route [5]. One study proposed AR-HUD system providing all of the detected vehicles and pedestrians information to drivers through a transparent display installed in front of the driver, fitting to the drivers' view [6].

We developed AR-HUD system and conducted experiments on driver's cognitive response behavior under daytime and nighttime conditions to confirm whether the AR-HUD is useful enhancing driving safety or not.

2 AR-HUD System

We developed AR-HUD system and it is composed of four S/W modules: Object detection module, Situation awareness module, Presentation module, and Registration module (Fig. 1).

2.1 Object Detection Module

Driving tasks start from detecting various kinds of road objects such as nearby vehicles, pedestrians, traffic sign, traffic signal, lane, and so on. Equipped with sensors like RGB camera, IR camera, and radar this object detection module detects and tracks road objects at various kinds of weather conditions like daytime, nighttime, rainy day, and foggy day.

Recognition rate and processing time are important factors for this module. To enhance recognition rate and reduce processing time we optimized image

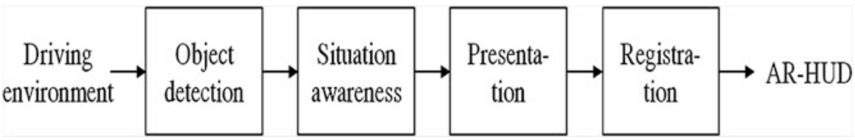


Fig. 1 System architecture of the proposed AR-HUD system

processing and feature detection algorithms, and classification algorithm such as Hough transform, HoG, and SVM. Utilizing GPU-based parallel processing and thread programming we could reduce processing time more.

2.2 Situation Awareness Module

All of the road objects detected by the object detection module should not be provided to driver because it can confuse driver by overlaying graphical object onto real traffic object, degrade driver performance, and as a result increase risk of accidence. So we have to aware driving situation by extracting meaningful context from the detected objects.

One of the meaningful situations is collision risk. This module computes time-to-collision (TTC) using location, velocity, acceleration, distance, and movement of road objects such as vehicle and pedestrian and ego-vehicle.

For collision risk warning strategy we referenced the international guidelines and regulations of NHTSA and EuroNCAP.

As another meaningful driving situation we identified the current driving lane. If we know which lane on I am driving now the more advanced navigation service can be made such as “make two lane changes to the left to turn left safely at the front intersection”.

The advanced driver assistance system (ADAS) data can also be integrated to this module to enhance situation awareness. We incorporated the blind spot detection (BSD) data to navigation with current driving lane awareness function. By doing this, the more advanced lane change assistance service such as “make one lane changes to the left attentively to the left side vehicle” could be possible.

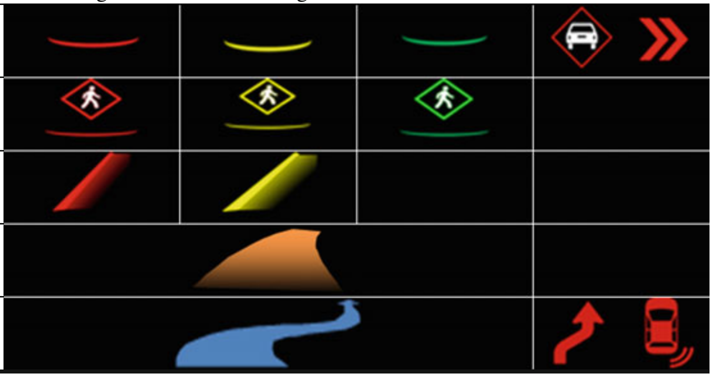
2.3 Presentation Module

Meaningful situations extracted from the situation awareness module are presented as graphics to HUD. The graphical components such as shape, color, brightness, and blink should be designed deliberately considering driving safety. It should be designed to be noticeable enough but not to take excessive cognitive attention.

We designed graphic objects and icons for several meaningful situations such as collision risk to vehicle or pedestrians, lane departure warning, lane curve warning, and lane change assistance (Table 1).

Table 1 Graphic objects and icons to present several meaningful driving situations

	Visual information			
Situation	Danger	Warning	Safe	Icons
Vehicle collision				
Pedestrian collision				
Lane departure				
Lane curve				
Lane change				



2.4 Registration Module

The shape and size of graphic objects generated from the presentation module are based on the objects detected from the object detection module. These objects are extracted from different images captured from various kinds of sensors like RGB camera, IR camera, or Radar.

Because the installation position of these sensors in vehicle and the sensor parameters differ from each other, the image coordinates are also different from each other. To present the graphic objects for driving situation fitted to driver's view, the image coordinate systems have to be transformed to the driver's view coordinate system and this process is called registration. We performed registration by extracting features from both sensor image and driver's view image, and matching those corresponding features to extract conjugate points, and then applying homographic projective transformation (Fig. 2).

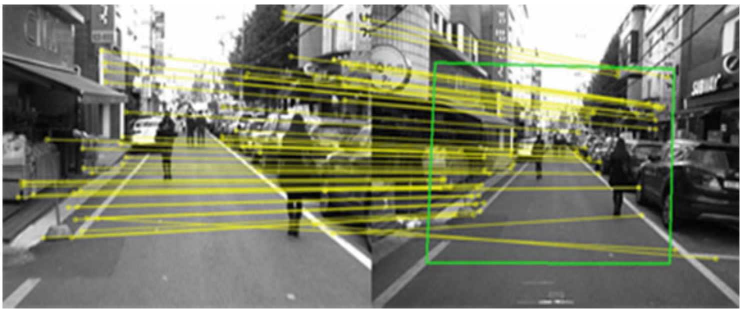


Fig. 2 Registration process between camera sensor image (*left*) and drive's view image (*right*). *Dots* in both images show conjugate points and lines show transformation

Fig. 3 Augmented reality services on HUD. This shows lane departure warning service, lane curve warning service, and lane change assistance service respectively from the *left*. These images are captured from our AR-HUD system by placing camera at the position of driver seat

2.5 AR Services on HUD

Through these four S/W modules, we implemented augmented reality services based on HUD. Those services are forward collision warning to vehicle and pedestrian, lane departure warning, lane curve warning, and lane change assistance. Figure 3 shows some examples of these AR-HUD services.

These services are expected to enhance driving safety and convenience. However, to verify whether it is true or not, we conducted experiments on driver's cognitive response behavior under daytime and nighttime conditions.

3 Experiment

3.1 Normal Visibility Condition

We had experimental study to ensure the effects of AR-HUD system on the risk perception behaviors of the drivers in the normal visibility condition (e.g. at daytime) [7].

The 31 drivers participated in this study, and we used a between-subject design. The participants were randomly assigned into the AR-HUD system usage condition and the control condition. We used the ANOVA to analyze the difference in risk perception behaviors between drivers for both conditions.

As a result, the response time of participants to risk events was a little fast, and the missing rate of risk events was fewer under AR-HUD condition than control condition. However, these differences were not statistically significant.

3.2 Low Visibility Condition

We also conducted experiment to confirm the effects of an AR-HUD system on the cognitive response behavior of the drivers under the low visibility weather conditions (e.g. at night, in rain).

The 35 male drivers participated in this experimental study. They participated in both in-vehicle AR-HUD system usage condition and control condition (within-subject design). Data for 4 drivers were excluded from the final analysis because they had fewer numbers of responses less than 50 % of the total experimental response. As a result, data of 31 drivers were used in the final analysis. The average age of the participants was 38.03 years (*SD* = 8.53, range: 25–29 age), and the average driving experience was 14.03 years (*SD* = 8.45, range: 2–33 years).

The experimental scenario was designed that drivers to response to the dynamic objects such as preceding vehicle and pedestrian under the low visibility conditions (night, rain, and fog). However, the preceding vehicle stimulus under the fog condition was not included in this study because of the safety reason, therefore, the number of experimental stimuli was totally 10. The arrangement of experimental stimulus is as follows:

> a pedestrian at night (control condition) → a pedestrian in rain (AR-HUD condition) → a preceding vehicle at night (AR-HUD condition) → a pedestrian in fog (control condition) → a preceding vehicle in rain (AR-HUD condition) → a preceding vehicle at night (control condition) → a pedestrian in rain (control condition) → a pedestrian at night (AR-HUD condition) → a preceding vehicle in rain (control condition) → a pedestrian in fog (AR-HUD condition).

We conducted this experiment in the indoor test-bed. It includes a 180-in. screen, a beam projector, a HUD device, a windshield, and a PC collecting the driver's response data and operating the simulation S/W (Fig. 4). Figure 5 shows examples of stimulus for control condition and AR-HUD condition used in this experiment.

We collected the response time of the participants to the preceding vehicles and the pedestrians for both conditions.

We used ANOVA to confirm the difference of the cognitive response behavior between the AR-HUD condition and the control condition. As a results, there was a statistically significant difference in response time to the dynamic objects between the AR-HUD condition (2.00 s) and the control condition (2.88 s; $t = 5.92$, $p < 0.001$).

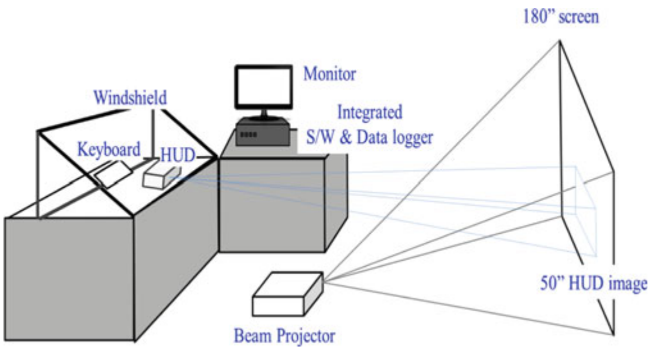


Fig. 4 The configuration of indoor test-bed

Fig. 5 Experimental stimulus. *Left* two are control condition and *right* two are AR-HUD condition. **a** pedestrian at night. **b** preceding vehicle in rain

In addition, the missing rate of the dynamic objects was less under the AR-HUD condition (11.11 %) than the control condition (27.60 %; $t = 5.76$, $p < 0.001$).

4 Conclusion

In this study, we developed AR-HUD system expecting that it will reduce driver's visual workload and distraction and finally enhance driving safety. To confirm whether the AR-HUD is useful enhancing driving safety or not, we conducted experiments on driver's cognitive response behavior under daytime and nighttime conditions. As a result, we found that AR-HUD is useful to have driver to response to the traffic situation faster and aware more of the driving situation. As a future work, we are planning to verity the usefulness of AR-HUD in real road situation for various kinds of weather and traffic conditions.

Acknowledgments This research was supported by the Industrial Strategic Technology Development Program of MOTIE [10040927, Driver-oriented vehicle augmented reality system based on head up display for the driving safety and convenience], and also supported by "Development of Advanced Digital Cockpit System for Driving Safety and Convenience" Project.

References

1. Lewin, I.: Driver training: a perceptual-motor skill approach. Ergonomics **25**(10), 917–924 (1982)
2. Rumar, K.: The role of perceptual and cognitive filters in observed behavior. In: Evans, L., Schwing, R.C. (eds.) Human Behaviour and Traffic Safety, pp. 151–165. Plenum Press, New York (1985)

3. Hibberd, D.L., Jamson, S.L., Carsten, O.M.J.: Managing in-vehicle distraction: evidence from the psychological refractory period paradigm. In: Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 4–11. Pittsburgh, (2010)
4. Liu, Y., Wen, M.: Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. Int. J. Hum. Comput. Stud. **61** (5), 679–697 (2004)
5. Victor, N.T.H., Karlin, B., Lee, B., Cuong, T., Rishabh, T., Srinath, S.: User-centered perspectives for automotive augmented reality. In: Proceedings of IEEE International Symposium on Mixed and Augmented Reality-Arts, Media, and Humanities, pp. 13–22. Adelaide, Australia, Oct 2013
6. Park, H.S., Park, M.W., Won, K.H., Kim, K.H., Jung, S.K.: In-vehicle AR-HUD system to provide driving-safety information. ETRI J. **35**(6), 1038–1047 (2013)
7. Hwang, Y.S., Park, B.J., Kim, K.H.: The Effects of toward AR-HUD system usage on drivers' driving behaviors. In: Proceedings of 2015 the Institute of Electronics and Information Engineers Summer Conference, pp. 1818–1820. IEIE Press, Seoul (2015)

Cognitive Constructs and Psychomotor Tracking

Rebecca Fosha, Lisa Durrance Blalock and Steven Kass

Abstract Psychomotor assessment tasks have been shown to be predictive of individual performance in aviation training. Predictive relationships have also been shown between performance in training and cognitive measures such as intelligence, processing speed, and spatial reasoning. The present study seeks to better understand the relationships between basic cognitive measures and a modern psychomotor tracking task. Participants completed a matrix reasoning task, a working memory capacity (WMC) span task, two reaction time tasks and a multiphase psychomotor tracking task. Only processing speed measures were correlated with psychomotor tracking tasks. Results support using psychomotor tracking tasks as a selection tool in addition to aptitude testing.

Keywords Psychomotor · Working memory · Processing speed · Fluid intelligence · Pilot selection

1 Introduction

A psychomotor task is one that involves a measurable muscular or motor activity paired with a perceptual or cognitive stimulus. The type of motor activity measured can vary from response speed in a simple tapping task to more complex tasks such as motor coordination while using a control stick and foot pedals to maneuver a vehicle [1, 2]. These tasks afford measurement of perceptual motor coordination and speed of discriminative reactions among a number of other measures that cannot be captured by a pencil and paper test. Thus, psychomotor tasks are used in

R. Fosha (✉) · L.D. Blalock · S. Kass
University of West Florida, 11000 University Parkway, Pensacola, FL, USA
e-mail: raf6@students.uwf.edu

L.D. Blalock
e-mail: lblalock@uwf.edu

S. Kass
e-mail: skass@uwf.edu

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_56

addition to aptitude testing in personnel selection for a number of professions including military aviation activities [3–5].

Early research in psychomotor performance utilized very simple mechanical apparatus tests. The psychomotor tasks currently in use for personnel assessment and selection are computer based cognitive and motor tasks that integrate perceptual stimuli with motor output. These more complex psychomotor tasks have shown good validity in personnel selection and training performance, adding predictive power of assessment tests beyond what is predicted by aptitude testing alone [1, 6–9]. Within the field of aviation, research has focused on the predictive power of psychomotor tasks for performance in military flight training [5, 8, 10, 11]. However, while research has focused on how psychomotor tasks predict performance, little research has examined the relationship between basic cognitive constructs and complex psychomotor performance. The present study seeks to examine the relationship between several simple measures of cognitive ability (processing speed, working memory capacity (WMC), and general fluid intelligence) and a modern computerized psychomotor task currently in use as a pilot selection tool.

2 Cognitive Constructs in Psychomotor Tasks

A literature review of psychomotor research implicates several cognitive constructs that may contribute to performance on psychomotor tasks. In addition, these constructs are known to be important to performance in complex vocations such as military aviation. The current review will focus on three key cognitive constructs that have been examined in the literature: processing speed, WMC, and general fluid intelligence.

2.1 Processing Speed

Processing speed is the speed in which an individual can make a predetermined response to a fairly simple stimulus [1, 6]. Processing speed is an important component to both psychomotor task performance and pilot training performance. In early psychomotor research, for example, Fleishman [1] consistently showed processing speed was a separate factor contributing to performance on psychomotor tasks, particularly when reaction time tasks were relatively simple [1, 2, 12, 13]. In addition, Fleishman [14, 15] found performance on reaction time tasks remained relatively stable with practice, indicating that it is a good measure of individual differences in ability.

Building on Fleishman's work, Ree and Carretta [9] looked for connections between cognitive function and psychomotor performance when they compared an established general aptitude test known as the Armed Services Vocational Aptitude Battery (ASVAB), and a recently implemented psychomotor test battery known as the Air Force Basic Aptitude Test (BAT). In one of the few studies directly comparing aptitude measures of general intelligence to psychomotor performance, they found a measure of reaction time to digit cancellation was more highly correlated with scores on the ASVAB than any tracking performance parameters measured. In a confirmatory factor analysis, they also found that reaction time showed the highest loading on a measure of intelligence referred to as *g*. Their results indicated that reaction time is a key component of performance on aptitude batteries even though pen and paper tests to not directly measure reaction time. Similarly, King et al. [16] found a modest correlation between training performance measures in USAF pilot trainees and simple reaction time tasks responding to visual and auditory signals. Much like the pilot training studied by King, psychomotor tasks such as the task involved in the present study require relatively quick reactions to score well, so it stands to reason that reaction time may be a key component of performance on these tasks.

Research indicates that the type of reaction time task used to study processing speed matters. Multiple studies have shown that performance on reaction time tasks that require higher-level executive functions, such as visual search, is susceptible to practice effects [6, 17]. Conway, Cowan, Bunting, Therriault, and Minkoff [18] suggested that visual search tasks require a representation to be maintained in short term memory and therefore draw on WMC to some degree. In addition, visual search can be seen as a skill, subject to learning [19]. Thus, this study relied on very simple reaction time tasks in an effort to identify reaction time as a separate construct from WMC.

2.2 Working Memory Capacity

WMC is described as a quantity of information kept in conscious awareness and important for problem solving and execution of complex tasks. This construct has been linked to performance on psychomotor tasks and pilot training performance. Chaiken et al. [6] examined the role of WMC in psychomotor performance using three WMC tasks: a spatial WMC task, a variation of a running span task, and verbal WM involving ordered word recall. They compared performance on these tasks to several different psychomotor test batteries; their confirmatory factor analysis yielded evidence for a single general psychomotor ability in contrast to Fleishman's earlier work [33]. They found working memory was highly correlated to a general cognitive ability, which was in turn highly correlated with the general psychomotor factor. They surmised this connection stems from the importance of WMC to a general intelligence known as *g*, and therefore WMC is a large component of psychomotor performance [6, 20].

The relationship between WMC and psychomotor performance may be influenced by an ability to deal with task complexity and novelty: simultaneous action of two limbs may require participants to apply two or more rules (complexity) and interpret novel movement of the target. Indeed, skill acquisition on new tasks has been related to cognitive resource availability [6]. This conclusion is also supported by King et al. [16] who showed moderate correlations between measures of memory, attention, and pilot training performance. In an effort to understand the contribution of WMC to psychomotor task performance independent of reaction time and fluid intelligence, we chose the Automated Operations Span (OSPAN) task, as it is well validated and has been shown to only measure WMC [21].

2.3 Fluid Intelligence

Fluid intelligence is commonly understood to be reasoning and problem solving abilities, and can be contrasted with crystallized intelligence, which is comprised largely of previously acquired knowledge such as word knowledge, arithmetic reasoning, and paragraph comprehension [22]. It is not surprising that a large body of research has linked measures of intelligence to performance in military aviation; general aptitude testing has been the main measure used in personnel selection for this profession [13, 16, 23, 24]. Modern multiple-aptitude selection test batteries such as the ASVAB and very similar Air Force Officer Qualifying Test (AFOQT) mostly measure a general intelligence known as *g* (comprised of both crystalized and fluid intelligence); the specific aptitudes that comprise these two batteries can be considered subsets of *g*, and add little if any contribution to its validity [25, 26]. Despite the important contribution psychomotor tests make to selection and their parallel assessment value (e.g., [16]), few studies have attempted to investigate the relationships between psychomotor tests and *g* and what little research has been done has heavily emphasized crystalized intelligence.

For example, Ree and Carretta [9] examined general cognitive ability assessed through the ASVAB and psychomotor performance on the BAT and showed a significant correlation between ASVAB scores (*g*) and BAT psychomotor tasks, including a two-hand coordination task very similar to the psychomotor task in the current study. Their factor analysis indicated performance on psychomotor tests was a higher order measure of *g*, along with scores on paper and pencil cognitive tests (also see [6, 27]). The specific aptitude tests they relied upon combined to measure *g*, but those tests relied heavily on the crystallized intelligence component of *g*.

King et al. [16] studied cognitive abilities with more influence from fluid intelligence such as attention/mental control, memory, reasoning/calculation, spatial processing and reaction time. While many of their reasoning tasks were heavily reliant on mathematics and other forms of crystallized intelligence, they also studied an object match task that attempted to measure abstraction and conceptual flexibility, which are less reliant on crystallized intelligence. This task showed only modest correlation with pilot training performance.

Research bears out a connection between WMC, fluid intelligence and a broader *g*; there is an ongoing debate as to the strength and nature of this relationship (see [28]). A meta-analysis shows that while WMC, fluid intelligence, and to a lesser extent processing speed share some variance, the relationship is not strong enough to conclude WMC and fluid intelligence are isomorphic abilities [28]. Therefore for this study, fluid intelligence was assumed to be measureable, relatively distinct from WMC, and was assessed using a matrix reasoning task very similar to well established fluid intelligence assessment tools, which draws on reasoning abilities but not on a capacity to maintain images or characters in memory as the OSPAN task does.

2.4 Current Study

The goal of this study was to understand how three cognitive constructs (processing speed as measured by simple and choice reaction time, WMC as measured by OSPAN, and fluid intelligence as measured by matrix reasoning) contribute to performance on a modern psychomotor task currently in use by the military for pilot selection. Our interest was in choosing simple and well validated measurements of each cognitive construct in order to obtain a clear picture of the contributions of each construct, keeping in mind that previous research has shown that there may be some overlap between these three cognitive abilities even in a very simple form. We then compared performance on these basic cognitive tasks with performance on a complex aviation selection test battery that has a heavy psychomotor component.

Given the varying influence and collective importance of processing speed, WMC fluid intelligence and psychomotor tracking to predicting pilot performance in training, we expect to find relationships between all measures and performance on the computer based psychomotor tracking tasks examined. If that is the case, consistently high correlations across the tests administered in this study will provide further support for the hypothesis that processing speed, fluid intelligence and WMC are important contributors to performance on psychomotor tracking tasks. On the other hand, if correlations of one or more of the cognitive constructs with psychomotor tasks are not significant, there may be more to understand about how and why psychomotor tasks lead to better performance in subsequent training.

3 Method

3.1 Participants

The study participants were university undergraduate students (N = 69, 16 males, 53 females) enrolled in psychology courses. The mean age of the sample was 22.96 years (SD = 6.22) with a range from 18 to 46. The participants were recruited through a research pool allowing extra credit points in psychology courses for participation in the study. Data for six participants were excluded due to software application errors that prevented completion of the psychomotor tracking test battery, resulting in 63 total participants. Each participant completed five cognitive tasks, the order of which was counterbalanced across participants using a Latin square. The tasks were a simple reaction time (simple RT) task, a choice reaction time (choice RT) task, an operation span task (OSPAN), a matrix reasoning task (matrix), and a psychomotor test battery.

3.2 Materials

Apparatus. The apparatus was a Dell Inspiron desktop computer with Intel 2.9 GHz processor, 15 in. CRT display at 1280 × 1024 resolution, and a Saitek X52 Pro Hands on Throttle and Stick (HOTAS) Flight controller. Reaction time, OSPAN, and psychomotor tasks were conducted using part or all of the apparatus. The matrix reasoning task was conducted with pencil and paper.

Reaction Time Tasks. In the simple RT task, each participant was asked to press a key on a keyboard when a green rectangle appeared on the screen. The rectangle was preceded by a "+" that was on screen for 800 ms followed by an open rectangle cue that appeared for 300, 600, 900 or 1200 ms prior to the target (randomized order). The choice RT task was similar to the RT task except that the stimulus was either a blue or green rectangle, and participants were asked to select a blue key (the "f" key on a standard keyboard labeled with a blue sticker) when the blue rectangle appeared or a green key (the "j" key on a standard keyboard labeled with green sticker) when the green rectangle appeared. Both tasks began with a short practice session of five target presentations. There were 60 trials for the simple RT task, 15 of each condition. Mean latency for each participant was measured. There were 120 trials of the choice RT task, 60 blue and 60 green with 15 of each delay (300, 600, 900, 1200 ms) in each color presented in a random order; mean latency and accuracy were measured.

Operation Span Task. An automated working memory operation span (OSPAN) task [21] was used to test WMC. The participants were presented with a simple mathematics equation with the solution represented by a question mark: 1/1–1 = ? followed by a number and two selection choices: "true" or "false". The participant had to choose if the presented number was the solution (true) or not (false). The participant was then presented with a letter. The trial phase varied from three to seven problem/letter sets. At the conclusion of a set the participant was asked to recall the letters in the set from a 3 × 4 grid of letters containing all of the letters presented with distracters. A practice equation set (without letter recall) determined the amount of time the equation screen was displayed. In the trial set if a "true/false" selection was not made within the mean response time from the math practice session plus 2.5 SD, the screen would automatically advance to the letter screen and the equation was counted as an error. Otherwise a mouse click on "true/false" advanced the screen to display a single letter for 800 ms. This practice equation set was followed by a practice letter recall set with no time restrictions.

The trial phase total math equation/letter recall set size was 75. The absolute OSPAN score of the total number of perfectly recalled sets was used [21].

Matrix Reasoning. 16-item matrix reasoning test was used to assess fluid intelligence. Each problem included eight objects presented in three rows and three columns, with the bottom right item missing. Eight possible answer choices were presented below the problem. The participant had to determine the relation among the items in the problem and then select the item that best completed the pattern, circling the correct answer on an answer sheet. Two practice problems were followed by 16 scored test problems with a 10 min time limit. The percentage of correct responses out of 16 was the measured score.

Psychomotor Tasks. This task had several components beginning with a spatial map orientation task. The participant viewed a tracker map on the left side of the screen showing an arrow that may point to any direction (north was always oriented at the top of the screen). On the right side of the screen, there was a perspective view that showed four squares ("parking lots") for each cardinal direction. The participant received an auditory cue directing them to select one of the four squares on the right panel based on the orientation of the arrow on the left panel.

Participants then completed a dichotic listening task. Participants were presented with a string of numbers and letters in each ear simultaneously and instructed to respond to specific cues (e.g., click the trigger on the joystick for even numbers, press a button on the throttle for odd numbers).

Following the dichotic listening task, participants completed a motor tracking task involving multiple components. Participants were asked to position a crosshair image over a target that moved vertically along the left side of the screen using the throttle with the left hand (throttle tracking) and position a crosshair image over a target that moved in two dimensions (up/down, left/right, and diagonally across the screen) over the center and right side of the screen using the joystick with the right hand (joystick tracking). In the most complex task, participants had to complete two tracking tasks and the dichotic listening task together.

Finally, participants completed the two tracking tasks simultaneously with emergency scenarios. In this task, participants were given instructions to resolve a fire, engine or propeller emergency by adjusting fuel and power indicators. In the trial phase, while conducting the vertical and aircraft tracking tasks, the participant received an audio "warning" cue for an emergency scenario. The to-be-executed procedure was cued by name at the bottom left portion of the screen. Upon receiving the auditory cue, the participant was required to execute the remembered procedure. If the actions were completed, the screen returned to normal. If the actions were not completed within a specific amount of time, the screen background changed to red and a warning appeared at the bottom of the screen indicating systems operating under duress. With no action or incorrect action taken, the screen would return to normal. This process repeated multiple times.

The psychomotor task software conducted a number of performance measurements. The measurements we focused on were the average stick tracking and throttle tracking scores across all tasks because our primary interest was in performance on the psychomotor task. Each score was a combination of the time the crosshair was in proximity of the target for each control axis and the number of redirects. A higher value would indicate that the participant was better able to track the target. The dichotic listening measurement was the total correct responses across tasks that involved dichotic listening. The spatial score used was the total correct responses to the spatial map orientation task.

3.3 Procedure

Participants completed a consent form prior to beginning the study. The allotted time for the study was 1.5 h, completion time varied due to participant latency in several portions of the test including reading the instructions for the psychomotor tasks. The order of tasks was pre-determined via a Latin square; the psychomotor tasks were a single task in the Latin square with the components presented in order (limitation of software program with a fixed order of presentation). Either one or two participants conducted the study simultaneously. Two computer stations were positioned facing away from each other and the participants conducted all tasks with headphones on to reduce distraction. At the conclusion of the study the participants received a debriefing form.

4 Results

4.1 Correlation Analyses

Descriptive statistics are contained in Table 1. A Pearson product-moment correlation was conducted between the joystick tracking score and the throttle tracking score to determine if the two measures were related. There was a significant positive correlation between the two variables ($r = 0.57$, $p < 0.01$). Thus, the two scores were averaged together and normalized for one overall psychomotor performance score. Correlations were also conducted between the cognitive constructs in the sample in order to determine if this sample displayed expected correlations found in the literature (see Table 2). WMC and Matrix reasoning were moderately correlated ($r = 0.25$, $p < 0.05$), as were simple and choice reaction time ($r = 0.26$, $p < 0.05$). This is in line with prior research and will be discussed further in the general discussion.

Bivariate correlations between cognitive constructs and scores on components of the psychomotor test battery revealed some significant results in addition to the expected relationships between cognitive measures. Simple and choice reaction times were negatively correlated with the combined psychomotor tracking score ($r_{simple} = -0.30$, $r_{choice} = -0.26$, $p < 0.05$) and dichotic listening ($r_{simple} = -0.26$, $r_{choice} = -0.26$, $p < 0.05$). WMC was moderately correlated with spatial reasoning

Table 1 Descriptive statistics for cognitive constructs and psychomotor scores

	Mean	SD	Minimum	Maximum
OSPAN	35.32	20.91	3.00	75.00
Matrix	0.54	0.17	0.12	0.94
Simple RT	344.97	150.25	246.65	1387.12
Choice RT	502.95	109.41	374.86	1127.78
Joystick	11.58	10.36	0.11	60.61
Throttle	31.31	15.75	1.76	94.92
Joystick/throttle combined	21.44	11.65	7.10	51.13
Dichotic listening	8.54	5.14	0.00	16.00
Spatial orientation	21.26	10.27	6.00	45.00

OSPAN = Total correct out of 75, Matrix = % correct out of 16, simple and choice RT = average latency in milliseconds, joystick and throttle = composite score related to total # of redirects and average distance of crosshair from target, joystick and throttle combined = average of joystick and throttle scores, dichotic listening = total correct, spatial orientation = total correct out of 48

Table 2 Bivariate correlations between cognitive constructs and psychomotor battery scores (N = 63)

	1	2	3	4	5	6
1. OSPAN						
2. Matrix	0.252*					
3. Simple RT	−0.149	0.168				
4. Choice RT	−0.202	−0.125	−0.264*			
5. Spatial task	0.324**	0.211	−0.128	−0.081		
6. Dichotic	0.227	0.219	−0.256*	−0.262*	0.321*	
7. Joystick/throttle	0.065	−0.040	−0.301*	−0.260*	0.322**	0.198

*$p < 0.05$
**$p < 0.01$

($r = 0.32$, $p < 0.01$) (Table 2). Spatial reasoning was correlated with both dichotic listening ($r = -0.32$, $p < 0.05$) and joystick and throttle tracking ($r = 0.32$, $p < 0.01$).

4.2 *Multiple Linear Regression Analysis*

A multiple linear regression was conducted to determine if simple reaction time, choice reaction time, OSPAN, and matrix reasoning (independent variables) predict psychomotor performance (dependent variable). All independent variables were entered simultaneously. These variables did not significantly predict psychomotor

Table 3 Multiple linear regression between cognitive constructs and combined joystick/throttle tracking score

	B	SE B	β	*t*	*p*
OSPAN	0.00	0.00	−0.10	−0.05	0.938
Matrix	−0.13	0.77	−0.22	−0.17	0.751
Simple RT	0.00	0.00	−0.25	−1.86	0.079
Choice RT	0.00	0.00	−0.20	−1.53	0.108

performance, $F(4, 58) = 2.15$, $p = 0.087$, $R^2 = 0.13$. None of the predictors significantly contributed to psychomotor performance, though some did trend in the expected direction (see Table 3).

4.3 General Discussion

We set out to learn more about the relationship between processing speed, WMC, fluid intelligence and performance on a complex psychomotor tracking task with the goal of understanding the basic cognitive constructs associated with psychomotor tracking. We expected to find a relationship between all cognitive constructs and psychomotor tracking task. While we did find a significant relationship between reaction time and psychomotor tracking, we did not find evidence of a significant predictive model between any of the cognitive constructs and psychomotor tracking.

While the overall hypothesis was not fully supported, the results did provide converging evidence of the relationships between the cognitive constructs and psychomotor performance. First, the demonstrated correlations between WMC and fluid intelligence are highly consistent with prior research (see [28]). The relationship between WMC and the spatial component of the psychomotor tasks is also supported by previous research (see [29]). This finding provides evidence that this sample displayed common relationships among the studied traits, and reduces the possibility that the variations specific to this sample produced abnormal results. Additionally, correlations between reaction time and dichotic listening (thought to be mediated by attentional control aspects of working memory) are also well established in literature (see [30, 31]).

Finally, correlations between simple and choice reaction time and joystick and throttle tracking tasks in this study support the importance of processing speed to psychomotor tasks. Multiple studies have shown processing speed to be important to both simple mechanical psychomotor tasks [1, 2, 6, 17] and modern computer based psychomotor tasks [9, 32]. Aptitude tests such as the ASVAB rely largely on declarative knowledge and crystallized intelligence and do not provide much opportunity to assess processing speed independent of those factors. Psychomotor tasks allow a measure of processing speed including physical reaction time that is completely independent of acquired knowledge. The results in this study suggest that the reaction time response measured by psychomotor tasks may be an important element to the predictive power of these tasks in training performance [16, 32].

A recurring theme in psychomotor research is a psychomotor factor or psychomotor equivalent to *g*, a factor comprised of several different characteristics and a cognitive ability in its own right. Fleishman repeatedly extracted nine to eleven factors related to basic abilities in early psychomotor tests [33]. In his early studies, a separate and independent factor, which could not be broken into more basic abilities, was deemed psychomotor coordination [2]. Wheeler and Ree [11] cited a later Fleishman work dismissing the idea of a general psychomotor skill. However, Wheeler and Ree found evidence for this skill in their study using linear models and avoiding potential errors of rotated factors favored by Fleishman. They found a psychomotor factor that was heavily influenced by a measure of reaction time, but not by other specific psychomotor abilities. A general psychomotor factor with only a small relationship to cognitive factors may underlie the importance of psychomotor task measurement [6, 20]. The results of this study support a general psychomotor factor in the absence of a relationship with cognitive measures. Additionally, the results of this study in combination with previous research also indicate that reaction time may be a key component of this ability.

5 Conclusion

In this study we set out to understand how three cognitive constructs (processing speed as measured by simple and choice reaction time, WMC as measured by OSPAN, and fluid intelligence as measured by matrix reasoning) contribute to performance in a modern psychomotor task in use by the military for pilot selection. The significant relationships found within measured cognitive constructs and between one cognitive measure and psychomotor task measure were expected and in keeping with well established understanding of cognitive abilities. Reaction time was found to have a moderate negative correlation with the psychomotor tracking task, supporting previous research indicating that reaction time is an important aspect of psychomotor tasks. However, the multiple regression analysis showed that none of the variables significantly predicted performance on the psychomotor tracking task.

The emergence of a significant relationship between spatial reasoning, and dichotic listening and joystick throttle tracking was not suspected in our initial hypothesis, nor was this relationship a primary interest. This significant result may be explained by the overlapping nature of cognitive abilities. It stands to reason that manipulating an object in space such as Fleishman's early tests [1] or positioning a joystick and throttle to move a representation on a screen would draw on spatial abilities, and this is demonstrated to some extent in the literature [20, 24].

The absence of expected significant relationships between WMC, fluid intelligence and the psychomotor task may be due to several factors. The prevailing trend in psychomotor research is to draw from pools of individuals selected for aviation training. It was hoped that a more diverse sample of the population would reduce the effects of range restriction and produce more pronounced significant relationships. Other studies also tend to be longer and more complex that the present study. Despite that difference, participants in this study often noted the difficulty of this study and in some cases demonstrated fatigue near the end of the 1.5 h. Although we accounted for order effects in the Latin square design and screened for outlying data greater than 2 SD (finding none that fell outside this range for more than one task), this difference in motivation may have diversified performance on the tasks in this study enough to stymie significant results with a small sample size. On the other hand, this finding may support the presence of a general psychomotor factor that is independent of the more well understood WMC and fluid intelligence. Some studies conclude that psychomotor factors are largely the same as cognitive abilities [27], while other research largely indicates that a psychomotor factor or ability has only limited shared variance with cognitive abilities [6, 23], and therefore is an important individual ability in its own right. In both cases cognitive skills were most often measured by broad aptitude tests that included a great deal of declarative knowledge in nearly every aspect of the test. The absence of a relationship between the simple cognitive tasks employed in the present study and psychomotor tracking supports the idea that there is a distinct difference between psychomotor skills and both WMC and fluid intelligence, while processing speed may be an important component of psychomotor skill.

Acknowledgments Permission to use the psychomotor tracking test battery in this study came from the Naval Aerospace Medical Institute, Naval Air Station Pensacola, Florida.

References

1. Fleishman, E.A.: Testing for psychomotor abilities by means of apparatus tests. Psychol. Bull. **50**, 241–262 (1953)
2. Fleishman, E.A.: Dimensional analysis of psychomotor abilities. J. Exp. Psychol. **48**, 437–454 (1954)
3. Causby, R., Reed, L., McDonnell, M., Hiller, S.: Use of objective psychomotor tests in health professionals. Percept. Motor Skill. **188**, 765–804 (2014)
4. Cuschieri, A., Hanna, G.B., Francis, N.K.: Psychomotor ability testing and human reliability analysis (HRA) in surgical practice. Minim. Invasive Ther. **10**, 181–195 (2001)
5. Griffin, G.R., Koonce, J.M.: Review of psychomotor skills in pilot selection research of the U. S. military services. Int. J. Aviat. Psychol. **6**, 125–147 (1996)
6. Chaiken, S.R., Kyllonen, P.C., Tirre, W.C.: Organization and components of psychomotor ability. Cogn. Psychol. **40**, 198–226 (2000)
7. Gibb, G.D.: Initial validation of a computer-based secondary selection system for student naval aviators. Mil. Psychol. **2**, 205–219 (1990)
8. Gibb, G.D., Dolgin, D.L.: Predicting military flight training success by a compensatory tracking task. Mil. Psychol. **1**, 235–240 (1989)

9. Ree, M.J., Carretta, T.R.: The correlation of general cognitive ability and psychomotor tracking tests. Int. J. Sel. Assess. **2**, 209–216 (1994)
10. Carretta, T.R., Ree, M.J.: Basic attributes test: psychometric equating of a computer based test. Int. J. Aviat. Psychol. **3**, 189–201 (1993)
11. Wheeler, J.L., Ree, M.J.: The Role of general and specific psychomotor tracking ability in validity. Int. J. Select. Assess. **5**, 128–136 (1997)
12. Fleishman, E.A.: The relationship between incentive motivation and ability level in psychomotor performance. J. Exp. Psychol. **56**, 78–81 (1958)
13. Fleishman, E.A., Hempel, W.E.: Factorial analysis of complex psychomotor performance and related skills. J. Appl. Psychol. **40**, 96–104 (1956)
14. Fleishman, E.A.: Psychomotor selection tests: research and application in the United States air force. Pers. Psychol. **9**, 449–467 (1956)
15. Fleishman, E.A.: A comparative study of aptitude patterns in unskilled and skilled psychomotor performances. J. Appl. Psychol. **41**, 263–272 (1957)
16. King, R.E., Carretta, T.R., Retzlaff, P., Barto, E., Ree, M.J., Teachout, M.S.: Standard cognitive psychological tests predict military pilot training outcomes. Aviat. Psychol. Appl. Hum. Factors **3**(1), 28–38 (2013)
17. Chaiken, S.R., Kyllonen, P.C., Tirre, W.C.: Organization and Components of Psychomotor Ability. United States Air Force Research Laboratory (1999)
18. Conway, A.R.A., Cowan, N., Bunting, M.F., Therriault, D.J., Minkoff, S.R.B.: A latent variable analysis of working memory capacity, short-term memory capacity, processing speed and general fluid intelligence. Intelligence **30**, 163–183 (2002)
19. Hillstrom, A.P., Logan, G.D.: Decomposing visual search: evidence of multiple item-specific skills. J. Exp. Psychol. Hum. Percept. Perform. **24**, 1385–1398 (1998)
20. Kyllonen, P.C., Chaiken, S.: Dynamic spatial ability and psychomotor performance. Int. J. Test. **3**, 233–249 (2003)
21. Unsworth, N., Heitz, R.P., Schrock, J.C., Engle, R.W.: An automated version of the operation span task. Behav. Res. Meth. **37**, 498–505 (2005)
22. Carpenter, P.A., Just, M.A., Shell, P.: What one intelligence test measures: a theoretical account of the processing in the raven progressive matrices test. Psychol. Rev. **97**, 404–431 (1990)
23. Carretta, T.R., Ree, M.J.: Pilot-candidate selection method: sources of validity. Int. J. Aviat. Psychol. **4**, 103–117 (1994)
24. DeKock, F., Schlecter, A.: Fluid intelligence and spatial reasoning as predictors of pilot training performance in the south african air force (SAAF). SA J. Ind. Psychol. (SA Tydskrif vir Bedryfsielkunde). **35** (2009)
25. Ree, M.J., Carretta, T.R.: The central role of g in military pilot selection. Int. J. Aviat. Psychol. **6**, 111–123 (1996)
26. Schmidt, F.L.: The role of general cognitive ability and job performance: why there cannot be a debate. Hum. Perform. **12**, 187–210 (2002)
27. Carretta, T.R., Ree, M.J.: Cognitive and Psychomotor Abilities: A Further Investigation of Their Relationship. United States Air Force Research Laboratory (1997)
28. Ackerman, P.L., Beier, M.E., Boyle, M.E.: Working memory and intelligence: the same or different constructs? Psychol. Bull. **131**, 30–60 (2005)
29. Logie, R.H.: Visuo-spatial Working Memory. L. Erlbaum Associates, East Sussex (1995)
30. Colflesh, G.J.H., Conway, A.R.A.: Individual differences in working memory capacity and divided attention in dichotic listening. Psychon. Bull. Rev. **14**, 699–703 (2007)
31. Conway, A.R.A., Cowan, N., Bunting, M.F.: The cocktail party phenomenon revisited: the importance of working memory capacity. Psychon. Bull. Rev. **8**, 331–335 (2001)
32. Carretta, T.R., Ree, M.J.: U.S. air force pilot selection tests: what is measured and what is predictive? Aviat. Space Environ. Med. **67**, 279–283 (1996)
33. Fleishman, E.A.: On the relation between abilities. Learn. Hum. Perform. Am. Psychol. **27**, 1017–1032 (1972)

Fatigue Driving Influence Research and Assessment

Gang Wu, Zhangwei Liu, Xiaodong Pan, Feng Chen, Meng Xu, Deshan Feng and Zhiguang Xia

Abstract Identification of different driving fatigue levels is critical for driving fatigue prediction and prevention, which will also promote the research and application of related driving assistance system. Although the driver's subjective fatigue feeling can be directly obtained by questionnaires, it varies a lot among different drivers and definition of unusual fatigue levels are sometimes arbitrarily determined. In order to get a relatively objective driver fatigue evaluation criterion, this study adopts a novel method which combines subjective fatigue level evaluation method (KSS) with driver face state video recognition technology to assess driver's fatigue level. Based on the analysis results, a new model was used to obtain the threshold value of driver fatigue levels, and the evaluation criteria of fatigue levels were established based on driver face state video recognition technology. And it is found that this method is more precise than that based on PERCLOS only.

Keywords Driver fatigue evaluation · Physiological features · Fatigue level

1 Introduction

The rapid progress of motorization has increased the number of severe traffic-related casualties. Fatigue driving has become an important cause of the traffic accidents. According to the accident investigation researches, nearly 60 % of the drivers surveyed admit that they have driven while fatigued at least sometimes, and nearly 2 % of them have been involved in fatigue related crashes [1, 2]. From crash data analysis, it is found that 26 % of all fatal and injury crashes in Ontario

G. Wu · X. Pan · F. Chen (✉) · M. Xu · D. Feng · Z. Xia
Laboratory of Road and Traffic Engineering of the Ministry of Education,
Tongji University, 4800 Cao'an Road, Shanghai 201804, China
e-mail: chenfeng.csu@gmail.com

Z. Liu
Shenzhen Urban Transport Planning Center, 3046 Aiguo Road,
Luohu District Shenzhen 518021, China

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_57

are fatigue related [3]. Approximately 20 % of all traffic accidents are related to fatigue driving in all over the world [4, 5]. However, potential risk factors of fatigue driving are still not clear, and it is not enough attention on the nature of fatigue driving and the accident influence of fatigue driving cause.

Summarized the existing related literature, which is mainly researched on the driving fatigue cause and performance, driving fatigue monitoring index and the prevention and treatment of driving fatigue. In the urban road driving fatigue accidents mainly happened in the morning and evening rush hour, the cause of which was considered as driving at night and lack of sleep [6]. Through recording the trucker condition continuously in highway, due to the highway landscape is monotonous and single operation in continuous the trucker appear easily fatigue above 2 h and the level of driving fatigue divided into clear-headed, slight fatigue and extreme fatigue [7, 8]. Meanwhile, it was discussed the relationship of eye movement with fatigue and discover the PERCLOS effectiveness when it was as the evaluation index of driving fatigue with average time closing eyes [9]. In laboratory simulation driving environment the driver's fatigue degree was defined by brain wave and blink data [10]. In the prevention and treatment of driving fatigue, a warning system was developed which will monitor the driver's eyes using camera and by developing an algorithm we can detect symptoms of driver fatigue early enough to avoid accident [11]. Using the developed rear-view mirror with CCD camera system insides and image processing system, the fatigue degree can be analyzed by detect driver's gaze direction and blinks [12]. How to analysis and prevent driving fatigue from the mechanism of production and impact effect is the matter of concern for researchers and the assessment of fatigue level is the fundament of fatigue driving influence research and selecting effective evaluation indexes and methods for accurate identification of driving fatigue.

This paper mainly confirms fatigue level according KSS asked subjective and external observation, then analyzed the relationship of brain waves, eye movements physiological characteristics and fatigue level. Using Rough Set Theory to integrate many indexes information such as driving period, driving duration, mouth opening state, blink frequency, aver-age time of eye closing and PERCLOS, and the level of fatigue criterion is determined.

2 Fatigue Driving Simulation Experiment

2.1 Experimental Objective

The relationship of the characteristic indexes and fatigue level is studied during the driving simulation experiment, and different fatigue state indexes threshold are determined. Then driving fatigue degree of decision rule is calculated by multi-source information data fusion.

2.2 Experimental Design

Select Experimenter. The experimenters are selected mainly considering the driver's gender and profession. In this experiment the number of effective experimenters are 31 whose age is between 21 and 48 and the average age is 29. The male is 26, the female is 5. Professional drivers are 8, non-professional drivers are 23. All drivers must be healthy, with a valid driver's license, the driving experience is more than two years, within one week before the trial no medication history, within 24 h before the test is not allowed to drink coffee, wine or other functional drinks.

Fatigue Control. In order to make the drivers appear in the limited experimental time driving fatigue state, two methods can be used to control the driver's fatigue before experiment, which is physiological trough method and sleep deprivation method. This experiment mainly chooses security guards on duty at night.

Experiment Equipment and Layout. The experiment equipment is used such as driving simulator *(Driving Simulator 2011)*, brain wave instrument, three computers, projector and two cameras. And driver security monitoring system and human body reaction time test system *(FD-HRT-A)* must be ready for this experiment. The equipment detailed arrangement and experiment field are shown in Figs. 1 and 2.

Experimental Procedure.

(a) Arrange experiment field, debug equipment and choose the driving scene of driving simulator 2011;
(b) Recording the basic information of experiment which include date, weather and indoor temperature;
(c) Driver wear brain all, familiar with the driving simulation platform and driving line;
(d) Starting experiment, all equipment is synchronous open, then test the driver's response time and ask the subjective evaluation fatigue level by KSS (shown in Table 1.) every 20 min. The last time is about 120 min;
(e) Save the video and experiment data in the end.

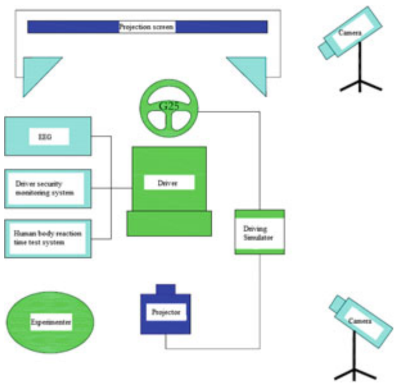


Fig. 1 Equipment detailed arrangement

Fig. 2 Experiment field arrangement

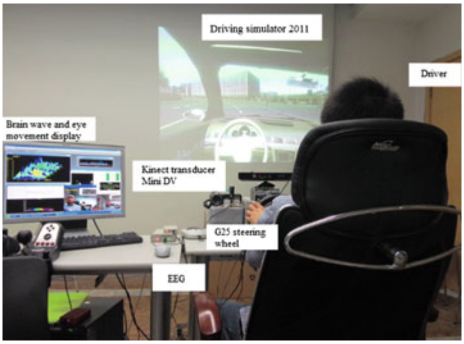


Table 1 KSS subjective fatigue cross-references

Subjective description	KSS fatigue level
Extreme awake	1
Every awake	2
Awake	3
Some awake	4
Neither awake nor drowsy	5
Began to show signs of sleepiness	6
Drowsy, without effort to stay awake	7
Drowsy, with some effort to stay awake	8
Sleepy, need great effort to stay awake	9
Heavy sleep	10

3 Research of Influence Driver's State on Fatigue

3.1 *The Relationship of Reaction Time and Driving Fatigue Level*

According the experiment data analysis, box diagram of reaction time distribution in different fatigue level was shown in Fig. 3. From the picture, the driver's reaction time has a tendency to increase with fatigue level deepened. The average reaction time is 601 ms when driver is awake. When driver is respective mild fatigue and drowsy it is 767 and 859 ms, that increase 166 and 258 ms compared with awake. The driver reaction time enlarges that will add the stopping sight distance, if the actual road conditions do not meet the stopping sight distance, it will exist potential safety hazard.

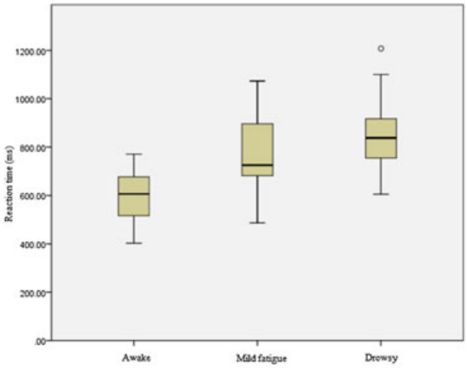


Fig. 3 Box plot of reaction time distribution in different fatigue level

3.2 The Relationship of Brain Wave and Driving Fatigue Level

It has been widely accepted that the value of (θ + α)/β power spectrum has a trend to gradually increasing and the value of slow α has a trend to gradually decrease with fatigue level deepened [13]. Figures 4 and 5 is respective the box diagram of (θ + α)/β and slow α power spectrum in different fatigue level. The average value of (θ + α)/β is respective 3.76, 3.61, 4.07 and 4.18 in the awake, mild fatigue, drowsy and sleep state. The average value of α is respective 402, 372, 270 and 155 in the awake, mild fatigue, drowsy and sleep state. The value of α reflects more clearly the fatigue level than (θ + α)/β for which downtrend is obvious.

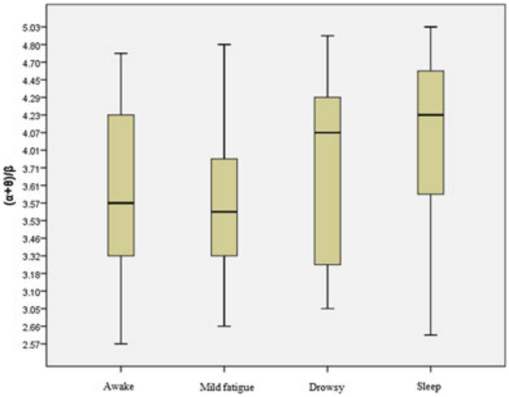


Fig. 4 Box diagram of (θ + α)/β power spectrum in different fatigue level

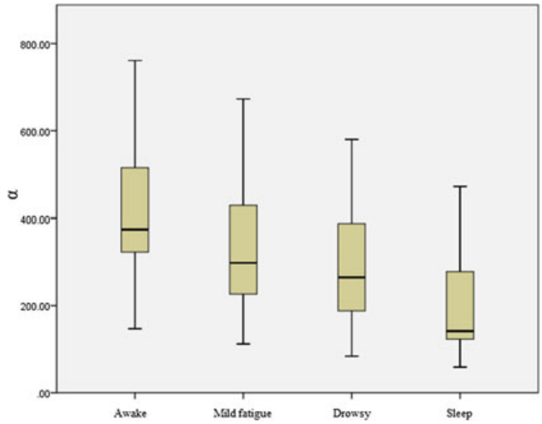


Fig. 5 Box diagram of α power spectrum in different fatigue level

3.3 *The Relationship of PERCLOS and Driving Fatigue Level*

According to statistical analysis of driver sample data, Box diagram of PERCLOS in different fatigue level is shown as Fig. 6. When driver is into the sleepy state, the PERCLOS increase. On the basis of the PERCLOS box diagram when the driver is awake or mild fatigue the value of PERCLOS mainly concentrates in between 0.05 and 0.10, and when the driver is mild fatigue or drowsy it mainly concentrates in between 0.07 and 0.25. This illustrates that the value of PERCLOS is effective using to distinguish between awake and mild fatigue.

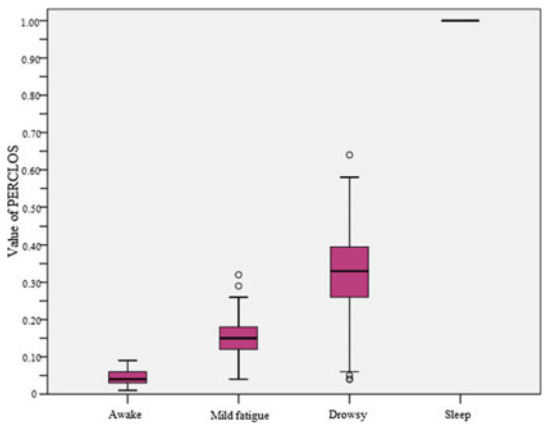


Fig. 6 Box diagram of PERCLOS in different fatigue level

4 Driving Fatigue Level Assessment Based on Rough Set Theory

4.1 PERCLOS Threshold Determine in Different Fatigue Level

This research uses the binomial Logistic regression analysis method to establish the fatigue level prediction model based on Logistic regression and determine PERCLOS threshold in different fatigue level. Define “awake” to “0” and “fatigue” to “1”, relation models were established between PERCLOS and fatigue state which were shown as Eqs. (1), (2).

$$\text{In}\frac{p_1}{1-p_1} = -15.848 + 147.570x_1 \tag{1}$$

$$\text{In}\frac{p_2}{1-p_2} = -9.796 + 38.345x_2 \tag{2}$$

where, P_1 and P_2 is respective the probability of fatigue and drowsy, $0 < P_1$, $P_2 < 1$; x1 is the PERCLOS threshold to divide awake and mild fatigue, $0 < x_1 < 1$; x_2 is the PERCLOS threshold to divide mild fatigue and drowsy, $x_1 < x_2 < 1$.

The Eq. (1) shows that smaller P_1, the lower the probability of fatigue, and PERCLOS threshold x_1 is smaller. Hence it is very key to determine PERCLOS threshold to keep system safety and reduce misjudgment rate. Select 0.1, 0.05, 0.02, 0.01 and 0.001 for P_1, plug this numbers into the Eq. (1). According to calculation results, when x_1 is 0.060 or 0.076, the awake recognition rate is a high level which is the ratio of the recognition awake number and total awake sample size and the awake misjudgment rate is a lower level which is the ratio of the number of fatigue misjudged as awake and total fatigue sample size. Similarly, P_2 is assigned as 0.1, 0.05, 0.02 and 0.01, mild fatigue recognition rate is one minus the drowsy misjudgment rate and mild fatigue misjudgment rate is one minus the drowsy recognition rate. Based on $x_1 = 0.060$, mild fatigue recognition and misjudgment rate decline with x_2 decreased and $x_2 = 0.198$ is as a reference threshold between mild and sleepy. The PERCLOS is 1 in sleeping state. So 0.06, 0.198 and 1 is respective as the threshold of distinguish awake, mild fatigue, drowsy and sleep.

According to the data of 167 fatigue samples appeared in simulation experiment, the result is that the awake recognition and misjudgment rate is respective 87.0 and 2.1 %, the mild fatigue recognition and misjudgment rate is respective 91.9 and 10.8 %, the drowsy recognition and misjudgment rate is respective 89.2 and 4.1 % and the sleep recognition and misjudgment rate is respective 93.6 and 3.5 %.

4.2 Driving Fatigue Level Assessment Based on Rough Set Theory

Considering the different data types and the large amount of data in the process of driver fatigue experiment, this paper adopted the distributed information fusion structure. First the original observation data is preliminarily analyzed and make judgment conclusion, then the conclusion information is passed to the fusion center, in the decision stage various data will be processed in terms of the rough set theory. At last it is concluded that the final judgment.

Attribute Discrete Normalization. These parameters are as condition attribute in rough set which are drive month, drive period, a single driving duration, mouth opening state, average time of eye closing and PERCLOS. The six indexes can be used to judge driver's fatigue state, and the index rule of discretization was shown in Tables 2 and 3. The discretization data is shown in Table 4 which is using the above rules to process experiment data.

Attribute Reduced of Decision Table. Attributes in decision table does not have the same importance, even certain attribute is redundancy, so some unrelated or unimportant knowledge must be deleted. Assume $U = \{u_1, u_2, \ldots u_3\}$, $C = \{a, b, c, d, e, f\}$, $D = \{Fatigue\ Output\}$, driver's fatigue level is judged by the above different attribute value. After simplify the decision table, simplified algorithm using the condition attribute was as follow:

- Step 1: calculate $pos_C(D)$;
- Step 2: $C_t = C - \{a_i\}$;
- Step 3: calculate $pos_{Ct}(D)$;
- Step 4: if $pos_C(D) = pos_{Ct}(D)$, a_i is deleted; if not, then keep it;
- Step 5: output a reduced set Condition attributes C for decision attribute D.

The reduced set simplify for attributes and the minimal decision arithmetic is shown as Table 5.

Table 2 The index rule of discretization

Condition attribute	1	0
a = Drive month	July and August	others
b = Drive period	22–8 h or 14–16 h	others
c = A single driving duration (min)	c ≥ 120	c < 120
d = Mouth opening state (aspect ratio)	d < 0.7 and keep 4 s	others
e = average time of eye closing (s)	e > 0.36	0 ≤ e ≤ 0.36

Table 3 The index rule of discretization

Condition attribute	0	1	2	3
f = PERCLOS	0 ≤ a < 0.060	0.060 ≤ a < 0.198	0.198 ≤ a<1.000	a = 1.000

Table 4 Discretization of driver's fatigue data

U (Sample capacity)	Time domain index			Physiological index			Fatigue output
	a	b	c	d	e	f	
u_1	0	0	0	0	0	0	0
u_2	0	0	1	0	0	0	0
u_3	0	0	0	1	0	0	0
u_4	0	0	1	1	0	0	0
u_5	1	1	0	0	0	0	0
u_6	1	1	1	0	1	0	0
u_7	1	1	1	1	0	0	0
u_8	0	0	0	0	1	0	0
u_9	0	0	1	1	1	0	1
u_{10}	0	0	0	0	0	1	1
u_{11}	1	1	0	0	0	1	1
u_{12}	0	0	1	0	0	1	1
u_{13}	1	1	1	1	0	1	1
u_{14}	0	0	0	0	1	1	1
u_{15}	0	0	0	1	1	1	1
u_{16}	1	1	1	0	1	1	2
u_{17}	1	1	0	1	1	1	2
u_{18}	0	0	1	1	1	1	2
u_{19}	1	1	0	0	0	2	2
u_{20}	0	0	1	0	0	2	2
u_{21}	1	1	0	1	0	2	2
u_{22}	0	0	1	1	0	2	2
u_{23}	1	1	1	1	0	2	2
u_{24}	1	1	0	0	1	2	2
u_{25}	0	0	0	1	1	2	2
u_{26}	1	1	1	0	1	2	3
u_{27}	1	1	0	1	1	2	3
u_{28}	0	0	1	1	1	2	3
u_{29}	1	1	1	1	1	2	3
u_{30}	0	0	1	1	0	3	3
u_{31}	0	0	0	0	1	3	3
u_{32}	0	0	1	1	1	3	3
u_{33}	1	1	1	1	1	3	3

According the Table 5, logical judgment of driver fatigue come true using the minimal decision arithmetic, meanwhile, when at least two value is "1" in driving fatigue time domain indexes, and the value of mouth opening state and average time of eye closing is “1”, the fatigue output of multiple attribute decision table is higher than that of single PERCLOS. Under other circumstance, the output result is keep an accordance.

4.3 Analysis of Assessment Result

According the data of 167 fatigue samples appeared in simulation experiment, the test result is shown as Table 6. The result demonstrate that the two methods have a high recognition rate for sleep, and for other fatigue levels the recognition rate of fusion multi-index decision table is all above 90 % which is superior to PERCLOS, the misjudgment rate is not more than 5 %. It indicates that fusing multi-index decision to judge the drive fatigue level is more accurate and comprehensive than single index detection method.

Table 5 The minimal decision arithmetic

Statistical index			Physiological index			Fatigue output
a	*b*	*c*	*d*	*e*	*f*	
0	0	0	0	0	0	0
0	0	1	0	0	0	0
0	1	1	1	0	0	0
0	1	1	1	1	0	1
0	1	1	0	0	1	1
0	0	1	1	1	1	1
1	0	1	1	1	1	2
0	1	1	1	1	1	2
1	1	1	1	1	1	2
0	1	0	0	0	2	2
1	1	1	0	0	2	2
0	0	1	1	1	2	2
0	1	1	1	1	2	3
1	1	0	1	1	2	3
–	–	–	–	–	3	3

(“–” on behalf of any value)

Table 6 Results comparison of fusion multi-indexes decision table and PERCLOS

Sample	Recognition and misjudgment rate	Multiple indexes logic judgment	PERCLOS
167	Awake recognition rate	91.4 %	87.0 %
	Awake misjudgment rate	2.0 %	2.1 %
	Mild fatigue recognition rate	93.8 %	91.9 %
	Mild fatigue misjudgment rate	5.3 %	10.8 %
	Drowsy recognition rate	92.5 %	89.2 %
	Drowsy misjudgment rate	3.2 %	4.1 %
	sleep recognition rate	95.1 %	93.6 %
	sleep misjudgment rate	2.8 %	3.5 %

5 Conclusion

This paper proposes a method of assessing driving fatigue level which is used to determine the driving fatigue level with the data that is from KSS objective fatigue level and amended by external observed fatigue evaluation standard.

1. The relationship of driver's state on fatigue and influence elements is studied in this paper. In fatigue state the driver's reactive time increase 0.167–0.259 s, the value of α reflects more clearly the fatigue level than (θ + α)/β for which downtrend is obvious, and the value of PERCLOS is effective using to distinguish between awake and mild fatigue.
2. Based on binomial logistic regression the PERCLOS threshold calculation method is proposed in different fatigue states. Preliminary threshold is certain using the relationship model of fatigue level and PERCLOS value, which is amended through sample recognition rate and misjudgment rate. The threshold was eventually corresponding to 0.06, 0.198 and 1 as mild fatigue, drowsy and sleep.
3. Driving period, driving duration, mouth opening state, average time of eye closing and PERCLOS were selected for the driving fatigue evaluation indexes. It indicated that the method which based on the rough set theory fusing multi-index decision to judge the drive fatigue level is more accurate and comprehensive than single index detection method.

Acknowledgments This research was jointly sponsored by Project 51578417 supported by the National Natural Science Foundation of China and the Fundamental Research Funds for the Central Universities.

References

1. Vanlaar, W., Simpson, H., Mayhew, D., Robertson, R.: Fatigued and drowsy driving: a survey of attitudes, opinions and behaviors. J. Saf. Res. **39**(3), 303–309 (2008)
2. Zhang, G., Yau, K.K.W., Zhang, X., et al.: Traffic accidents involving fatigue driving and their extent of casualties. Accid. Anal. Prev. **87**, 34–42 (2016)
3. Elzohairy, Y.: Fatal and injury fatigue-related crashes on Ontario's roads: a 5-year review. In: Highway Safety Roundtable and Fatigue Impairment, Driver Fatigue Symposium, Toronto, May 16th 2007
4. MacLean, A.W., Davies, D.R., Thiele, K.: The hazards and prevention of driving while sleepy. Sleep Med. Rev. **7**(6), 507–521 (2003)
5. Fernandes, R., Hatfield, J., Job, R.S.: A Systematic investigation of the differential predictors for speeding, drink-driving, driving while fatigued, and not wearing a seat belt, among young drivers. Transp. Res. Part F Traffic Psychol. Behav. **13**(3), 179–196 (2010)
6. Fell, D.L., Black, B.: Driver fatigue in the city. Accid. Anal. Prev. **79**(4), 463–469 (1997)
7. Williamson,A.M., Feyer, A.M., Friswell, R.: The impact of work practices on fatigue in long distance truck drivers. Accid. Anal. Prev. **28**(6), 709–719 (1996)
8. Hattori, H., Matsuura, Y., Nammiya, K., et al.: Fatigue and fatigue research: the Australian experience (1987). Paper Presented to 7th Biennial Australasian Traffic Education Conference, Speed, Alcohol, Fatigue, Effects. Brisbane (Feb 1998)
9. Xu, L., Li, S., Bian, K., et al.: Sober-drive: A Smartphone-Assisted Drowsy Driving Detection System. In: Proceedings of 2014 International Conference on Computing, Networking and Communications, pp. 398–402 (2014)
10. Yin, Y.: Research on driver fatigue state based on analyzing EEG and eye blink in stimulated driving environment. Tongji university master's degree thesis, (2008)
11. Singh, H., Bhatia, J.S., Kaur, J.: Eye tracking based driver fatigue monitoring and warning system. In: India International Conference on Power Electronics, IICPE (2010)
12. Adachi, K., Itoh, M., Mori, S., et al.: A driving simulator with driver monitor system. Reports of the Faculty of Science and Technology, Meijo University vol. 45, pp. 63–9 (2005)
13. Pfurtscheller, G., Aranibar, A.: Event-related cortical desynchronization detected by power measurements of scalp EEG. Electroencephalogr. Clin Neurophysiol. **42**(6), 817–826 (1977)

Research on Design Pattern of City Tunnel Side Wall Based on the Driver Visual Effect

Zhiguang Xia, Yonggang Lv, Xiaodong Pan, Feng Chen, Meng Xu, Gang Wu and Deshan Feng

Abstract When driving in the tunnel drivers often feel that the visual environment is not comfortable and the ability of speed perception drops. At the same time the enclosed environment of tunnel makes drivers feel nervous. All of these characteristics of driving in the tunnel are the causing factors for the accidents. This study focused on the design pattern of city tunnel side wall and investigated the design of city tunnel side wall in Shanghai city. On the base of color psychology and human factors engineering, the design pattern of city tunnel side wall including design style, color and the spacing of the pattern was analyzed and studied. The fuzzy elevation method was applied to elevate the new design patterns. The comfort and rationality of the new design patterns had been verified and the results provide guidance to the design pattern of city tunnel side wall.

Keywords Traffic safety city tunnel · Driver guidance system · Design pattern of city tunnel side wall · Fuzzy elevation method

1 Introduction

In China to release traffic pressure, the number of the urban tunnels is increasing. The constructions of urban tunnels bring a lot of advantages. Several studies [1–3] have researched on them and acknowledged the utilization of urban tunnels, especially at important traffic hub and regional business center, saving surface land; cutting vehicle mileage; and reducing traffic in important city streets. However the road tunnel traffic accidents bring enormous loss at the same time. To ensure the

Z. Xia · X. Pan · F. Chen (✉) · M. Xu · G. Wu · D. Feng
Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai 201804, China
e-mail: chenfeng.csu@gmail.com

Y. Lv
CCCC Highway Consultants Co., Ltd., No. 85 Deshengmenwai Street, Xicheng District 100888, Beijing, China

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_58

safety standard of the tunnel, we need to improve the driving environment and make it more comfortable. In this study focuses on which pattern and color collection design of city tunnel side wall is more suitable for the drivers.

The specification of design pattern of tunnel side wall has not been put in China. So the different tunnels have different designs. We made a research on the Yan'an East Road Tunnel and East Fuxing Road across-river tunnel in Shanghai China and paid more attention on the design pattern of the side wall and the traffic facilities. There were three features we found out: (1) the air tight environment; (2) the color of the side wall was white; (3) there was no pattern and tag on the side wall. These features have effect on the drivers. The condition of these two tunnels are presented in Figs. 1 and 2. When driving in the air tight environment drivers need to keep distance from the side wall to prevent crashing on the wall. This sustained situation makes drivers nervous and want to drive out of the tunnel quickly. The refection is called side wall effect [4]. If drivers are long be in the side wall effect, they will react slowly and have difficulty in controlling the cars correctly. In addition the closed environment will make drivers drop the speed sense. Driving on the open road, drivers can justice their speed relay on the surrounding scenery. While considering the features of inside tunnel environment, the speed sense drops seriously. Drivers depend on instrument panel to control the speed. Some research find that the fixation points often distribute on the middle-blow area and it means that drivers spend more time on the instrument panel [5]. To avoid the visual distraction there is no pattern on the side wall but this condition makes the surroundings monotonous.

The environment of the tunnel is very important for the traffic safety so many studies have worked on it and analyzed its influence on the drivers [6]. M.P. Manser

Fig. 1 The Yan'an East Road tunnel

Fig. 2 The East Fuxing Road tunnel

and P.A. Hancock sought to determine if the type of visual pattern and presence of texture applied to transportation tunnel walls differentially affected driving performance. Then they found out that the visual patterns with decreasing width made drivers slow down. It proved that the typical transportation environment and texture placed on the walls of the tunnel had a marked impact on speed perception and subsequent driving behaviors [7]. Katja Kircher, Christer Ahlstrom selected the tunnel lighting design as the subject. The results showed visual attention which was given to the driving task is the more crucial factor for giving rise to significant changes in both driving behavior and visual behavior. The results also indicated that light-colored tunnel walls are more important than strong illumination to keep the drivers' visual attention focused forward [8, 9]. As long tunnels become more and more, the periods of monotonous driving is longer and this condition leads drivers to be boredom and fatigue.

Considering the particularities of the road tunnel environment, the road tunnel accidents are different from the open road accidents. If the road tunnel accidents happen, the consequences are often severe. Now most research efforts have been allocated to accident analysis. When drivers approach to the tunnel portal, they need to change their driving style and they often focus on the entrance which cause the loss of information. Drivers generate anxiety when driving within the dark narrow and monotonous situation. These behaviors have negative on the safety [10]. Nussbaumer also mentions that according to police reports, the main reasons for tunnel crashes are lacking vigilance (e.g. fatigue distraction, inattention) and behavior aspects (e.g. safety distance, lane keeping, overtaking) [11].

Generally speaking topics of road tunnel are fall into two categories: the road tunnel accident rate and the traffic facilities design. The aim of this study was to investigate the influence of the color and design pattern of the tunnel side wall on the attentive and visually distracted drivers. The investigation was divided into two parts: the color and the pattern dimension. To choose the fitful color collection and pattern design, a simulator study was conducted. The fuzzy elevation method is applied to elevate if the new design patterns is better than the using urban tunnel.

2 Materials and Methods

Some studies of the road tunnel found that setting the patterns by certain distance on the side wall will improve the speed sense of the driver. Different types of visual pattern and presence of texture differentially affected driving performance [6]. The characteristics of the visual attractions was big bright and colorful. A suddenly visual attraction will attract the drivers especially when the stimulus appear on the peripheral of the vision [12]. However the qualitative analysis has accounted for a large proportion of the studies so far. The quantitative analysis like the setting space of the pattern is less. This work sought to find out the suitable design pattern of the tunnel side wall and put forward the recommended dimension by the theoretical analysis and experiment.

2.1 Pattern Design Research

Pattern Form. The pattern form should be simple and consistent. If the pattern form is complicated the drivers need to spend time on identifying, so it may distract the drivers and weaken the security. In addition the consistency takes an important part in visual induction. We chose wave and arrow as the pattern form design which are presented by Figs. 3 and 4. Choosing these two patterns for three reasons.

1. These two kinds of patterns are simple and sharp silhouette. Therefor drivers are easy to recognize.
2. The continuous waves is strongly consistency and the arrow has the mean of instruction. All of them can lead driving.

Pattern Dimension. There two requirements for the simulation to attract drivers: one is stimulus intensity the other one is enough action time. Generally speaking the bigger the patterns are the drivers are more easily to recognize. But if the pattern is so big the pattern cannot coordinate with the surroundings. Chinese specification "Road traffic signs and markings" stipulated the relationship between Chinese characters' Height and the design speed [13]. It is showed on Table 1.

Although the pattern dimension design is different from the Chinese characters' height but all of them require the high recognition. So we can consult the specification to study the range of pattern dimension. In the pattern design we suppose that

Fig. 3 The wave pattern

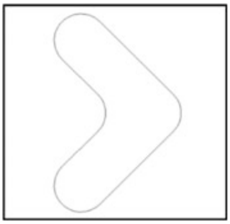


Fig. 4 The arrow pattern

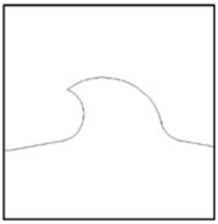


Table 1 The relation between Chinese character's height and design speed

Design speed (km/h)	Chinese character's height (cm)
100–120	60–70
71–99	50–60
40–70	35–50
<40	25–30

the design speed is 60 km/h, the height of the pattern is 40–80 cm and the width corresponds with the height. The details of the design are showed in Figs. 5 and 6.

Distance between the Road Surface and the Pattern. To attract the drivers the distance between the road surface and the pattern should be as high as the sight. The height of eye level of small vehicles is 1.2 m and the truck's is 2.5 m. Pattern need to be arrayed in two rows. One of the row is arrayed on the height of 1.2 m and the other one on the height of 2.5 m. The design plan imitates the small vehicles so the distance between the road surface and the pattern is 1.4 m. The details of the design are shown in Figs. 7 and 8.

The Spacing of Pattern. During driving drivers' physiological activities always follow the process that stimulation-feeling-judgement-operation [14]. Stimulation is the starting of drivers getting the information. When some objects appear in the drivers' vision as certain rule and drivers can receive visual pulse signal. Flash rate

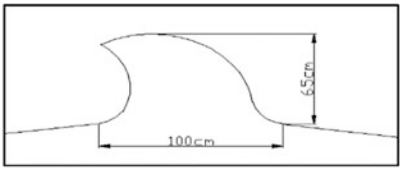


Fig. 5 The size map of the wave pattern

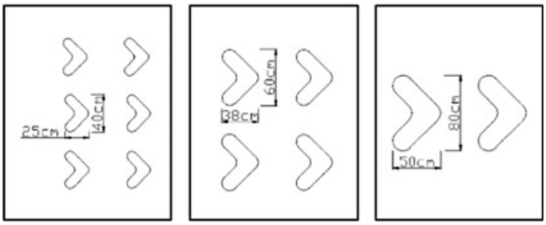


Fig. 6 The size map of the arrow pattern

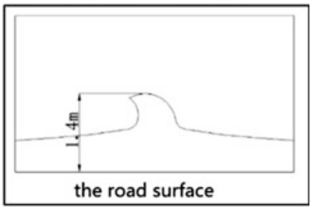


Fig. 7 Distance between the road surface and the wave pattern

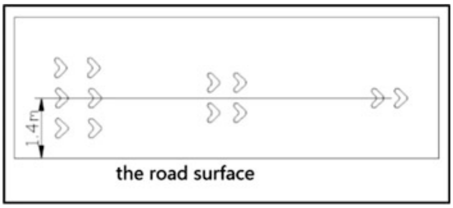


Fig. 8 Distance between the road surface and the arrow pattern

Table 2 The flash rate of the lane boundary dotted line

Road classification	Design speed V (km/h)	Dotted lines $d_s + d_k$ (m)	Flash rate (Hz)
Freeway	120	6 + 9	2.2
First-class highway	100	6 + 9	1.9
Second-class highway	80	2 + 4	3.7
Urban expressway	80	6 + 9	1.5
Secondary road	50	2 + 4	2.3

means the appearance frequency of the subject which generates the visual stimulation [15]. Drivers pay more attention on the road when they are driving. The lane boundary line is the dotted line so they appear into the drivers' vision by the certain flash rate. Chinese specification "Road traffic signs and markings" stipulated the flash rate of the dotted line [13]. It is based on the different design speeds and the different road grades like freeway and first-class highway and second-class highway and so on. The data has been listed on Table 2.

d_s means the length solid part of the dotted line; d_k means the length between the adjacent solid part. The calculating formula of the flash rate is as follow.

$$f = \frac{V}{3.6 * (d_s + d_k)} \tag{1}$$

The meaning of the symbols:

f (Hz) the flash rate
V (km/h) the design speed
d_s (m) the length solid part of the dotted line
d_k (m) the distance between the adjacent solid part.

From the table we can find out that the flash rate is smaller than the maximum limit which the American research reported (the flash rate of the urban road is lower than 8 Hz and the suburban roads' is lower than 4 Hz). In the flash rate term the dotted line is set reasonably. If the pattern of the tunnel side wall appear to the drivers with suitable flash rate the spacing of pattern must be set appropriately. From the analyze we know that the maximum limit of the flash rate is 4 Hz. We can use this value to make out the minimum spacing of pattern. Through the Formula (1) we can get the calculating formula of distance between the adjacent patterns.

$$L_{min} = \frac{V}{3.6 * f_{max}} \tag{2}$$

In the design plan we assumed the design speed was 60 km/h and the f_{max} was 4 Hz. Then we can get the minimum spacing of pattern is 4.2 m. To get the best visual reaction the spacing of pattern should be shorter than the L_{min} and cannot be too long as well. The suitable range of the spacing will keep the continuity and enhanced speed sense.

2.2 The Color of the Pattern

In the tunnel the color of the side wall and the lighting device take important part in building visual environment. They have influence on the drivers' psychology [16]. Some researchers found that the dark tunnel walls produced more very long glances away from the forward roadway than the light-colored tunnel walls. One possible explanation could be that the drivers used more foveal vision in order to track the road edges, as they were more difficult to see when the walls were dark [7]. Based on the consideration of this the color of the side wall always choose light-color especially the white. In fact the white color will become dazzling and stark under the tunnel lights. In this condition the white may bring distracted and uncomfortable feeling to the driver. So which kinds of color are more ideal for the tunnel side wall? We chose four common color of traffic and made a comparison of the colors.

The Comparison of Four Colors. Firstly we contrasted psychological influence physiology and meaning of the four colors. As we all know that different colors give people different feelings and they all have different symbol.

1. Blue: this color always makes people concentrate their mind and stabilizes emotions. It also means lucidity intellect and calm.
2. Green: this color makes people feel relaxed and cure tiredness. Green is the main color of the nature so it represents vitality and fresh. Sometimes it means peaceful and serene as well.
3. Red: this color brings people the feeling like excited and make people increase vigilance. Red color is high penetrating and has the meaning of prohibiting and dangerous.
4. Yellow: this color make people feel warm and remind the dangerous. It always means sunlight happiness and hope.

Some researchers used the experiment to demonstrate the different influence of different colors. The results showed that when people were in the room whose walls were painted by blue and green the temperature of the skin decreased 1–2 °C and pulse slowed. On the contract the temperature of the skin increased and the pulse quickened when people were in the room whose walls were painted by red and yellow. From the results we can conclude that the blue and green color will release the pressure and relieve the tension. It will help for the traffic safety.

Secondly we contrast the optical illusion of the colors. It is said that different color will change the eyesight distance. When we keep distance with the cars the eyesight distance is different for different colors. With the cold color like blue and green, the eyesight distance is further than the real distance. With the warm color like red and yellow, the eyesight distance is shorter than the real one. The depressing circumstances of the tunnel is an important problem for the safety. Considering the optical illusion of the color if we paint the tunnel side wall with cold color the depressing circumstances will be released. However the optical illusion of the color is also related to the brightness and concentration of the color. When we use the optical illusion of the color we need to control the brightness and concentration.

Fig. 9 The color collocation

From the comparisons we can find out that the blue and green are suit for the tunnel side wall. Therefore we choose blue and green to make combination and pick up the most suitable one.

The Collection of the Color. The color of design pattern should accord with the color of tunnel side wall. To avoid the landscape being monotonous, we made a combination of the blue and green color. When the tunnel side wall was blue the pattern would be green. This design would bring sharp contrast and strong visual impact. It also emphasizes the meaning of the pattern. The color collection is shown in Fig. 9.

2.3 Participants

Participants in this study were 15 females and 17 males between 20 and 27 years of age (mean = 22.28, standard deviation = 1.63) recruited from the staff and student body of the Tongji University. All of these participants took part in two experiments. One of the experiment was for the test of the spacing of the pattern and the other one was for the color combination of the pattern and the side wall. Participants did not receive any monetary compensation or class credit for their involvement. Participants possessed a valid state of Chinese driver's license and had the experience of driving on the tunnel. There was no colorblind of the participants and their vision was good.

2.4 Driving Simulator and Driving Scenario

The studies were conducted by the driving environment simulator—SCNeR simulator of Tongji University. The SCNeR simulator consisted of spherical cockpit and steel bracket and control platform. The screen was 250° curved screen so it can create real driving environment. We used SCNeR to establish the driving scenarios. The driving scenarios employed in all experimental conditions was representative of a typical urban roadway segment consisting of a two lanes 1.00 km in length. Each lane was 3.00 m wide, colored light gray, and separated from the adjacent lane by a dotted line. Shoulders on each side of the roadway were 0.25 m wide and colored dark gray. The altitude of all the tunnel did not change. There was no curve sequence of the tunnel. The ceiling of the tunnel was colored gray with lights only. The design pattern of the tunnel side was the only changing element of the tunnel.

2.5 Procedures

After completing the establishment of the driving environment participants began the experiments. The experiments consisted of two parts.

The first experiment was about the spacing of the design pattern. As we have discussed we chose the arrow and wave as the tunnel side wall pattern. Then we set five kinds of spacing distance and they were 5, 10, 15, 20 and 30 m. Through the calculation we got the minimum distance of spacing was 4.2 m. So the minimum distance which we set was 5 m. Every participant needed to take part in 10 groups. The participants drove with the speed of 60 km/h in the scenario and then picked up best one which gave them the pretty good visual effect by their feeling.

The second experiment was about the assessment between the new design and the real using urban tunnel. On the base of the result of the first experiment we concluded the most comfortable distance of spacing then we put it into the design plan. All the participants also needed to drive in the real urban tunnel. We used fuzzy elevation method to appraise the advantages between the design plan and the real condition.

3 Results

3.1 The Test of Spacing of the Design Pattern

The participants drove with the speed of 60 km/h and the height of the eyesight is 1.2 m. Table 3 demonstrated the results. For the wave pattern there was 35.8 % of the participant thought the 10 m distance brought the best visual effect. For the arrow pattern there was 33.9 % of the participant thought 20 m distance was the best. The difference of the pattern style and the dimension may contribute to the different results. For many participants the best spacing is 10 or 20 m and the 5 and 30 m were chosen less. Therefore we can conclude that the range of 10–20 m is the ideal one. In the design plane the spacing of the wave is 10 m and the spacing of the arrow is 20 m. Figures 10 and 11 display it.

Table 3 Proportional distribution (%)

Pattern type	Distance of the adjacent patterns (m)				
	5	10	15	20	30
Wave pattern (%)	7.1	35.8	33.9	16.1	7.1
Arrow pattern (%)	1.8	26.8	28.6	33.9	8.9

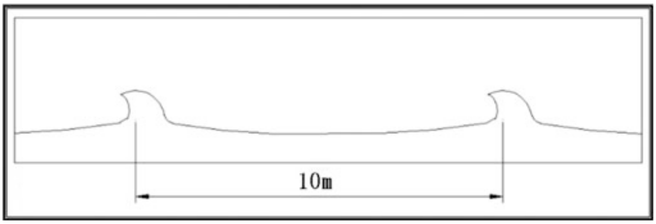


Fig. 10 The spacing of the waves

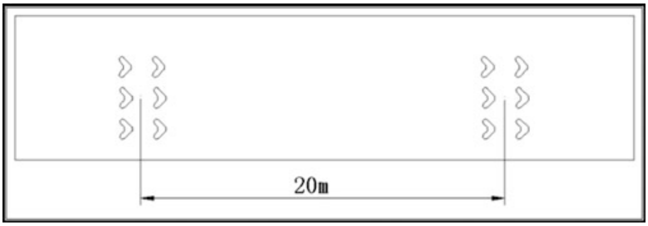


Fig. 11 The spacing of the arrow

3.2 The Contract Test Between the New Design and the Original Design

Drivers driving in the new design plan simulated environment and the real urban tunnel respectively. This study used fuzzy elevation method to appraise the advantages between the design plan and the real condition. The elements of the visual effect were divided into three parts which were safety and comfort and beauty. Safety contained the clarity of the driving environment and the speed sense. Comfort meant the relaxed and happy feelings and whether the drivers feel nervous and depressive. Beauty was mainly considered as the coordination (Figs. 12 and 13).

Taking these elements as the evaluation factors and forming the discussion range U

$$U = \begin{cases} \text{Safety}\,(U_1) \\ \text{Comfort}\,(U_2) \\ \text{Beauty}\,(U_3) \end{cases}$$

The participant appraised these three elements and the assessment was consisted of four levels—pretty good, good, acceptable and bad. They form the discussion range V

$$V = \{\text{Pretty good}\,(V_1),\ \text{Good}\,(V_2),\ \text{Acceptable}\,(V_3),\ \text{Bad}(V_4)\}$$

The result of the assessment is shown in Table 4.

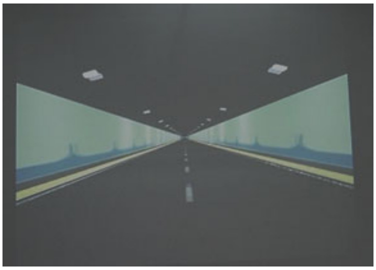


Fig. 12 The design plan

Fig. 13 The original design

Table 4 Proportional distribution (%)

Evaluation factor	Content	Pretty good (V_1) (%)	Good (V_2) (%)	Acceptable (V_3) (%)	Bad (V_4) (%)
Safety (U_1)	Design	71.9	21.9	6.2	0
	Original	0	0	18.8	81.2
Comfort (U_2)	Design	78.1	18.8	3.1	0
	Original	0	0	9.4	90.6
Beauty (U_3)	Design	75.0	21.9	3.1	0
	Original	0	0	0	100

From the table we gained two evaluation matrixes R_1 and $R_{2.}$ R_1 is for the design plan and R_2 is for the real urban tunnel.

$$R_1\begin{cases}(0.719, 0.219, 0.062, 0)\\(0.781, 0.188, 0.031, 0)\\(0.750, 0.219, 0.031, 0)\end{cases}\quad R_2\begin{cases}(0, 0, 0.188, 0.812)\\(0, 0, 0.940, 0.906)\\(0, 0, 0, 1)\end{cases}$$

To comprehensively display the participants' feelings, we need to set different weightings on the three evaluation factors. The weighting of the safety is 0.25 and the comfort's is 0.4 and the beauty's is 0.35. The weightings form the fuzzy vector A:

$$A = \{0.25,\ 0.40,\ 0.35\}$$

With the fuzzy vector A we can get the visual effect evaluation matrixes B_1 and B_2. B_1 is for the design plan and B_2 is for the real urban tunnel.

$$B_1 = A * R_1(0.25, 0.40, 0.35) * \begin{Bmatrix} 0.719 & 0.219 & 0.062 & 0 \\ 0.781 & 0.188 & 0.031 & 0 \\ 0.750 & 0.219 & 0.031 & 0 \end{Bmatrix}$$
$$= (0.75, 0.21, 0.04, 0)$$

$$B_1 = A * R_1(0.25, 0.40, 0.35) * \begin{Bmatrix} 0 & 0 & 0.188 & 0.812 \\ 0 & 0 & 0.094 & 0.906 \\ 0 & 0 & 0 & 1 \end{Bmatrix} = (0, 0, 0.08, 0.92)$$

Through the results of the assessment we found out that the new design plan of the urban tunnel was better than the real tunnel in all the aspects. The real tunnel which had no decoration on the side wall gave drivers undesirable visual environment. It had the effect on the traffic safety. For the new design plane the design pattern created the comfortable and beautiful environment. So the new design achieved the goals and it was very suit for the drivers.

4 Conclusion and Discussion

It had been proved that the new design patterns created comfortable visual environment of the tunnel for the drivers. So the new pattern design could be applied in the tunnel design for the purpose of improving the traffic safety of the urban tunnel. But this study still had some aspects which deserved to be discussed. Firstly we only made research on one type of design speed (60 km/h). In China the top speed limit of the urban tunnel is 80 km/h. So we could choose different design speeds to make more researches in the future. Secondly in the study we discussed two kinds of patterns which were simple and consistent. Actually there were many similar patterns. Therefore different types of patterns can apply to different conditions and researches of different patterns will need to be studied. Finally the color we used were blue and green because of their physiological influence. There are many kinds of color in the nature so the suitable combination of the color will be so many. Except the color the brightness of the color should be analyzed as well. In conclusion the new design pattern of the urban tunnel side wall has pretty good effect on the driving safety In general conclusion the design plane achieve the goal that improve the driving environment of the urban tunnels.

Acknowledgments This research was jointly sponsored by Project 51578417 supported by the National Natural Science Foundation of China and the Fundamental Research Funds for the Central Universities.

References

1. Sterling, R.: Underground technologies for livable cities. Tunn. Undergr. Space Technol. **12**, 479–490 (1997)
2. Hoeven, F.: A review of the impact of the RingRing study for multifunctional motorway tunnels in the Netherlands and beyond. Tunn. Undergr. Space Technol. **26**, 292–309 (2011)
3. Huang, X.: Traffic safety analysis on urban underpass tunnels. Southwest Jiaotong University (2013)
4. Song, Z.X., Pan, X.D.: Study on driving safety evaluation based on tunnel side wall effect. Highw. Eng. **35**, 10–15 (2010)
5. Zhuo, Y.C., Shao, F.: Driver's fixation characteristics during driving through urban tunnels. Traffic Inf. Saf. **6**, 101–106 (2014)
6. Manser, M.P., Hancock, P.A.: The influence of perceptual speed regulation on speed perception, choice, and control: tunnel wall characteristics and influences. Accid. Anal. Prev. **39**, 69–78 (2007)
7. Katja, K., Christer, A.: The impact of tunnel design and lighting on the performance of attentive and visually distracted drivers. Accid. Anal. Prev. **47**, 153–161 (2012)
8. Thiffault, P., Bergeron, J.: Monotony of road environment and driver fatigue: a simulator study. Accid. Anal. Prev. **35**(3), 381–391 (2003)
9. Ting, P.H., Hwang, J.R., Dong, J.L., Jeng, M.C.: Driver fatigue and highway driving: a simulator study. Physiol. Behav. **94**(3), 448–453 (2008)
10. Caliendoa, C., Guglielmoa, M.L.D.: Accident rates in road tunnels and social cost evaluation. Proc. Soc. Behav. Sci. **53**, 166–177 (2012)
11. Nussbaumer, C.: Comparative analysis of safety in tunnels. In: Young Researchers Seminar, Brno (2007)
12. Liu, B.: Mechanism Study on the Vision-Based Speed Control and Lane Keeping. Wuhan University of Technology (2008)
13. Mandatory National Standard of the People's Republic of China GB 5768.2-2009, Road traffic signs and markings: Road traffic signs and markings—Part 2: Road traffic signs (2009)
14. Ye, H.F.: The Technology of Speed Control for the Ordinary Highway. Southeast University (2009)
15. Wu, Q., Pan, X.D.: Research on driving safety experiment of tunnel based on tunnel sidewall effect. Highw. Eng. **35**, 11–103 (2013)
16. Hu, Y.Z.: On the Side Wall Decoration Design of Urban Tunnel Chinese and Overseas Architecture, vol. 12, pp. 99–100 (2015)

Research on Drivers' Recognition Ability of Obstacles Under Lighting Environment of Tunnel Entrance Sections

Meng Xu, Zewen Yu, Xiaodong Pan, Zhiguang Xia, Feng Chen, Shaoshuai Li and Gang Wu

Abstract The tunnel entrance section is where the visual environment changes abruptly and influences the driver's judgment of the traffic situation in front. The lighting design of tunnel entrance sections currently usually takes the same brightness standards at the entrance and the entrance section. It makes drivers may not find the obstacles at a distance from the entrance, which is extremely detrimental to traffic safety. In this paper, a reasonable obstacle-recognition experiment under driving situation is designed to study the influence of visual environment (mainly the lighting environment) on visual adaptability in the entrance section of the tunnels and analyze the relationship between the visual stimulation caused by the sharp change of luminance and drivers' recognition ability of obstacles. The analysis of the obstacle-recognition experiment results provides basis for further study on improvement of visual environment of tunnel sections, and is of great significance to improving traffic safety of tunnels.

Keywords Tunnel entrance section · Obstacle recognition · Visual adaptability · Lighting environment

M. Xu · X. Pan · Z. Xia · F. Chen (✉) · G. Wu
Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai 201804, China
e-mail: chenfeng.csu@gmail.com

Z. Yu
Hunan Communications Research Institute, Changsha 410015, Hunan, China

S. Li
Shanghai Municipal Engineering Design Institute (Group) Co., Ltd., 901 North Zhongshan Road (2nd), Yangpu District, Shanghai, China

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_59

1 Introduction

Tunnel sections are usually accident prone locations on highway. While vehicles running in highway tunnel section, the significant luminance difference between inside and outside the tunnel environment can cause abrupt changes of visual environment that leads to the decrease of drivers' visual ability and makes drivers cannot make perception and judgment of the distance, speed, obstacles accurately and in time and induces road traffic accident.

The current design standard of tunnel lighting brightness level takes the same brightness values at the entrance and the entrance section. In practice, under the influence of natural light outside, the brightness inside the tunnel decline gradually according to certain rules from the entrance during the day and has a sudden change at a certain distance from the entrance. In accordance with the design method of fixed brightness value, the brightness at the tunnel entrance is high enough for drivers to recognize obstacles at the entrance at required distance. But if the obstacles appears at a distance from the entrance, because of sudden change of the light conditions, there is no guarantee that drivers can found obstacles within the scope of enough stopping sight distance, that will be adverse to driving safety.

Therefore, in this article, the obstacle recognition experiment under reasonable driving conditions is designed to analyze the influence of the visual environment (mainly lighting environment) of the tunnel entrance section on drivers' visual adaptation and discuss the relationship between the tunnel entrance brightness changes and the driver's obstacles recognition ability. Standard obstacles is selected to test the obstacle recognition distance of several testers under different driving speed through driving experiment. The relationship between the light illumination conditions and the obstacle recognition distance is studied from the point of view of traffic safety to provide basis provide for further study on improvement of visual environment of tunnel sections, especially lighting environment.

2 Literature Review

The light conditions of the tunnel entrance section is the basis for tunnel lighting requirements of other parts, and the driver's visual feature is unique at the tunnel entrance. Therefore, it is very important to study the tunnel light environment for driving safety. The Netherlands Organization for Applied Scientific Research pointed out that the various elements of the tunnel design, such as the length and width of the tunnel, the number and the curvature of the tunnel route curve and the lighting conditions and so on, will affect driving behavior and traffic safety [1]. French scholar Bourdy et al. [2, 3] used simulation methods to simulate the brightness variable environment of the tunnel entrance and obtained the driver's static visual threshold and dynamic visual threshold under the visual adaptation condition. Adrian [4, 5] did researches on the adaptation brightness and the required brightness of the driver close to the tunnel entrance. Narisada [6] made records point distribution and the fixation time of drivers close to the tunnel entrance and put forward the demand brightness.

Whether the lighting environment of the tunnel entrance section can meet the driver's obstacle recognition requirement is one of the important indices to evaluate traffic safety of the tunnel entrance section. Scholars have did some relevant research on the driver's ability to identify obstacles. French scholar Mayeur et al. [7, 8] indicated that the complexity of the background had significantly influence on the visibility of the target object. The required distance for drivers to recognize obstacles was measured through real vehicle experiment and the suggested values of visibility level were given. Brémond [9] set up a nighttime highway environment simulation test scenarios and simulate obstacles of different sizes, such as vehicles, pedestrians, signs. Japanese scholars Ekrias et al. [10] utilized digital camera and set up the test environment with vehicle lamp lighting and road lighting. The visibility of obstacles under the environment was evaluated. Bacelar [10] studied the effect of the existence of road lighting on the obstacle visibility.

There have been some researches focusing on the tunnel entrance section illumination environment and the driver's visual ability, but there is few scholars pay attention to the effect of light environment at the tunnel entrance and the entrance section on drivers' obstacles recognition. Based on the research methods and results above, a reasonable obstacle-recognition experiment under driving situation is designed in this paper, to analyze the relationship between the visual stimulation caused by the sharp change of luminance and drivers' recognition ability of obstacles. Through site measurements and driving experiment results, the influence of natural light outside the tunnels, driving speed and illumination change on obstacle recognition have been studied.

3 The Experimental Research

3.1 The Experiment Site

The experiment site was an unopened highway, which provided necessary conditions for driving experiments. There were 4 tunnels in this section of the highway, that were all double hole design, with total length of 2103, 455, 1643, 1298 m. The experiment tunnels were designed by 80 km/h. The side wall of the tunnels was bonded with yellow brick within 2.5 m height, with reflectivity of about 50 %. The reflection coefficient of the arch part was about 25 %. The tunnel pavements were all asphalt pavements (Fig. 1).

Fig. 1 The lighting environment of the experiment tunnel

3.2 *The Experimental Personnel and Equipments*

This experiment needed operating staff of two, respectively responsible for the obstacles' location and illumination measurement inside and outside the tunnel; driver of one; testers of eight.

This experiment instruments included a luxmeter, a non-contact speedometer of the motor vehicle and an experimental vehicle of a Mazda PREMACY car.

Most of the reflection coefficient of objects on the road is 3–30 %, as shown in Fig. 2 [12]. The choice of the obstacles should guarantee that the reflection coefficient was typical and representative. Tunnel environment was a relatively closed environment with a high maintenance coefficient. The possible obstacles in tunnels were usually objects falling from vehicle, such as cartons, plastic foam and so on, with reflection coefficient of 20–35 %. Refer to Fig. 2, the standard obstacle of the experiment was set as size of 20 cm × 20 cm × 20 cm and reflection coefficient of 0.3, as shown in Fig. 3.

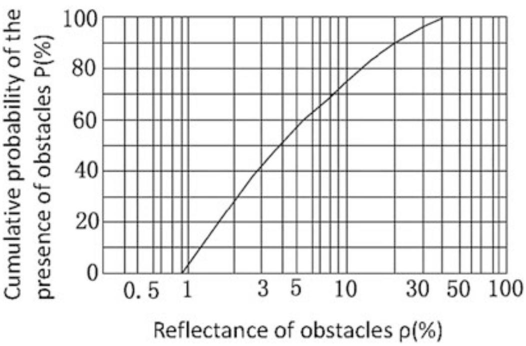


Fig. 2 The cumulative probability of the presence of obstacles on road [12]

Fig. 3 The standard obstacle of the experiment

3.3 The Experiment Content

The tester's recognition distance of the standard obstacle set at a certain distance from the tunnel entrance was recorded by the non-contact speedometer of the motor vehicle (CTM-8C) equipped on the experimental car, under the present lighting condition and a certain driving speed. The illuminance values of three places (outside the tunnel, the location where the obstacle was set and the surface of the obstacle) were measured by the luxmeter meanwhile every time the driving experiment was repeated.

4 Results and Discussion

4.1 The Illuminance Change of the Tunnel Entrance Section Under the Influence of Natural Light

Affected by the natural light outside the tunnel, the illuminance inside within a distance from the entrance is actually close to it outside. With the distance increases over a certain value, the illuminance inside the tunnel decreases immediately, even a sudden change may occur. A sudden decrease in illumination in short time may cause temporary visual disturbance, which will affect driving safety.

The illuminance changes of the tunnel entrance section under the circumstance of daytime of sunny day and cloudy day with the tunnel lighting facilities in normal running was measured to analyze the illuminance change rule within a distance of the tunnel entrance section under the influence of nature light. The measured data is processed and the fitting results are shown in Figs. 4 and 5.

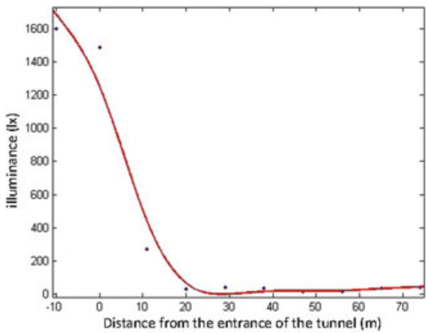


Fig. 4 Illuminance changes of sunny day

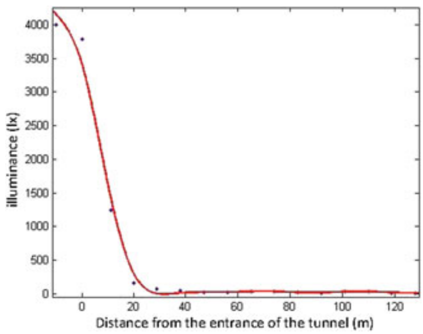


Fig. 5 Illuminance changes of cloudy day

As shown in Figs. 4 and 5, it can be found that the illuminance changes very quickly within the scope of a certain distance from the tunnel entrance, from thousands of lux to dozens of lux within dozens of meters. In the sunny day, the illuminance of the entrance is 94.5 % of the illuminance outside the tunnel while the illuminance 20 m from the entrance decreases to 4.15 % of the illuminance outside. In the cloudy day, the illuminance of the entrance is 92.9 % of the illuminance outside the tunnel while the illuminance 20 m from the entrance decreases to 1.88 % of the illuminance outside. The illuminance sudden change is of high decrease rate within the scope of 20 m from the tunnel entrance that has adverse effect to traffic safety.

It can be concluded that the effect of natural light on the tunnel illuminance is very obvious and has strong regularity: the tunnel lighting designed according to the present design methods can meet the driving visual demand well at the entrance, but the rapid variation of the illuminance within the scope of a certain distance (the scope of the experimental tunnel is 10–30 m) has bad effect on the driver's obstacle recognition. Therefore, the lighting of this area should be appropriately strengthened from the prospect of traffic safety. The former research had also proved it well [13]: a lot of tunnel driving experiment results showed that the driving visual load is the largest within scope of a distance from the tunnel entrance evaluated by the pupil area changing speed, and the lighting of the area should be strengthened.

4.2 Relationship Between Obstacle Recognition Distance and Driving Speed

According to the preliminary analysis of the experimental results, there is some relationship between the obstacle recognition distance and the vehicle speed. Effective mathematical methods are selected to analyze function relation or the relative relationship between the experiment data of the speed and the obstacle recognition distance. The results of recognition distances of different speed are shown in Fig. 6.

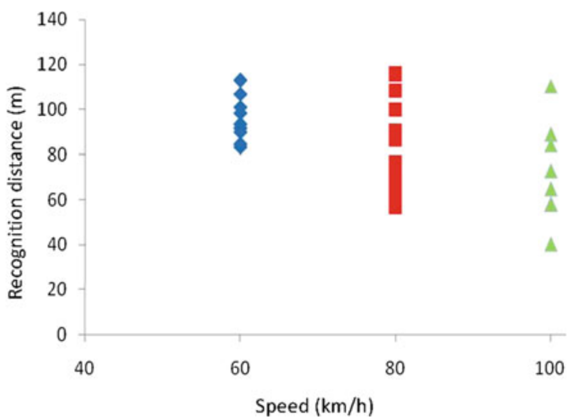
The results show that the obstacle recognition distances have discreteness and the main reasons are:

1. The psychological physiological differences of testers lead to different experimental results even if under the same objective conditions.
2. The illuminance of the environment changes constantly in each experiment, which is one of the important reasons for the existence of the data discreteness.

4.2.1 Discreteness Analysis

The two figures indicate that the obstacle recognition distances within 95 % CI are in a reasonable range and the dispersion of the data is acceptable. However, the median, extreme value, standard deviation, and confidence interval of the obstacle recognition distance data of 80, 60, 100 km/h varies a lot while the basic trend of recognition distances of different speeds is 60 km/h > 80 km/h > 100 km/h. Therefore, it is very necessary to make correlation analysis on the driving speed and the obstacle recognition distance (Figs. 7 and 8).

Fig. 6 Obstacle recognition distances of different driving speed



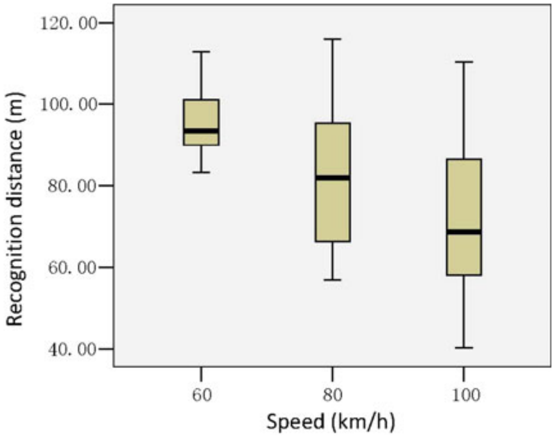


Fig. 7 Box plot of obstacle recognition distances under different speeds

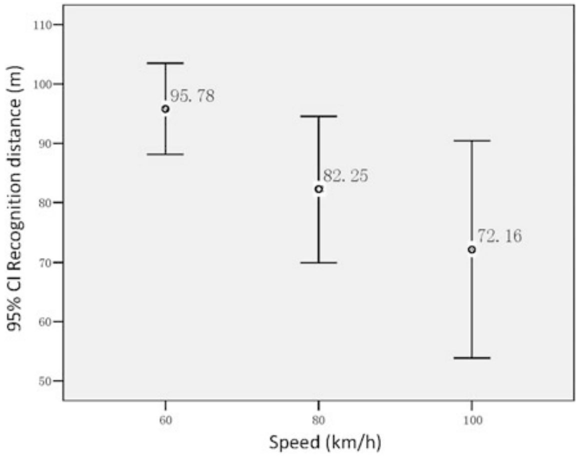


Fig. 8 95 % confidence intervals chart of obstacle recognition distances under different speeds

4.2.2 Correlation Analysis

Pearson Correlation Analysis. Pearson correlation analysis is used to analyze the relationship between continuous data. The computation formula is:

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (-1 \leq r \leq 1) \tag{1}$$

r can be used as the estimated value of the correlation coefficient. $r < 0$ indicates negatively correlated, $r < 0$ indicates positively related and $r = 0$ indicates irrelevant. However, due to the existence of sampling error, hypothesis test should be carried out.

Table 1 Pearson correlation analysis results

		Speed	Recognition distance
Speed	Pearson correlation	1	−0.473*
	Sig. (1-tailed)		0.005
	N	29	29
Recognition distance	Pearson correlation	−0.473*	1
	Sig. (1-tailed)	0.005	
	N	29	29

*Correlation is significant at the 0.01 level (1-tailed)

The SPSS software is used to do the Pearson correlation analysis of driving speed and obstacle recognition distance and the results are shown in Table 1.

The results show that the correlation coefficient is −0.473 and the significant test result is 0.005, which is much less than 0.05, that shows that the correlation between speed and recognition distance is highly significant.

Partial Correlation Analysis. The partial correlation analysis refers to the statistical method that while studying the relationship between two variables, the other variables associated with the two variables are controlled unchanged. The illuminance of outside the tunnel, inside the tunnel and obstacle surface are taken as the control variables. The SPSS software is used to do the partial correlation analysis of driving speed and obstacle recognition distance and the results are shown in Table 2.

The results show that the correlation coefficient is −0.465 and the significant test result is 0.017, which is much less than 0.05, that shows that the correlation between speed and recognition distance is highly significant.

According to the analyses above, driving speed is an important factor affecting drivers' obstacle recognition distance and the basic trend of recognition distances of different speeds is 60 km/h > 80 km/h > 100 km/h.

Table 2 Partial correlation analysis results

Control variables			Recognition distance	Speed
Illuminance of obstacle and illuminance inside the tunnels and illuminance outside the tunnel	Recognition distance	Correlation	1.000	−0.465
		Significance (2-tailed)		0.017
		df		24
	Speed	Correlation	−0.465	1.000
		Significance (2-tailed)	0.017	
		df	24	0

4.3 Relationship Between the Obstacle Recognition Distance and the Change of Illuminance

According to the preliminary analysis results, the correlation between the obstacle recognition distance and a single type illuminance is not obvious. The reason is mainly related to the dark adaptation stage in the process of recognition that the illuminance of the tunnel environment is constantly changing.

Therefore, in this paper, we put forward the change rate of the illuminance from outside tunnel to the obstacle surface to indicate the driver's visual adaptation of characterization. Due to the difference of orders of magnitude between the illuminance of the tunnel and the obstacle surface illuminance is bigger, certain mathematical methods are used to eliminate error. The illuminance change rate v is defined:

$$v = \frac{\left|\lg E_{out} - \lg E_{object}\right|}{\lg E_{out}} \tag{2}$$

E_{out} is the illuminance outside the tunnel, E_{object} is the obstacle surface illuminance. According to the above analysis, the illuminance decreases with entering the tunnel while the illuminance change rate increases. The relationship between the illuminance variation rate and the distance from the tunnel entrance is as shown in Fig. 9.

The measured data of the obstacle recognition distance and the change rate of illuminance is processed and the fitting results are shown in Fig. 10.

The following results can be concluded from Fig. 10:

1. Under different speed conditions, when the change rate of illuminance is in a certain range, the obstacle recognition distance is in a lower range compared to other regions. When the speed is 60 km/h, the change rate of illuminance is

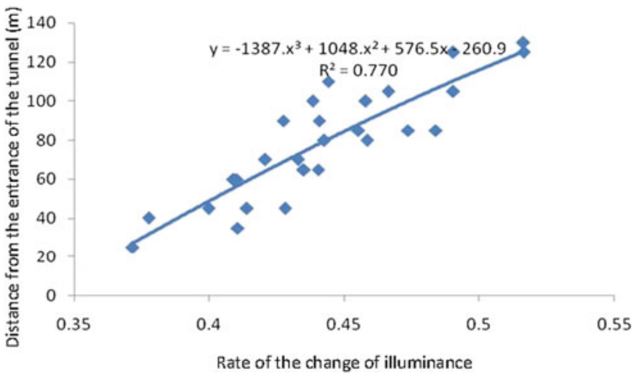


Fig. 9 Relationship between the distance from the tunnel entrance and the change of illuminance

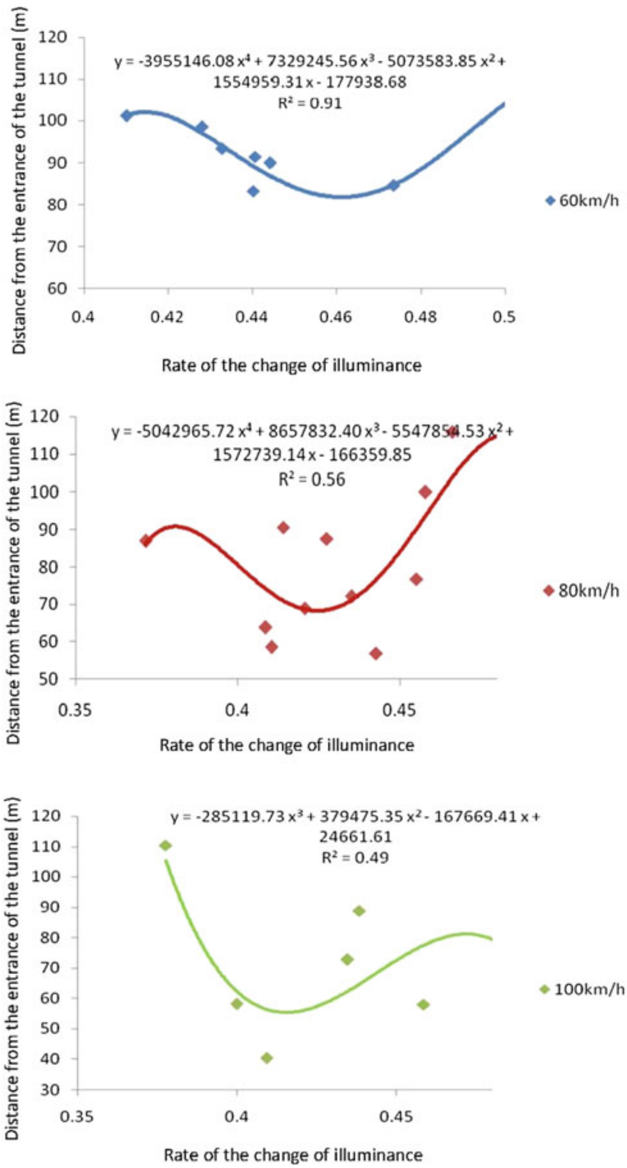


Fig. 10 Relationship between the obstacle recognition distance and the change of illuminance under driving speed of 60, 80, 100 km/h

0.4612, the obstacle recognition distance reaches the lowest point 82.4 m; 80 km/h with 0.4246 and 68.2 m and 100 km/h with 0.4145 and 56.0 m. The design speed of the experimental tunnel is 80 km/h and the longitudinal slope is −3 %. The lowest recognition distance is 68.2 m, which cannot meet the requirements of the parking sight distance 110 m [14] and may lead to traffic accidents.

2. According to the fitting formula in Fig. 9, the places where the distance of obstacle recognition is shortest under different driving speed: 91.8 m from the entrance of the tunnel when the speed is 60 km/h, 66.6 m from the entrance of the tunnel when the speed is 80 km/h, 59.3 m from the entrance of the tunnel when the speed is 100 km/h.
3. The obstacle recognition distance does not show a decreasing trend with the increase of the intensity of illuminance. When reaching a certain range, it presents the opposite result. The reason is mainly related to the dark adaptation stage in the process of obstacle recognition. The driver can recognize the obstacle deep inside the tunnel from a long distance because of that the driver has been in the tunnel for a while and dark adaptation process has completed or partially completed.

5 Conclusion

In this paper, focusing on the existing defects in the lighting design of the tunnel entrance section, a typical obstacle recognition experiment is carried out. According to the detailed analysis of the experimental data, the main conclusions are as follows:

1. The effect of natural light on the tunnel illuminance is very obvious and has strong regularity: the rapid variation of the illuminance within the scope of a certain distance (the scope of the experimental tunnel is 10–30 m) has bad effect on the driver's obstacle recognition. Therefore, the lighting of this area should be appropriately strengthened from the prospect of traffic safety.
2. Driving speed is an important factor affecting drivers' obstacle recognition distance and the basic trend of recognition distances of different speeds is 60 km/h > 80 km/h > 100 km/h.
3. Under different speed conditions, when the change rate of illuminance is in a certain range, the obstacle recognition distance is in a lower range compared to other regions. When the speed is 60 km/h, the change rate of illuminance is 0.4612, the obstacle recognition distance is to reach the lowest point 82.4 m, 80 km/h with 0.4246 and 68.2 m and 100 km/h with 0.4145 and 56.0 m.

The analysis of the obstacle-recognition experiment results provides basis for further study on improvement of visual environment of tunnel sections, especially lighting environment, and is of great significance to reducing traffic accidents due to vision environment mutations and improving traffic safety of tunnels.

Acknowledgments This research was jointly sponsored by Project 51578417 supported by the NSFC and Project 201519 supported by Department of Transportation of Hunan Province.

References

1. Baart, M., Driever, H.: Tunnel safety. TNO Traffic and Transport (2002)
2. Bourdy, C., Chiron, A., Cottin, C., Monot, A.: Visibility at a tunnel entrance: effect of temporal adaptation. Light. Res. Technol. **19**, 35 (1987)
3. Bourdy, C., Chiron, A., Cottin, C., Monot, A.: Visibility at a tunnel entrance: effect of temporal luminance variation. Light. Res. Technol. **20**(4), 199 (1988)
4. Adrian, W.: Investigations on the required luminance in tunnel entrances. Light. Res. Technol. **14**, 151 (1982)
5. Adrian, W.K.: Adaptation luminance when approaching a tunnel in daytime. Light. Res. Technol. **19**, 73 (1987)
6. Narisada, K., Yoseoikawa, K.: Tunnel entrance lighting—effect of fixation point and other factors on the determination of requirements. Light. Res. Technol. **6**(9), 9–18 (1974)
7. Mayeur, A., Brémond, R., Bastien, J.M.C.: Effects of the viewing context on target detection. Implications for road lighting design. Appl. Ergonomics **41**(3), 461–468 (2010)
8. Mayeur, A., Brémond, R., Bastien, J.M.C.: The effect of the driving activity on target detection as a function of the visibility level: implications for road lighting. Transp. Res. Part F Traffic Psychol. Behav. **13**(2), 115–128 (2010)
9. Brémond, R., Bodard, V., Dumont, E., Nouailles-Mayeur, A.: Target visibility level and detection distance on a driving simulator. Light. Res. Technol. **45**(1), 76 (2013)
10. Bacelar, A.: The contribution of vehicle lights in urban and peripheral urban environments. Light. Res. Technol. **36**, 69 (2004)
11. Lecocq, J.: Calculation of the visibility level of spherical targets in roads. Light. Res. Technol. **31**(4), 171–175 (1999)
12. Kawakami, G.: Proceedings of Road Lighting, vol. 151 (1982)
13. Pan, X., Song, Y., Yang, Z., Zhang, G.: Visual environment improving scope at entrance and exit of highway tunnel based on visual load. J. Tongji Univ. (Nat. Sci.) **37**(6) (2009)
14. Ministry of Transport of the People's Republic of China: Guidelines for Design of Lighting of Highway Tunnels. JTG/T D70/2-01-2014 (2014)

Research on Optimized Design of Bridge-Tunnel Linkage Sections on Mountainous Highways Considering the Effect of Crosswind on Driver Behavior

Deshan Feng, Kai Wu, Feng Chen, Xiaodong Pan, Meng Xu, Gang Wu and Zhiguang Xia

Abstract This research investigates how different shoulder widths affect driver behavior and driver's heart rate, on which the optimized design of bridge-tunnel linkage sections is based. We conduct a simulation driving experiment on an 8-degrees of freedom simulator which creates a high fidelity virtual and motive driving environment. In this study, the max countering steering wheel angle (MCSWA) is chosen to characterize the safety degree of driver's manipulation while the variability of max heart rate (VMHB) is used to describe driver's mental and physiological state. After an overall consideration of relations among shoulder width, MCSWA and VMHB, the most suitable shoulder width and the optimized design of bridge-tunnel linkage sections are proposed.

Keywords Bridge-tunnel linkage sections · Simulation driving · Driver behavior · Heart rate · Optimized design

1 Introduction

Nowadays, the proportion of bridges and tunnels in mountainous highways has increased dramatically. For some highways in western China, alternative bridges and tunnels occupy at most 40 % of their length. Strong crosswind is very common in mountainous area and it is even more powerful on high-pier bridges. However, in contrast, there is little crosswind in tunnels. The change of crosswind may make it

---

D. Feng · F. Chen (✉) · X. Pan · M. Xu · G. Wu · Z. Xia
Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai 201804, China
e-mail: chenfeng.csu@gmail.com

K. Wu
Imperial College London, London SW7 2AZ, UK

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_60

more difficult for drivers to control the vehicles. Under normal conditions, drivers tend to steer oppositely to counteract the effect of sudden change of crosswind. However, some of them would probably fail to respond or overact and accidents may happen. To enhance the overall safety of mountainous highway, the importance of investigating driver behavior under the effect of crosswind at the bridge-tunnel linkage sections is apparent.

Various studies have been done to investigate how crosswind affects the driving safety from different aspects. The research methods they took mainly included numerical simulation [1–4], wind tunnel test [5, 6], simulation driving [7, 8] and field test [9, 10]. Based on these methods, different kinds of conclusions were made, including risk management [2, 5, 11], lateral displacement [3, 8–10, 12, 13], heading error [4, 5, 8–10], steering angle [3, 8, 10, 14], etc. These findings to a certain extent explained some phenomenon and meanwhile gave proposals to improve driving safety, mainly including vehicles' speed limitation, decision-making of expressway closure, steering assistance, etc. Among all research methods, the numerical simulation was most frequently used because of its simplicity and safety. However, numerical simulation ignores drivers' reaction to crosswind and this made the findings based on this method less convincing. More researchers are now adopting the simulation driving method which enables drivers feel the crosswind more authentically. In this paper, simulation driving is based on an 8-degrees of freedom simulator, creating a high fidelity virtual and motive driving environment.

Bridge-tunnel linkage sections are typical and peculiar from the perspective of diving safety not only for the frequent changing crosswind, but also for the intricate light environment near tunnel's entrance and exit and the feeling of suppression caused by tunnel. According to one survey, among all the accidents on bridge-tunnels linkage sections of mountainous highway, 25 % of them are caused by the mismatch between vehicles' deviation caused by crosswind and transitional cross section design of bridges and tunnels. Cross section design is vital for driving safety, especially the shoulder width. Ben-Bassat, T. and Shinar, D. studied the shoulder width on drivers' perception and behavior through simulation driving [15]. It is a pity that road design regulations in China haven't given enough concern to cross sectional design of bridge-tunnel linkage sections from the view of crosswind and the existing rules are relatively general and ambiguous. In order to improve driving safety, this paper proposes a distinctive angle of optimized design of bridge-tunnel linkage sections, which is different from the suggestions above.

This study primarily investigates how different left shoulder widths of tunnel' exit and entrance affect driver behavior considering the effect of crosswind at bridge-tunnel linkage sections. To make the conclusions more convincing and effective, we conduct simulation driving experiments based on an 8-degrees of freedom simulator to make drivers feel the crosswind as real as possible. The max countering steering wheel angle (MCSWA) when crosswind changes suddenly is selected to characterize the safety degree of driver behavior. The shoulder width is chosen as the main factor of cross section design that we assume affects the driving safety. Then, the relationship of MCSWA and left shoulder width at tunnel entrance and exit is established. According to the critical state equation, the most suitable shoulder width is calculated. In this research, to get a better insight about driver behavior, we also investigate drivers' heart rate which characterizes the mental and physiological state. At last, an optimized design of bridge-tunnel linkage sections is provided for a mountainous highway.

2 Method

2.1 Effect of Shoulder Width on Driving Safety

The distance between the outside wheel and guardrail is an important index which directly reflects the drivers' demand of side space. If the side space is too small, drivers tend to steer oppositely instinctively, which is dangerous if the vehicle is at a high speed and also imposed with strong lateral force. That is exactly the condition that drivers may confront at the bridge-tunnel linkage sections. In addition, when vehicles are at tunnel's entrance or exit, drivers' physiological and mental load may raise because of strong crosswind, which on the other hand decrease the accuracy and stability of manipulation. From the perspective of tunnel design, the shoulder width is the key factor that we should focus. To investigate how should width affects driver behavior is extremely important for the optimized design of bridge-tunnel sections, which we ultimately hope can provide a safe and comfortable driving environment.

2.2 Driving Simulator

This study is based on Traffic Behavior and Cooperative Virtual Reality System of Tongji University. This driving simulator is featured with 2.5 tons of capacity and 8 degrees of freedom motion (XY 5×20 m + 6 DOF motion) which can perfectly simulate the driving feeling like acceleration, deceleration. The cockpit is a sealed rigid structure with the minimum radius of 5.5 m and the vehicle used for driving is right in the center of the cockpit. The vehicle used for this system is Renault Megane III However, the engine was taken out Other facilities which are used for data collection of steering wheel, brake, gear, etc. are added.

Five projectors on top of the vehicle creates a broad screen projective system with the horizontal angle of 250° and the vertical angle of 40°. The resolution ratio of each projector is 1400×1050 and the refresh frequency is 60 Hz. Three rear-view mirrors are LCD screens. In addition, the high quality sound system in the cockpit provides a real acoustic environment with the sound of tyres, engine, brake, etc. The control software for the driving simulator is SCANeR Studio developed by

OKTAL Corp, which has functions of integrated control, scenario design, data acquisition, data analysis, etc.

The dynamic system and other supplementary devices above together create an exceedingly real driving environment. However, it is possible that you may feel a little bit dizzy if you stay long in the cockpit. So researchers have to take this into consideration when designing experiments.

2.3 Experiment Design

The cross section design of simulation driving scenario is based on a highway in western China. The length of whole scenario is 15,000 m long straight highway with alternative tunnels and bridges. Each section of tunnels and bridges is 450 m long. Before the first tunnel, there is a 600 m long ordinary road set for drivers to get accustomed for the experiment. In this research, the shoulder width is the main factor of cross section design that we assume affects driver behavior and driving safety.

As is illustrated in Fig. 1, taking the shoulder width into account, we design a new transitional form of cross section at bridge-tunnel linkage sections. The original shoulder width is 0.75 m for both bridges and tunnels. In this experiment, left shoulder width of tunnels near the exit increases from 0.75 m (inside tunnel) to 0.75 + Xm (tunnel exit) at with the variance ratio of 1/25, which is recommended by road design regulations in China. Here the value of X varies from 1.75 m to 0. It is the same at tunnel's entrance. During the whole process, the driver would encounter different combinations of bridge and tunnel, whose shoulder widths vary at the linkage sections.

This research focuses on how different left shoulder width affects driver behavior considering the effect of crosswind. So the crosswind speed is only set 60 km/h to simulate the worst situation, which also corresponds to the real wind level on high-pier bridges on mountainous highways. As is shown in Fig. 2, the whole process is consisted of two parts. The crosswind come from the right side in the first

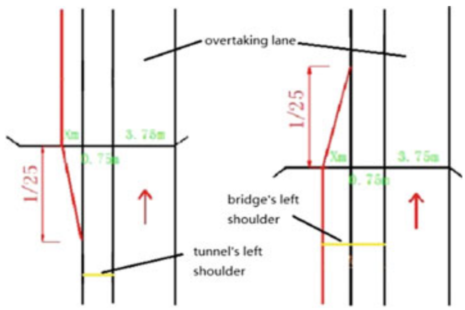


Fig. 1 Transitional form of cross-sectional design at tunnel's entrance and exit

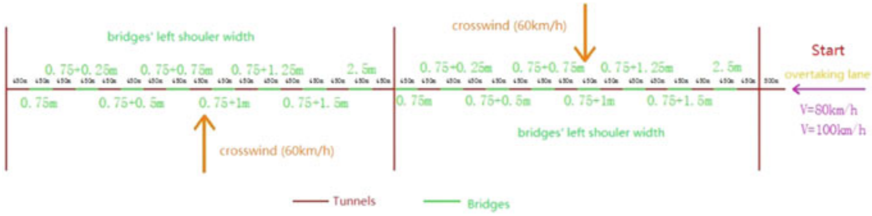


Fig. 2 The whole process of the simulation driving

part while coming from the left part in the second part. The direction of crosswind is always vertical to the road.

Ten drivers aged between 23 and 44 and consisting seven male and three female were recruited from Tongji University. They were all in good health and had driver's license. Four of them are novice drivers and the rest of them are skilled at driving. The reason why nearly half of the recruiters are freshmen is that many accidents are caused by the improper manipulation of novice drivers. It is meaningful to prove into the driving behavior of freshmen, so the ratio of freshmen was intentionally amplified in this study.

2.4 Experiment Procedure

Each driver had to drive in the scenario for twice, once is with the speed of 80 km/h and another is 100 km/h. There are several steps to follow before and during the experiment.

First of all, to help drivers get adapted to the driving simulator, they were asked to drive for about 10 min. On the hand, drivers knew the force feedback of the gear, accelerator, brake and steering wheel better. On the other hand, this helped the drivers to adapt to the virtual scenario and found out if any recruiter felt very uncomfortable in the cockpit.

The second step was placement and debugging of relative experiment instruments, including SmartEye EMR, Polar Heart Rate Watch and Electroencephalograph. These instruments were used to monitor the physiological state during the whole process and some of them may be chosen to analyze interior reason for their driving behavior. There were two points worth attention. The first was to make sure these instruments didn't influence the driving. The second was make sure these instruments were on the same time domain, which would be very helpful to the later data analysis.

The third step was driving process. The drivers were told that there would be crosswind during the experiment. However, they didn't know how strong the crosswind was. All they had to do was to keep the car in the overtaking lane. Each driver drove once under the speed limit of 80 and 100 km/h respectively.

The forth step was a short questionnaire. The drivers were asked how they felt about the driving simulator, whether they ever confronted this kind of crosswind situation, whether the scenario was real enough, etc. These questions were very important for later analysis.

3 Data Analysis

3.1 Max Countering Steering Wheel Angle (MCSWA)

The actions that drivers take to counteract the effect of crosswind at tunnel's entrance and exit are complicated, including acceleration, deceleration, steering, etc. The foremost one we think is steering because it largely affects the stability of vehicles. Usually, the steering angle fluctuates around 0° in tunnels. At tunnel's exit, the crosswind exerts strong force on vehicles, which results a sudden whirl of steering wheel. Instinctively, drivers will whirl the steering wheel reversely to pull the vehicle back to the lane. In most conditions, drivers can handle it well. However, for some who seldom experienced this situation, it would be different. Some of them fail to react while the others overreact. It is the same at tunnel's entrance.

During the experiment, we find that most of them took actions in time and the primary difference is the degree they whirl the steering wheel. In previous researches, the steering wheel angle, the steering wheel force feedback, the steering wheel acceleration, etc. were proposed to characterize the action of steering and each index has its own features. In this paper, we choose the max countering steering wheel angle (abbreviated as MCSWA below) as the main index indicating the safety of driver behavior. It should be noted that MCSWA is usually larger than the proper steering angle. So drivers have to adjust the steering angle if it reaches its maximum because they realize they have overreacted.

In this experiment, the crosswind speed is constant and left shoulder width changes. Different values of should width have different effects on driver behavior, which can be observed from MCSWA. If it is too large, we can presume that the corresponding left shoulder width is inappropriate from the perspective of driving safety. So MCSWA is chosen to characterize the safety degree of driver behavior. Based on this assumption, the data analysis of MCSWA is illustrated below.

Tunnel' Exits. At tunnel's exit, drivers experience the process from feeble crosswind to strong crosswind. The simulation driving is conducted under the speed limit of 80 and 100 km/h. Figure 3 shows the distribution of MCSWA of ten drivers considering different shoulder width at tunnels' exits.

Then we use quadric curve fitting and cubic spline fitting for average MCSWA ($\bar{y}$) and left shoulder width (x).

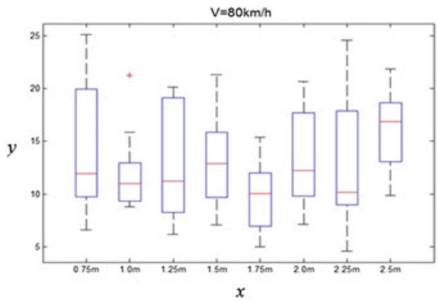


Fig. 3 The distribution of MCSWA of ten drivers considering different shoulder widths at tunnel's entrance (80 km/h). x refers to the left shoulder width and y refers to MCSWA

(1) The speed limit of 80 km/h—

Quadric curve fitting:

$$\begin{aligned} \bar{y} &= 4.073x^2 - 12.39x + 20.99 \\ R^2 &= 0.6307 < 0.8,\ \text{SSE} = 7.475 \end{aligned} \tag{1}$$

Cubic spline fitting:

$$\begin{aligned} \bar{y} &= 3.2x^3 - 11.53x^2 + 11.11x + 10.26 \\ R^2 &= 0.704 < 0.8,\ \text{SSE} = 5.99 \end{aligned} \tag{2}$$

The effect of cubic spline fitting, not ideal though, is better than quadric curve fitting. By solving the Eq. (2), the value of $\bar{y}$ reaches the minimum 11.5° when x equals 1.735 m. We presume that the most suitable shoulder width for driving safety is 1.735 m.

Under the speed limit of 80 km/h, it can be seen from Fig. 4 that MCSWA doesn't always increases as the shoulder width decreases. If the shoulder width is too small or too large, drivers tend to react more fiercely. The reason for this phenomenon may be that drivers would not feel much tension if the shoulder width is large, which offers rich lateral space. However, when the vehicle deviates from

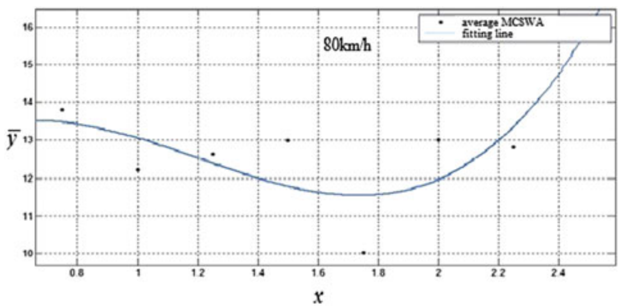


Fig. 4 Cubic spline fitting for average MCSWA ($\bar{y}$) and left shoulder width (x) (80 km/h)

the lane too much, drivers would whirl the steering wheel more to pull the vehicle back to the lane. This condition is different from what we discuss above that drivers would feel nervous if the lateral clearance is too restricted.

(2) The speed limit of 100 km/h—

Cubic spline fitting:

$$\begin{aligned} \bar{y} &= -3.472x^3 + 22.8x^2 - 43.16x + 35.3 \\ R^2 &= 0.8365 > 0.8,\ \text{SSE} = 5.805 \end{aligned} \tag{3}$$

The effect of cubic spline fitting is relatively good. By solving the Eq. (3), the value of $\bar{y}$ reaches the minimum 10° when x equals 1.384 m. We presume that the most suitable shoulder width for driving safety is 1.384 m (Fig. 5).

Under the speed limit of 100 km/h, the changing behavior of MCSWA resembles what it is under the speed limit of 80 km/h. However, the difference is that the most suitable shoulder width of 80 km/h is larger than that of 100 km/h, which means the higher the vehicle speed, the more lateral clearance needed. Considering that the speeding behavior is common on highway in China especially on the overtaking line, we tend to focus on the results of 100 km/h when designing the tunnel's cross section.

Tunnel's Entrance. At tunnels' entrance, drivers experience the process from strong crosswind to feeble crosswind. Figure 6 shows the distribution of MCSWA of ten drivers considering different shoulder width at tunnels' entrances.

Then we use cubic spline fitting for average MCSWA ($\bar{y}$) and left shoulder width (x).

(1) The speed limit of 80 km/h—

Cubic spline fitting:

$$\begin{aligned} \bar{y} &= -6.278x^3 + 33.42x^2 - 55.56x + 36.51 \\ R^2 &= 0.7831 < 0.8,\ \text{SSE} = 3.09 \end{aligned} \tag{4}$$

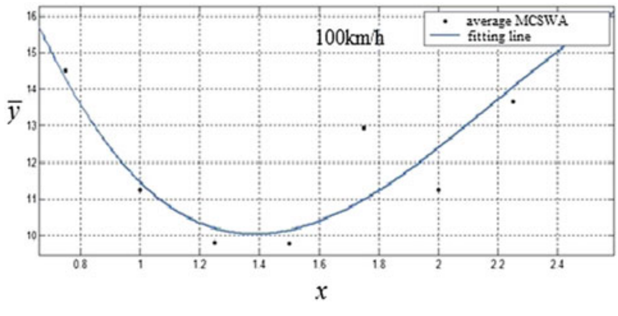


Fig. 5 Cubic spline fitting for average MCSWA ($\bar{y}$) and left shoulder width (x) (100 km/h)

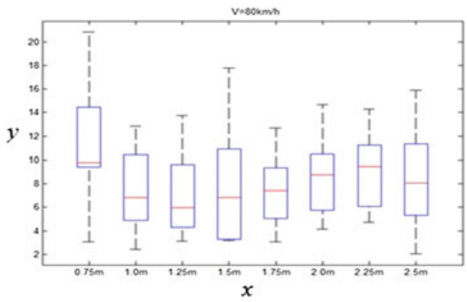


Fig. 6 The distribution of MCSWA of ten drivers considering different shoulder widths. *x* refers to the shoulder width at tunnels' entrances. *x* refers to the shoulder width and *y* refers to MCSWA

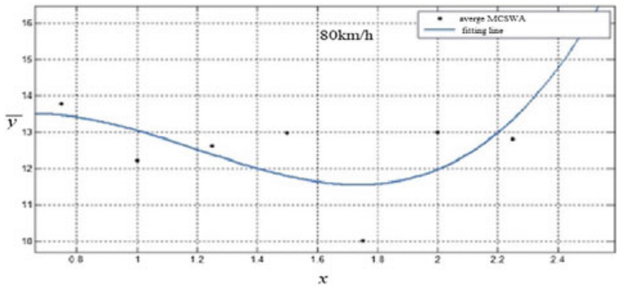


Fig. 7 Cubic spline fitting for average MCSWA ($\bar{y}$) and left shoulder width (x) (80 km/h)

By solving the Eq. (4), the value of $\bar{y}$ reaches the minimum 6.96° when x equals 1.33 m. We presume that the most suitable shoulder width for driving safety is 1.33 m (Fig. 7).

(2) The speed limit of 100 km/h—

Cubic spline fitting:

$$\begin{aligned} \bar{y} &= 0.001152x^3 + 2.527x^2 - 8.918x + 13.51 \\ R^2 &= 0.8409 > 0.8,\ \text{SSE} = 1.037 \end{aligned} \tag{5}$$

By solving the Eq. (5), the value of $\bar{y}$ reaches the minimum 5.65° when x equals 1.762 m. We presume that the most suitable shoulder width for driving safety is 1.33 m.

3.2 Heart Rate

The heart rate of people in resting state ranges from 60 to 75/min while it could be over 100/min if people are stimulated by outside factors. So heart rate is an important index that reflects drivers' physiological and mental state to a certain extent. Drivers of too heavy mental and physiological load are inclined to make mistakes and accidents follows. Due to the individual differences, the basic heart rate of drivers varies. The variability of heart rate has been used in some researches to show how outside factors affect driving safety. In this research, we primarily observe drivers' heart rate near the tunnels' entrances and exits. Specifically, variability of maximum heart rate (abbreviated as VMHR below) during this process is the index we investigate.

At tunnel's exit, caused by the change of driving scenario and crosswind, the VMHR of drivers raise usually. However, the degree that drivers' physiological state fluctuates differs under different conditions of left shoulder widths. Figure 8 shows how different left shoulder widths affect drivers' VMHR at tunnel's exit. If the shoulder width is relatively large (>1.5 m), VMHR is small and there is no apparent difference between 80 and 100 km/h. In contrast, VMHR raises dramatically as the left shoulder width decreases when it is smaller than 1.5 m, which to a certain extent explains the changing behavior of MCSWA. If the lateral clearance is too restrained, drivers would feel physiologically and mentally overloaded, tending to whirl the steering wheel more, which may endanger the vehicle's stability. Specifically, for the speed limit of 80 km/h, the max VMHR is 6.3 % while it is 8 % for 100 km/h.

At tunnel's entrance, the changing behavior of VMHR resembles to that at tunnel's exit. When the left shoulder width is relatively large (>1.5 m), the VMHR is around 3–4 %. However, if it is less than 1.5 m, the lines become much steeper. Compared with the tunnel's exit, the VMHR is generally smaller when getting into tunnels. That means the disappearance of crosswind causes less effect to drivers than the appearance of it. This finding means a lot for driving safety on bridge-tunnel linkage sections (Fig. 9).

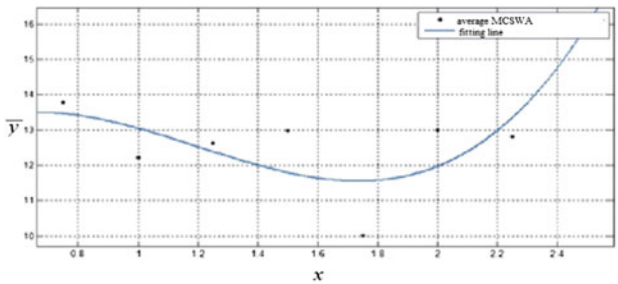


Fig. 8 Cubic spline fitting for average MCSWA ($\bar{y}$) and left shoulder width (x) (100 km/h)

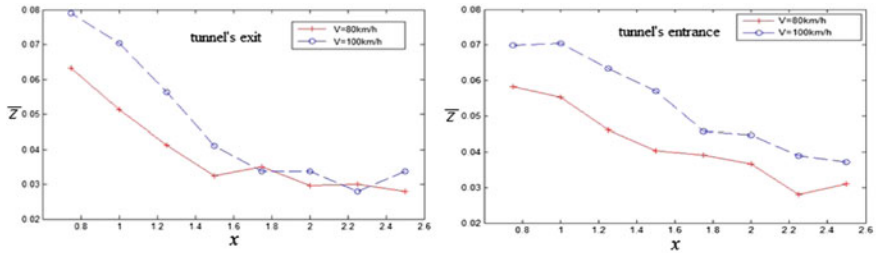


Fig. 9 Scattering plot of average VMHR ($\bar{z}$) and shoulder width (x). *Red* line refers to the condition of 80 km/h while the *blue* line refers to that of 100 km/h

3.3 Optimized Design of Bridge-Tunnel Linkage Section

In Sects. 3.1 and 3.2, we have discussed the max countering steering wheel angle and heart rate of drivers to investigate how different left shoulder widths at bridge-tunnel linkage sections affect driving safety from the perspective of drivers. These findings above is helpful to propose an optimized design of cross section. On top of that, we have to ensure drivers experience a comfortable transitional process. That means keeping MCSWA and VMHR at a comparatively low level. It has to be mentioned that the simulation driving is conducted under a certain condition (the road is straight, the wind speed is 60 km/h and is vertical to road), so the given design below only provides a general method.

According to the findings in Sect. 3.1, too large or too small shoulder space may cause negative effects on drivers' manipulation. From the perspective of MCSWA, the shoulder width should be around 1.4 m at tunnel's exit and 1.3 m around at tunnel's entrance if the speed limit is 100 km/h. Taking the drivers' heart rate into consideration, neither 1.4 nor 1.3 m is the best choice because the VMHR is a little bit large. Combining the two factors above and construction difficulty, we make a tradeoff that shoulder width of tunnel's exit and entrance are both 1.5 m (Fig. 10).

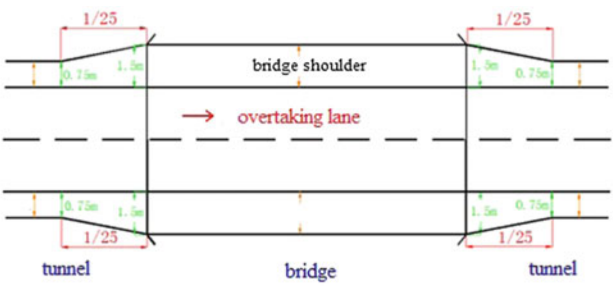


Fig. 10 Optimized design of bridge-tunnel linkage sections

We all know that wind barriers are often used in mountainous highways to reduce the effect of crosswind on vehicles. Compared with that, the way of changing cross section design has its advantage. As we all know, snow and ice would cause friction to change, especially near the tunnel's exit and entrance. From the view of fault-tolerant, more lateral clearance gives the space to adjust vehicles. Though this way may add construction difficulty, it can be seen as a useful and innovative way to improve driving safety of mountainous highway.

4 Conclusion

We focus on the driving safety of bridge-tunnel linkage sections on mountainous highway. Vehicles will experience the sudden change of crosswind while getting in or out of the tunnel, which may induce accidents if not handled appropriately. Different from previous researches, we propose an optimized design of bridge-tunnel linkage sections. Among many indexes of road section, the left shoulder width is chosen as the main factor we think affects driving safety a lot. There do have some researches on lateral deviation caused by crosswind and then offer suggestions about lateral clearance, however, they more or less ignore the driver's state. In this study, we study the driver's manipulation of steering wheel and driver's heart. These two factors can to a large extent reflect driving safety from the perspective of drivers.

After the simulation driving experiment and data analysis, we do have some findings. First of all, too large or too small left shoulder width would cause drivers to whirl more fiercely to counteract the effect of sudden change of crosswind. Purely from the view of steering manipulation, we presume that the left shoulder width is most proper for driving safety if the max countering steering wheel angle (MCSWA) reaches its minimum. Then the appropriate left shoulder width at tunnel's exit and entrance is calculated. Apart from MCSWA, driver's mental and physiological state is also vital for driving safety. So we adjust the value of most suitable left shoulder width ever calculated previously by keeping the value of variability of max heart rate (VMHB) at a relatively low level. At last, an optimized of bridge-tunnel linkage sections is offered.

There is no denying that some parts of this paper need improvement and further investigating. The optimized design mentioned above is only suitable for the crosswind of 60 km/h, which is vertical to the road. However, this sample at least provides an innovative thinking that we can enhance driving safety from the design phase. What' more, driving safety is more an issue that involves driver behavior much, so we should never forget to take drivers' feelings into account.

Acknowledgments This research was jointly sponsored by Project 51278361 and 51508409 supported by the National Natural Science Foundation of China and the Project-sponsored by SRF for ROCS, SEM.

References

1. Bettle, J., Holloway, A.G.L., Venart, J.E.S.: A computational study of the aerodynamic forces acting on a tractor-trailer vehicle on a bridge in cross-wind. J. Wind Eng. Ind. Aerodyn. **91**(5), 573–592 (2003)
2. Cheung, M.M.S., Chan, B.Y.B.: Operational requirements for long-span bridges under strong wind events. J. Bridge Eng. **15**(2), 131–143 (2010)
3. Kim, K., Choi, J., Yi, K., et al.: Lateral disturbance compensation using motor driven power steering. In: IEEE 75th Vehicular Technology Conference (VTC Spring) (2012)
4. Han, Y., Shu, S.Q., Tan, D.: Numerical simulation on aerodynamic characteristics of road vehicles on bridges under cross winds. Adv. Technol. Manuf. Eng. Mater., Pts 1–3 **774–776**, 241–247 (2013)
5. Pinelli, J.P., Subramanian, C., Plamondon, M.: Wind effects on emergency vehicles. J. Wind Eng. Ind. Aerodyn. **92**(7–8), 663–685 (2004)
6. Coleman, S.A., Baker, C.J.: High sided road vehicles in cross winds. J. Wind Eng. Ind. Aerodyn. **36**, Part 2(0), 1383–1392 (1990)
7. Maruyama, Y., Yamazaki, F.: Driving simulator experiment on the moving stability of an automobile under strong crosswind. J. Wind Eng. Ind. Aerodyn. **94**(4), 191–205 (2006)
8. Rodriguez, J.M., Codjoe, J., Osman, O., et al.: Experimental modeling of the effect of hurricane wind forces on driving behavior and vehicle performance. J. Emerg. Manage. **13**(2), 159–172 (2014)
9. Balsom, M., Wilson, F., Hildebrand, E.: Impact of wind forces on heavy truck stability. Transp. Res. Rec. J. Transp. Res. Board **1969**, 115–120 (2006)
10. Kim, K., Kim, B., Go, Y., et al.: An investigation on motor-driven power steering-based crosswind disturbance compensation for the reduction of driver steering effort. Veh. Syst. Dyn. **52**(7), 922–947 (2014)
11. Young, R.K., Liesman, J.: Intelligent transportation systems for operation of roadway segments in high-wind conditions. Transp. Res. Rec. **2000**, 1–7 (2007)
12. Baker, C.J.: The quantification of accident risk for road vehicles in cross winds. J. Wind Eng. Ind. Aerodyn. **52**, 93–107 (1994)
13. Kim, K.W., Lee, S.B., Park, C.S., et al.: Estimation of lateral disturbance under parameter uncertainties. Int. J. Automot. Technol. **16**(3), 427–433 (2015)
14. Saito, T., Itoh, M., Inagaki, T., et al.: Effects of environmental factors on the steering operation when driving a car. In: SICE Annual Conference, pp. 2360–2365 (2013)
15. Ben-Bassat, T., Shinar, D.: Effect of shoulder width, guardrail and roadway geometry on driver perception and behavior. Accid. Anal. Prev. **43**(6), 2142–2152 (2011)

Part VIII
Road and Rail—Accidents and Pedestrian Modeling

Reducing Reversing Vehicle Incidents in Australian Fleet Settings—A Case Study

Darren Wishart, Klaire Somoray and Bevan Rowland

Abstract Reversing vehicle incidents is a significant but often overlooked issues in organisations. Utilising three Australian organisations, this study aimed to evaluate the efficacy of reversing aids and a behaviour-change program in reducing reversing-related crashes in fleet settings. Reversing-related incidents increased from Time 1 to Time 2 in the organisation that did not implement a specific strategy to reduce their reversing-related crashes and in the organisation that implemented the reversing aids intervention. However, the increase was only statistically significant in the organisation that utilised the reversing aids technology. In this organisation, the odds of its drivers getting involved in a reversing incident has almost doubled from Time 1 and Time 2. In contrast, the frequency of reversing incidents in the organisation implementing the behaviour-change program has significantly decreased, with less than 50 % chance of its drivers being involved in a reversing incident from Time 1 to Time 2. The implications associated with these results will be discussed.

Keywords Reversing-related incidents · Work-related driving safety · Technology-assisted reversing aids · Behaviour change

D. Wishart (✉) · K. Somoray · B. Rowland
Centre for Accident Research and Road Safety Queensland (CARRS-Q), Queensland University of Technology, 130 Victoria Park Road, Kelvin Grove, Brisbane, QLD 4059, Australia
e-mail: d.wishart@qut.edu.au

K. Somoray
e-mail: k.somoray@qut.edu.au

B. Rowland
e-mail: b.rowland@qut.edu.au

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_61

1 Introduction

Driving for work is potentially one of the riskiest activities undertaken in the course of a person's work, which is evidenced by the over-representation of crashes and injuries involving the operation of motor vehicles while undertaking work activities [1–5]. Research has also demonstrated that work drivers are consistently involved in a higher level of crash involvement in comparison to private car drivers [6–8].

Within the work-driving literature, research has indicated an overrepresentation of reversing related incidents across various fleet settings [9–11]. Health and Safety Executive in the UK [11] reported that a quarter of all work-related fatalities in the UK workforce involved reversing vehicle incidents. Similar trends are seen in Australia. For instance, it is estimated that nearly half (46 %) of fatalities involving heavy trucks that occurred in Australian worksites from 2003 to 2012 is attributed to a reversing vehicle [12]. In light vehicle fleet settings, Davey et al. [9] examined five organisations that operate light fleet vehicles and found that reversing-related incidents were one of the top three incident types across all of the participating organisations, accounting for 17–28 % of their total crashes. In Davey et al.'s study, the work drivers reported that low speed manoeuvring and reversing-related incident is a significant issue in their organisations especially when visiting residential environments (e.g., delivering goods, providing trade service) [9]. Accumulation of such incidents also often results in expensive insurance claims for employees and organisations [9]. Yet, reversing-related incidents are often viewed as minor crashes despite the risk of death or serious injury and expensive costs.

Although some reversing-related incidents occur in the workplace, many reversing incidents also occur in residential driveways [13, 14]. Reversing-related incidents that occur in domestic settings often involve other vulnerable road users such as pedestrians, cyclists, and motorcyclists, and research indicates that the majority of people injured as a result of reversing incidents are children and older people [15–18]. Previous research has also shown that many low speed vehicle deaths that involved children have been a result of reversing vehicles [14, 19, 20]. Data from the Australian Road Deaths Database showed that, on average, seven pedestrians aged 0–14 were killed and 60 were seriously injured each year from 2001 to 2010 due to a reversing vehicle [18].

1.1 Strategies to Reduce Reversing Vehicle Incidents

While reversing-related incidents, particularly within the work driving setting, appears quite prevalent, there is a lack of research examining the types of intervention strategies designed and implemented to reduce reversing-related crashes. Technology assisted reversing aid systems, such as audible warning systems and visual cameras, are commonly utilised by organisations to reduce the risk of reversing-related incidents [21]. These technologies are designed to assist drivers in parking and manoeuvring while driving in reverse. Recently, the United States passed a legislation that requires car manufacturers to install reversing aids to new vehicles [22]. However, there is limited research that supports the efficacy of reversing aid technology in reducing reversing incidents. These technologies may even increase the risk of fleet vehicle crashes due to an over-reliance on the technology, instead of improving driving safety behaviours. Research has shown that, even though most work vehicles are equipped with audible reversing warning systems and cameras, there are a number of challenges that may compromise the use of some of these technologies within work settings. For example, reversing alarms may not always be clear especially in a noisy work environment and workers have reported turning off the reversing warning systems in their work vehicles due to its annoying sound [10, 12].

Other research have argued that interventions should focus on improving workers' driving behavior and attitudes along with organisational safety climate and management processes [9, 11]. For instance, occupational guidelines have suggested the use of colleagues and passengers to act as spotters and marshals to assist in guiding the driver while reversing and the "walk around" procedure prior to reversing to prevent the occurrence of a reversing related crash [9, 11, 23, 24]. The "walk around" procedure involves drivers, in the absence of a colleague or passenger walking around the vehicle to check for hazards before manoeuvring or reversing [9, 11, 23].

1.2 Current Study and Aim

While research have suggested that changes in drivers' behaviour and management processes lead to safer occupational driving [1, 3], in contrast to simply only changes in vehicle equipment and technology, there is no current research that specifically evaluates behavior change programs and technology assisted reversing aids to reduce the frequency of reverse-related incidents. Consequently, there is a need for research to investigate and compare the use of reversing aids such as audible reversing sensors and cameras with behavioural strategies that target reversing driving behaviours such as walk around and marshalling procedures. The aim of this research is to evaluate various reversing intervention strategies adopted by three organisation and assess the efficacy of each strategy in reducing reversing crashes within the work driving setting.

2 Methodology

2.1 Participants and Procedure

Three local government organisations geographically located within city-based environments and situated in South East Queensland, Australia agreed to participate in the research. Each organisation undertakes a range of similar activities relating to responsibilities associated with the care and maintenance of services required by the community. Initially, each organisation participated in a workshop and agreed upon the information required to record a crash event. The workshop was carried out to ensure that each organisation reached a level of consensus in regards to defining, recording, and reporting of crashes. The first stage of the research was to examine the nature of work driving crashes within each organisation (e.g., type of crash, costs, who's at fault, crash location, crash description). The acquired crash data allowed the researchers to examine the impact of the interventions implemented by each organisation between two time periods based on 12 months apart.

2.2 *Interventions*

Each organisation and the interventions implemented are described and listed below. Organisation A did not implement any intervention strategies specifically aimed at addressing reversing issues, but rather continued general work driving safety initiatives already in place. Organisation B embarked upon a program integrating reversing aids such as reversing sensors and reversing cameras across their light vehicle fleet. The organisation also embarked upon a retro fitting program to incorporate reversing technology to vehicles that previously did not possess reversing sensors or cameras.

Organisation C implemented a behaviour-change program, which incorporated new policies, procedures, and managerial processes relating to reversing vehicles. This program consisted of light vehicle fleet operators asking other workers to act as a "marshal" to guide the driver while engaging in reversing manoeuvres or light vehicle fleet operators conducting "walk arounds" when approaching a vehicle to check for potential obstacles prior to reversing, if they are working on their own. A third process that was incorporated into the program was an increase in managerial processes associated with those involved in a reversing crash. For example, this process included activities such as crash investigations and interviews with a manager to seek explanations as to why a reversing crash still occurred when an employee other than the driver was present. This process was undertaken to determine if the above-mentioned policies and procedures were being adhered to. An education and promotion campaign relating to the implementation of these new policies and procedures were introduced to ensure that all light vehicle fleet operators were aware of the policy and procedure requirements.

3 Results

3.1 *Descriptives*

Crash Data at Time 1. Crash data obtained from each organisation at Time 1 indicated that reversing-related incident was the most prevalent crash type for all

Table 1 Most common incident types in time 1

	1st most common incident	2nd most common incident	3rd most common incident	Total crashes
Organisation A	Reversing: 26 (19.0 %)	Hitting an object: 22 (16.1 %)	Damaged whilst parked: 18 (13.1 %)	137
Organisation B	Reversing: 29 (34.9 %)	Rear end: 16 (19.3 %)	Parking: 11 (13.3 %)	83
Organisation C	Reversing: 41 (34.7 %)	Damage whilst parked: 19 (16.1 %)	Hit by an object: 10 (8.5 %)	118

three organisations (ranging from 19.0 to 34.9 %, refer to Table 1). Reversing-related incidents are defined as workers reversing into an object or another vehicle, while performing work-related activities or being hit by another reversing vehicle. Other incident types which were highly represented were, damage whilst parked, rear end incidents, hitting an object, hit by an object and parking (refer to Table 1).

The proportion of the crash repair costs relating to reversing related incidents were also calculated in Time 1 for each organisation. Reversing-related incidents accounted for between 14.5 and 37.2 % of the total crash repair costs. The average crash repair costs of reversing crashes ranged from $900 to $1200 per incident (Table 2).

Crash Data in Time 2. Crash data of the three organisations were analysed at Time 2, with results demonstrating that reversing was still the most common type of crash ranging from 21.3 to 47.9 %. The next most common types of crashes were: damage whilst parked, rear end incidents, hitting an object, hit by an object, parking and turning/merging (refer to Table 3 for the proportion of each crash type in Time 2).

The proportion of the crash repair costs relating to reversing related incidents were also calculated in Time 2 for each organisation. Reversing-related incidents in Time 2 accounted between 14.9 and 22.8 % of the total crash repair costs. In Time 2, the average crash repair costs of reversing crashes ranged from $500 to $1500 per incident. The results also demonstrate that there were increases in total repair cost of

Table 2 Total costs of crashes and the proportion of reversing crash costs of each organisation in time 1

	N of reversing crashes	Total crash costs	Total reversing crash costs (%)	Maximum costs of reversing crashes	Mean costs of reversing crashes
Org A	26	$181,285.76	$29,984.78 (14.5 %)	$5465.92	$1153.26
Org B	29	$172,786.43	$26,372.86 (15.3 %)	$3924.70	$909.41
Org C	41	$133,277.11	$49,542.80 (37.2 %)	$5242.48	$1208.36

Table 3 Most common incident types in time 2

	1st most common incident	2nd most common incident	3rd most common incident	Total crashes
Organisation A	Reversing: 26 (21.3 %)	Hit by an object: 17 (13.9 %)	Hitting an object: 12 (9.8 %)	122
Organisation B	Reversing: 57 (47.9 %)	Rear end: 13 (10.9 %)	Turning/merging: 11 (9.2 %)	119
Organisation C	Reversing: 34 (21.4 %)	Parking: 26 (16.4 %)	Damage whilst parked: 23 (14.5 %)	159

Table 4 Total costs of crashes and the proportion of reversing crash costs of each organisation in time 2

	N of reversing crashes	Total crash costs	Total reversing crash costs (%)	Maximum costs of reversing crashes	Mean costs of reversing crashes
Org A	26	$202,068.61	$396,35.81 (19.6 %)	$22,961.14	$1524.45
Org B	57	$321,353.22	$731,43.04 (22.8 %)	$5674.26	$1283.21
Org C	34	$117,021.64	$17,458.32 (14.9 %)	$1708.77	$5,13.48

crashes from Time 1 to Time 2 for Organisation A and Organisation B, but a decrease in total costs for Organisation C. In addition, the total costs of reversing crashes increased in Organisation A and Organisation B from Time 1 to Time 2, while Organisation C decreased the total costs of reversing crashes (Table 4).

3.2 Data Analysis

Chi-square tests were carried out to determine if the proportion of reversing crashes compared to the non-reversing crashes, changed from Time 1 to Time 2 between the three organisations. Using a binary logistic regression, the crash data was further analysed to test the predictive power of the model in each organisation, using time as a predictor.

Analysis 1: Chi-Square Tests. There was a total of 738 crashes that occurred in the three organisations overall from both Time 1 and Time 2. Reversing crashes accounted for 28.7 % ($N = 213$) of the total crashes recorded in both Time 1 and Time 2 across all participating organisations. Table 5 shows the frequency and proportion of non-reversing and reversing incidents of each organisation at each of the two time points.

Table 5 Frequency and proportion of non-reversing and reversing crashes of each organisations over the two-year period

	Organisation A (no specific strategy)		Organisation B (reversing aids)		Organisation C (behaviour change program)	
	Non-reversing (%)	Reversing (%)	Non-reversing (%)	Reversing (%)	Non-reversing (%)	Reversing (%)
Time 1	111 (81.0 %)	26 (19.0 %)	54 (65.1 %)	29 (34.9 %)	77 (65.3 %)	41 (34.7 %)
Time 2	96 (78.7 %)	26 (21.3 %)	62 (52.1 %)	57 (47.9 %)	125 (78.6 %)	34 (21.4 %)
	$\chi^2(1) = 0.10$, $p = 0.755$		$\chi^2(1) = 2.85$, $p = 0.091$		$\chi^2(1) = 5.47$, $p = 0.019$	

Chi-square tests with Yates' continuity correction, revealed that, there was an increase in reversing-related crashes in Organisation A (no strategy) and Organisation B (reversing aids) from Time 1 to Time 2, but the change was not statistically significant. For Organisation C (behaviour change), the proportion of reversing incidents was lower at Time 2 compared to Time 1 and the change was statistically significant (refer to Table 5).

To test the homogeneity of odds ratio across the groups, the Breslow-Day test was conducted. The Breslow-Day test showed that, from Time 1 to Time 2, the change in the proportion of reversing crashes for Organisation A was similar to the change in the proportion of reversing crashes for Organisation B, $\chi^2(1) = 0.84$, $p = 0.359$. This is perhaps not surprising, given that over time, both organisations experienced an increase in reversing crashes that did not reach statistical significance. In contrast, the test demonstrated that, over time, the change in the proportion of reversing crashes for Organisation B (reversing aids) is different from the change in the proportion of reversing crashes experienced by Organisation C (behavior change), $\chi^2(1) = 9.15$, $p = 0.002$. Again, this result is not surprising since Organisation B experienced an increase in reversing crashes, while Organisation C experienced a decrease in reversing crashes from Time 1 to Time 2. This difference in effect of time is also found between Organisation A (no strategy) and Organisation C (behavior change), $\chi^2(1) = 3.93$, $p = 0.048$, with Organisation A experiencing an increase in reversing crashes and Organisation C experiencing a decrease in reversing crashes from Time 1 to Time 2.

Analysis 2: Chi-Square Tests on At-Fault Reversing Crashes Only. Further analysis of the crash databases of the three organisations showed that the reversing-related incidents can be divided into two types: (1) reversing-related incidents at-fault and (2) reversing-related incidents not-at-fault. When reversing incidents are at-fault, the work driver of the participating organisation reversed into an object or another vehicle. When reversing incidents are not-at-fault, other third party drivers have reversed into the vehicle of one of the participating organisations. There was a total of 13 reversing-related incidents not at fault in Time 1 and 17 reversing incidents not at fault in Time 2 across the three organisations. To investigate the significance of the difference between these two reversing-related incident types, further in-depth analyses were conducted treating the not-at-fault reversing-related incidents as non-reversing crashes. When the Chi-square tests were conducted using the above parameters, there were no statistically significant differences between analysis 1 and analysis 2. However, the increase in reversing crashes from Time 1 to Time 2, without the third party at fault crashes, was approaching statistical significance in Organisation B, $\chi^2(1) = 3.74$, $p = 0.053$.

To test the homogeneity of odds ratio across the groups, further Breslow-Day tests were conducted, focusing only on the at-fault reversing-related incidents. Results indicated that there were no significant changes between analysis 1 and analysis 2. Due to the small difference in the two analyses, the researchers decided to carry out the subsequent analyses treating crashes as reversing-related incidents where the participating organisations' drivers were at-fault only.

Analysis 3: Binary Logistic Regression Analysis. Using a binary logistic regression, the crash data was further analysed to test the predictive power of the model and a significant interaction was found between the variables of organisation and time, $\chi^2(2) = 14.09$, $p = 0.001$. Therefore, the crash data for each organisation was analysed separately to assess the effect of time in each organisation. As shown in Table 5, there was an increase of reversing crashes in Organisation A (no strategy) and Organisation B (reversing aids) between Times 1 and 2, as indicated by the positive beta-weights. However, the increase in reversing crashes from Time 1 to Time 2 was only statistically significant in Organisation B (reversing aids). The odds ratio of reversing crashes in Organisation B (reversing aids) was 1.89. This result suggest that, the chances of being involved in a reversing-related incident has almost doubled from Time 1 to Time 2 for work drivers in Organisation B. In contrast, there was a significant decrease in the frequency of reversing-related incidents in Organisation C (behavior change) as indicated by the negative beta-weights. The odds ratio of the reversing crashes in Organisation C was 0.41, suggesting that the chances of their work drivers in having a reversing-related incident has decreased by 41 % from Time 1 to Time 2 (Table 6).

Table 6 Logistic regression analysing the effect of time on each organisation

Organisation	B (*S.E.*)	Odds ratio	Wald χ^2	*P*-value	95 % CI for odds ratio
Organisation A (no specific strategy)	0.25 (0.33)	1.28	0.58	0.448	[0.68, 2.42]
Organisation B (reversing aids)	0.64 (0.31)	1.89	4.28	0.038	[1.03, 3.46]
Organisation C (behaviour change program)	−0.89 (0.29)	0.41	9.30	0.002	[0.23, 0.73]

4 Discussion

Reversing-related incidents are over represented in work driving light vehicle fleet settings. In addition, not only is the cost of work-related reversing crashes considerably high, but the potential for a reversing incident to escalate to a serious injury or fatality is extremely plausible [9]. This research conducted a field-based study whereby two participating organisations implemented specific intervention strategies designed to reduce the frequency of reversing-related incidents in their organisations. Organisation B implemented a strategy of fitting technology-assisted reversing aids to all vehicles, while Organisation C embarked upon a behaviour change program, which involves drivers performing a walk around, use of other workers as marshals while reversing and a managerial process. A third organisation continued with their usual work driving safety initiatives and did not implement any initiatives specifically targeting reversing-related incidents.

Interestingly, the results of this study revealed that the fitting of reversing aid technologies to vehicles actually resulted in an increase in the frequency of reversing crashes in Organisation B following 12-month period. Given the costs associated with the retro fitting, this outcome questions the supposed benefits associated with the fully-fledged implementation of this intervention, especially in regard to the increase in costs associated with the increase of reversing crashes.

In contrast, in implementing a behavior change program and a managerial process, Organisation C experienced a significant reduction in their reversing-related incidents and a substantial decrease in the associated reversing crash repair costs. The results experienced by Organisation C indicate that, in regards to a potential cost benefit, the minimal cost associated with implementing the behaviour change program would appear considerably smaller than the cost savings of repairs through a decrease in reversing crashes. This result also provides further support to previous studies that found improved driving behaviours and safety records in work drivers as a result of behaviour focused programs and managerial processes [3, 25, 26].

4.1 Strengths and Implications

Although these results appear to assert that technology assisted reversing aids (e.g., audible alarms and reversing cameras) appear to have little impact on preventing reversing crashes within the work setting, there are a number of considerations worthy of further discussion. Firstly, while it is acknowledged that the fitting of technology assisted reversing aids can provide benefits, simply taking a silver bullet approach [1] and fitting technology to a vehicle without the initiative being supported by educational programs instructing drivers on the use of the technology is unlikely to result in expected benefits. Research suggest that interventions and strategies in improving work drivers safety should be proactive and multidimensional, with an aim at addressing both behavioural and organisational factors contributing to the crash [1]. In this study, the organisation that installed reversing aids in their vehicle fleets did not provide educational program to its drivers on how to use the new technology fitted in the vehicles nor any managerial process or crash investigations. Therefore, it would appear that any strategies implemented should be supported by further initiatives such as behavioural change and educational programs along with managerial and crash investigation processes.

The results of the study also indicated that while the number of reversing crashes decreased in Organisation C after implementing the behavior change process specifically aimed at reversing-related incidents, the organisation experienced an increase in the overall reported crashes from Time 1 to Time 2. It is suggested that the increase in overall crashes could be in part attributed to improved safety initiatives and awareness within the organisation resulting in improved reporting of crashes. Alternatively, the increase in overall crashes could be as a result of the organisation's intervention strategies focusing only on reversing-related incidents as against to other crash types. Nevertheless, although the number of the total crashes increased for Organisation C at Time 2, the total costs associated with these crashes decreased. This results suggest that even though the organisation experienced a higher frequency of crashes, they were likely less severe crashes than previously.

This study offers a particular number of strengths associated with the results. Firstly, it has demonstrated that a low cost multi-faceted intervention appears to have been successful in decreasing reversing crashes and costs. Therefore, the practical implications of these results are particularly beneficial to organisations that lack resources required to implement technological solutions. This study also offers a particular strength in that the study evaluated different reversing prevention strategies in real world light fleet environments with organisations that conduct similar work activities in similar geographical environments.

4.2 Limitations and Conclusions

As in all studies, the results obtained must be viewed in conjunction with study limitations. Although the results obtained are associated with interventions directly aimed at preventing reversing crashes, there are other potential factors inherent within a dynamic road environment that could not be controlled for and all results obtained should be viewed accordingly. It should also be noted, that although retro fitting of reversing assisted technology resulted in an increase in the frequency of reversing crashes, the researchers were not involved in the selection or fitting of the technology, and consequently, could not conduct any evaluation on the types of technology used or the manner in which they were used. Additionally, while the organisations had similar fleet sizes and conducted similar work activities, the exact number and nature of their work vehicles were not provided.

Notwithstanding the limitations associated with this study, the results do provide some practical assistance and intervention strategies to industry conducting light vehicle fleet operations. For example, any intervention with a silver bullet approach is unlikely to have a substantial effect on reducing reversing crashes, thus any intervention implemented to improve drivers safety at work should contain various stages and consist of a multifaceted approach. Future research could expand upon this current study by investigating the different types of technology available and compare the use of these technologies alongside other strategies.

References

1. Wishart, D.E., Davey, J.D.: A research based case study approach to the development of fleet safety interventions in large vehicle fleets. In: Proceedings Safety in Action Conference, Melbourne (2004)
2. Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A.A., Jarawan, E., Mathers, C.: World Report on Road Traffic Injury Prevention. World Health Organization (2004)
3. Haworth, N., Tingval, C., Kowadlo, N.: Review of best practice road safety initiatives in the corporate and/or business environment. In: Monash University Accident Research Centre (2000)
4. Wishart, D.E., Rowland, B.D., Freeman, J.E., Davey, J.: Fleet safety: risk reduction for work related vehicle use. In: Proceedings of 9th Industrial and Organisational Psychology Conference, Brisbane (2011)
5. Wishart, D.E.: The challenge of developing a fleet driving risk assessment tool: what can be learned from the process? Dissertation, Queensland University of Technology (2015)
6. Australian Transport Council: National Road Safety Strategy: 2011–2020. Department of Infrastructure and Regional Development, Australian Transport Council (2011)
7. Downs, C., Keigan, M., Maycock, G., Grayson, G.: The Safety of Fleet Car Drivers: A Review. Transport Research Laboratory (1999)
8. Lynn, P., Lockwood, C.: The Accidental Liability of Company Car Drivers. Transport Research Laboratory (1999)
9. Davey, J., Rowland, B., Freeman, J.: QIPC Low Speed Run-Overs of Young Children Research Project Final Report. Centre for Accident Research and Road Safety, Queensland (2010)
10. Debnath, A.K., Blackman, R.A., Haworth, N.L.: Understanding worker perceptions of common incidents at roadworks in Queensland. In: Proceedings of the 2013 Australasian Road Safety Research, Policing and Education Conference, pp. 1–13, Brisbane (2013)
11. Health and Safety Executive: A Guide to Workplace Transport Safety. HSE (2014)
12. Safe Work Australia: Work-Related Fatalities Involving Trucks, Australia, 2003 to 2012. SWA (2014)
13. Anthikkat, A.P., Page, A., Barker, R.: Low-speed vehicle run over fatalities in Australian children aged 0–5 years. J. Paediatr. Child Health **49**, 388–393 (2013)
14. Armstrong, K.A., Obst, P.L., Davey, J., Thunstrom, H.: Developing guidelines for intervention to reduce risk of low-speed vehicle run-overs of young children. Inj. Prev. **16**, A90–A90 (2010)
15. Oxley, J., Corben, B., Fildes, B., O'Hare, M., Rothengatter, T.: Older Vulnerable Road Users—Measures to Reduce Crash and Injury Risk. Monash University Accident Research Centre (2004)

16. Griffin, B., Watt, K., Kimble, R., Wallis, B., Shields, L.: Systematic literature review of incidence rates of low-speed vehicle run-over incidents in children. Worldviews Evid. Based Nurs. **11**, 98–106 (2013)
17. Williamson, A., Irvine, P., Sadural, S.: Analysis of Motor Vehicle Related Fatalities Involving Children Under the Age of Six Years (1995–2000). Motor Accidents Authority (2002)
18. Bureau of Infrastructure Transport and Regional Economics (BITRE): Child Pedestrian Safety: 'Driveway Deaths' and 'Low-Speed Vehicle Run-Overs', Australia, 2001–10. Department of Infrastructure and Regional Development, BITRE (2012)
19. Murdoch, J.: Paediatric Driveway Injury: Vulnerable Groups and Prevention Strategies. The Centre for Accident Research and Road Safety (2008)
20. Davey, J., Dingle, G., Clark, M., Johnston, T., Woods, S., White, J., Freeman, J.: The prevalence and characteristics of paediatric driveway accidents in Queensland. J. Aust. Coll. Road Saf. **18**, 34–40 (2007)
21. Gallagher, V.A.: Backup alarms: preventing injuries caused by reversing vehicles. Prof. Saf. **48**, 28–32 (2003)
22. National Highway Traffic Safety Administration: Federal Motor Vehicle Safety Standards; Rear Visibility, pp. 19177–19250. Department of Transportation. U.S. Government Printing Office, Washington, DC (2014)
23. The Royal Automobile Club of Queensland Limited: Submission to Parliamentary Travelsafe Committee: Investigation into Low Speed Run-Overs in Queensland. The Royal Automobile Club of Queensland (2007)
24. Cooper, D.L., Duffy, S., Orrick, P., Ragland, D.R.: Develop Methods to Reduce or Prevent Backing Crashes. UC Berkeley Traffic Safety Center (2010)
25. Murray, W., Ison, S., Gallemore, P., Nijjar, H.: Effective occupational road safety programs: a case study of Wolseley. Transp. Res. Rec. J. Transp. Res. Board **2096**, 55–64 (2009)
26. Murray, W., Dubens, E., Rea, M.: Work-related road safety: good practice cases from around the world. In Australian College of Road Safety National Conference, Perth (2009)

The Relationship Between Traffic Rule Violations and Accident Involvement Records of Drivers

Mohamed Shawky, Yousef Al-Badi, Iyad Sahnoon and Hussain Al-Harthi

Abstract This paper aims to explore the relationship between the at-fault drivers involved in traffic accidents and their history of traffic violation records as a function of drivers' behavior. The employed data was integrated from different dataset systems in Abu Dhabi Traffic Police including traffic violations, accident information and drivers' licenses data systems. About 713,783 drivers involved in the analysis process with total accident number of 690,697 and total violation number of 2,762,011 during five years from 2010 to 2014. The analysis addressed two main parameters; accident rate per drivers and ratio of drivers with accident. Each parameter is investigated in terms of different variables; total number of violations, number of hazard violations, number of violations with penalty points and a cumulative number of traffic penalty points. The regression analysis shows a very strong relationship between the two parameters and the explanatory variables. In conclusion, the results indicated that the driver risk to be involved in future accidents can be predicted from prior driving records for traffic violations.

Keywords Traffic accident risk prediction · Traffic violation · Traffic accident rate · Drivers behavior in Abu Dhabi

1 Background

Despite the significant number of researches in traffic accidents and road safety topics, the studies that addressed the interaction relationship between the drivers' behavior in terms of their historical records of traffic violation and accidents involvement are relatively few in the literature. Peck and Kuan found that person-centered driving record variables (such as prior record of violations and

M. Shawky (✉) · Y. Al-Badi · I. Sahnoon · H. Al-Harthi
Traffic and Patrol Directorate, Abu Dhabi Police, Abu Dhabi, United Arab Emirates
e-mail: m_shawky132@hotmail.com

M. Shawky
Faculty of Engineering, Ain Shams University, Cairo, Egypt

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_62

accidents, age, gender, socioeconomic) and driving exposure made unique contribution to accident prediction [1]. In addition, driving record variables and exposure measures were approximately equally efficient as accident prediction. Hauer et al. [2] showed that the information of at-fault accidents is an efficient way in identifying the drivers who have a high risk of accident in the future. The same findings were pointed out by applying regression model of about 1,998,347 British Columbia drivers' records [3]. It was found that the relative importance of various violation types in predicting future culpable accident risk and thus can sever as a basis for setting penalty point levels.

Two studies were conducted in California that aimed to assess the accuracy of predicting future accidents risk using various combinations of demographic and prior driving records variables [4, 5]. Approximately 140,000 records of licensed drivers including their age, gender and driving record variables were analyzed. About 17 regression models were developed using various combinations of demographics and prior driving variables. The results indicated that the models that use prior total accidents as a predictor variable of future accidents perform better than other models that do not use total accident as a predictor and the same results was found for using historical violation. It also found that adding demographic variables of the drivers to the model increase the performance of the model. Accordingly, the total historical accidents and violation records can be used as a measure to predict high-risk drivers in order to intervene before this risk is realized, which support the current point count strategy in California which attempts to optimize the identification of driver having a high probability of subsequent accidents involvement.

Daigneault et al. [6] attempted to determine accident predictors for older drivers using prior violations or accidents and concluded that prior accidents are a better predictor for accident risk than prior violations. Chandraratna et al. [7] developed an accident prediction model and showed that drivers that had a recent accident history, very young or very old, males and drivers with both speeding and non-speeding citations have higher likelihood to be at fault in a accident. Zhang et al. [8] investigated the associated risk factors with traffic violations and accident severity in China. It was found that traffic violations are one of the major risks threatening road safety. In addition, drivers have 2 years or less driving experience display a significant higher risk of traffic violations.

Other studies addressed the impacts of a certain traffic violation type on the accident rates such as speeding, alcohol usage, set belt, red light violations, etc. For example, the speeding behavior (as a measure by violation conviction) was utilized to identify drivers as high risk for becoming culpably involved in speed-related accidents [9]. The results identified a clear distinction between the conviction categories of "exceeding the speed limit" and "excessive speed" in terms of these accident-violation relationships. In addition, Dissanayake and Lu [10] and Baker et al. [11] proved that the seat belt usage and alcohol usage have major effects on the severity of accidents.

Alver et al. [12], explored the interaction between socio-demographic characteristics of traffic rule violators (four types of traffic violations records were aggregated) and accident history for young drivers (18–29 years old) by applying binary logit models. The analysis showed that 23.9 % of drivers were involved in at least one traffic accident in last three years. This accident rate increases to 38.3 % for those who received at least one traffic citation/violation in last three years and peaks to 47.4 % for those who were fined for seat belt violations.

Self-reported questionnaire approach was used by some researchers to investigate the relationship between violation records and accident involvements of drivers [e.g., 13–16]. For example, Winter and Dodou [13] conducted a study to investigate the relation of errors and traffic violations to accident involvement using survey of more than 45,000 respondents. The results, through a meta-analysis, showed that violation predicated accidents with an overall correlation of 0.13. It also provided the validity of the driver behavior questionnaire to be used by researchers and road safety practitioners who seek to obtain insight into driving behavior of a population on interest.

A recent study [17] presented a comprehensive review in the previous published studies from 1970 to 2014 that addressed the relationship between impulsivity of the drivers and at least one driving related outcomes (e.g., a self-report measure of driver behavior) were included. The definition of the impulsivity has been listed and it can be easily defined as the “tendency to act with little forethought, without deliberation and evaluation of consequences”. About 38 studies out of 288 studies are reported, however, the studies that tackled the interaction of the traffic offences as a measure of impulsivity and the accidents are about 5 studies only.

2 Data Preparation

The employed data in this paper was extracted from the Emirate of Abu-Dhabi (AD) the capital of UAE traffic police databases during five years from 2010 to 2014. The required data were integrated from different sets of databases; (1) traffic violation data, (2) traffic accident data and (3) driving license information. Both the property damage only (POD) accidents and severe accidents (i.e., any accident with at least one injury of fatality) were used in the research. Any driver has at least one record in the violation and/or accidents database systems during the five years was involved in the analysis process.

Around 1.1 million registered driving licenses exist in the driving licenses database. However, about 83 % of them were found in one or both of the data base systems (i.e. violation or accident database). That means 83 % of the drivers involved in traffic violation or/and accidents. Due to the dynamic change in the drivers’ population in AD (about 85 % of the drivers are not local drivers) the drivers’ information and records that have at least one record of traffic violation and/or accident were involved in this study. Accordingly, about 713,583 drivers’ records were examined with a total number of traffic violations and accidents of 2,762,011 and 690,697, respectively.

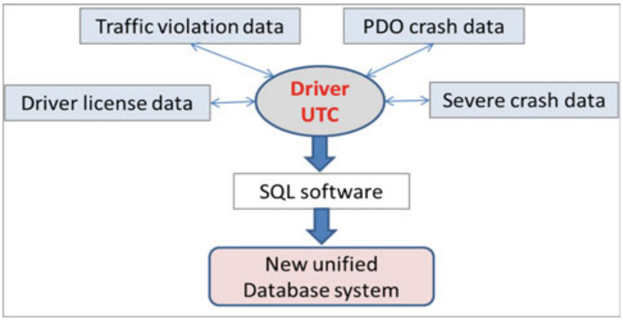


Fig. 1 Data sources and preparation methodology

An integration process among the different database systems was conducted by using a Unified Traffic Code (UTC) which is a unique number that has been given to each driver when he/she issues the diving license in United Arab of Emirates (UAE). Also any driver uses another country/international driving license, like tourists, has also given a UTC. Based on this unique number of each driver, a new unified database system was created by using SQL software package. As shown in Fig. 1, four sources of data were utilized to crate the required data of this study.

3 Data Analysis

3.1 Accident and Violation Frequencies

The data analysis addressed a total number of accident involvements of 690,697 during the 5-year period, among these about 8602 severe accidents (i.e., an accident with at least one injury) and 682,095 PDO accidents. Table 1 shows the frequency of at-fault drivers involved in these two types of traffic accidents. An about 46.5 % of the drivers involved in at least one PDO accident against 1.15 % involved in at least one severe accident. The average rate accident involvement per driver is 0.968 accidents during five years. It is mean about 193 accidents per 1000 driver per year and 2.4 severe accidents per 1000 driver per year. Figure 2 shows the frequency of at-fault drivers based on their number of accident involvements record during the five years.

A total number of violation involvements of 2,762,011 are included in the analysis. Among these, about 979,859 violations are considered as hazard violation class based on AD traffic police classifications. Hazard violation list includes about 25 types of violations out of 173 violations listed in the traffic law in AD. The 25 hazard violations includes for example aggressive driving, exceed the limit speed

Table 1 Frequency of at-fault drivers involved in traffic accidents

No. of accidents	PDO accidents		Severe accidents	
	Number	Percentage (%)	Number	Percentage (%)
0	414,354	53.5	765,667	98.85
1	187,419	24.2	8721	1.13
2	88,408	11.4	152	0.02
3	41,793	5.4	4	0.00
4	20,058	2.6	1	0.00
5	10,148	1.3	0	0.00
>5	12,365	1.6	1	0.00

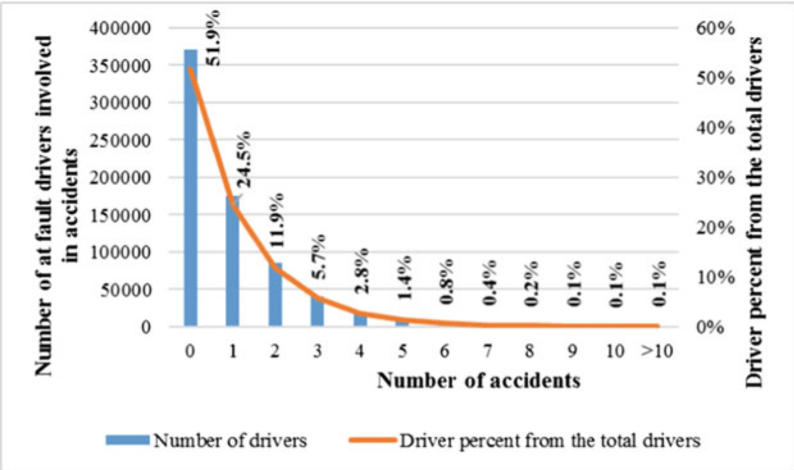


Fig. 2 Frequency of at-fault drivers involved in accidents at five years

by 60 kph or more, read light crossing, sudden lane changing, tailgating, alcohol usage and mobile usage during driving, etc. The recorded violations indicate that the violation rate per driver over five years is 3.989 and about 0.798 violation per driver per year. In addition, violation rate of the hazard violation class is 1.373 violations per driver over five years and 0.275 violations per driver per year.

Figure 3 shows the frequency of drivers involved in traffic violations during the period of five years. In addition, Fig. 4 shows the number of drivers in terms of their record of violations and classed based on the number of accidents involvements. It shows that the total number of drivers significantly decrease with increasing the number of violations record and with increasing the number of accident involvements.

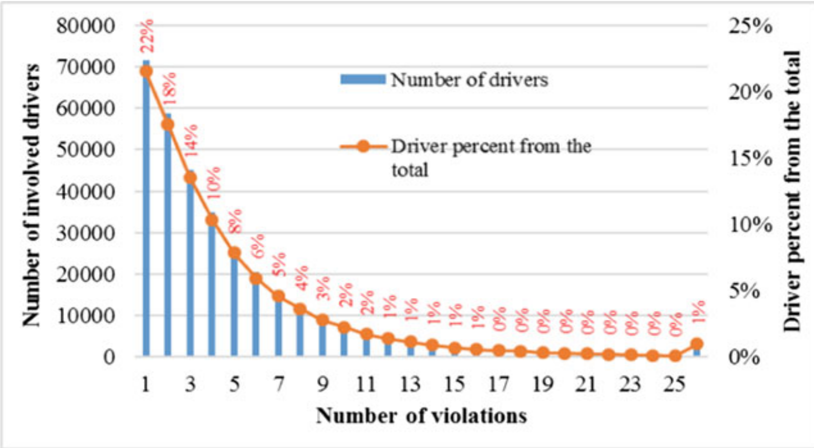


Fig. 3 Frequency of drivers involved in traffic violations at five years

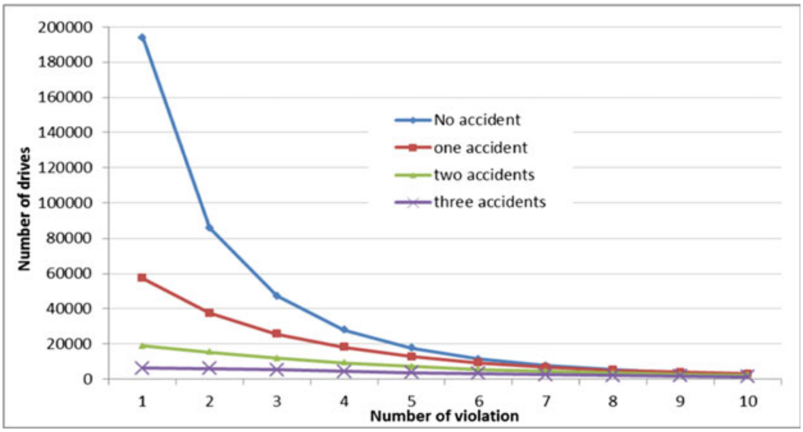


Fig. 4 Number of drivers at different violation records and accident number

3.2 Accident Rate Per Driver Estimation

The accident rate per driver was calculated as the ratio between the total numbers of accidents of at-fault drivers who have a given violation records to the total number of drivers who have the same violation records. Figure 5 shows the accident rate per driver in terms of his/her historical violation records. It shows a very strong correlation (R^2 = 0.991) between accident rate and violation records at a driver base calculations. The accident rates of drivers significantly increase with increasing the historical records of traffic violations. The same analysis was conducted in terms of

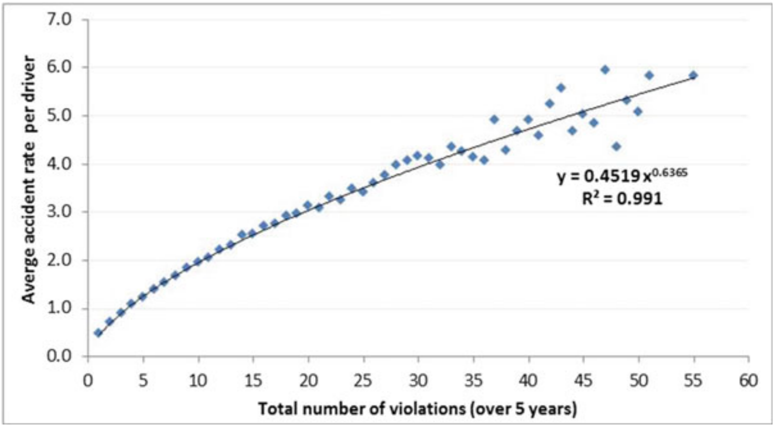


Fig. 5 Accident rate per driver based on the historical total violation records during five years

the number of hazard violations and the violations that have penalty points and the total number of penalty points of the driver given in five years.

Figure 6a–c show the relationship between accident rate per driver in terms of three more variables; (1) number of hazard violations, (2) number of violation that have penalty, (3) aggregated number of penalty points of the driver of five years. It shows a strong correlation between accident rate of the drivers and these three variables based on the logistic regression fitting analysis. Figure 6d shows the estimated accident rate based on the fit model of the three variables. This figure indicates that the accident rate at a certain number of hazard violations is significantly high compared by the same number of other violations which support the selectin of the 25 violation as hazard of traffic rule violation class in AD Traffic police.

3.3 Accident Probability Estimation

The probability of a driver to be involved in a traffic accident can be calculated as the ratio (or the percentage) of the drivers who were involved in an accident to the total number of drivers at a given number of violation involvement. Figures 7 and 8 show the percentage of drivers with accidents in terms of the total number of traffic violations and the numbers of hazard violations, respectively. These figures show that the logit regression model strongly fit the data (R^2 = 0.99 and 0.98), which mean that the logistic regression model will be the best model to be used for estimating the accident probability per driver as a function of his/her historical violation records.

Figure 9 shows a comparison between the estimated percentage of drivers involved in accidents based on two different variables; total violations and hazard

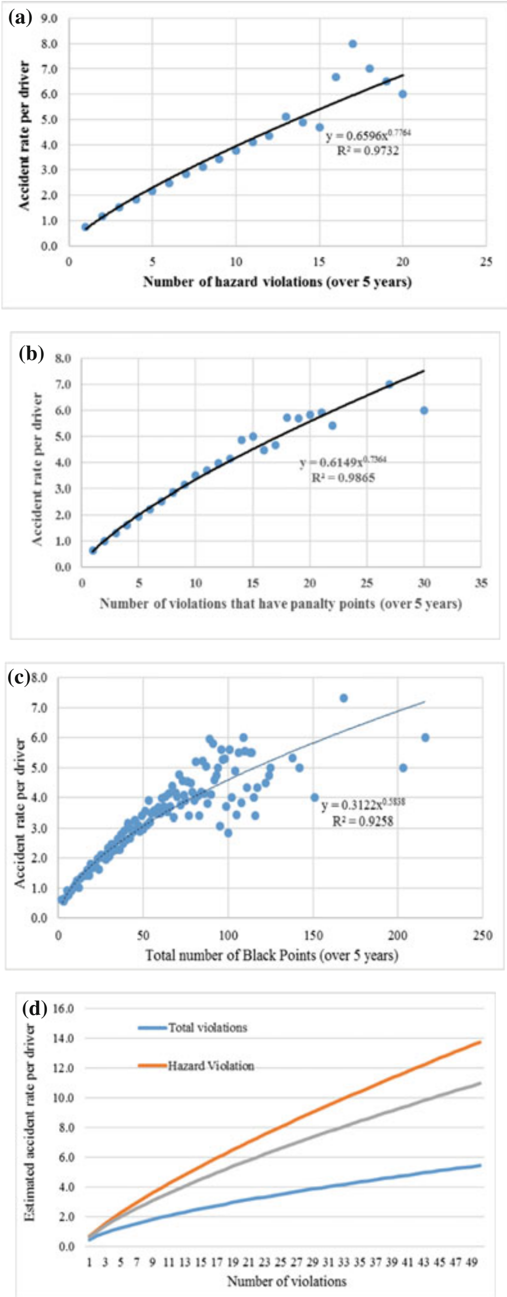


Fig. 6 Accident rate per driver based on the historical total violation records during five years. **a** In terms of hazard violations. **b** In terms of violations that have penalty points. **c** In terms of total number of penalty points. **d** Comparison between different violations categories

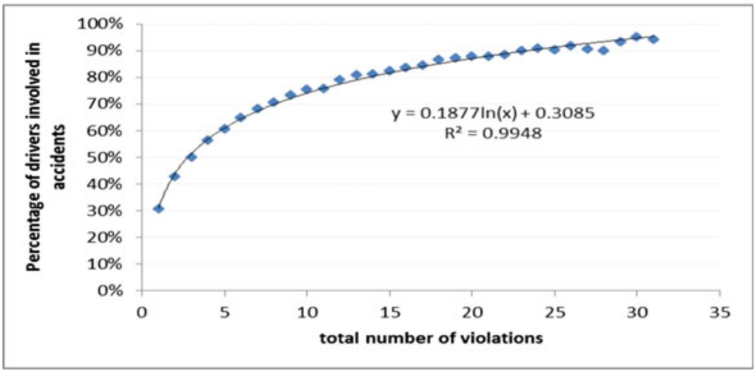


Fig. 7 Percentage of drivers involved in accidents with their total number of violations

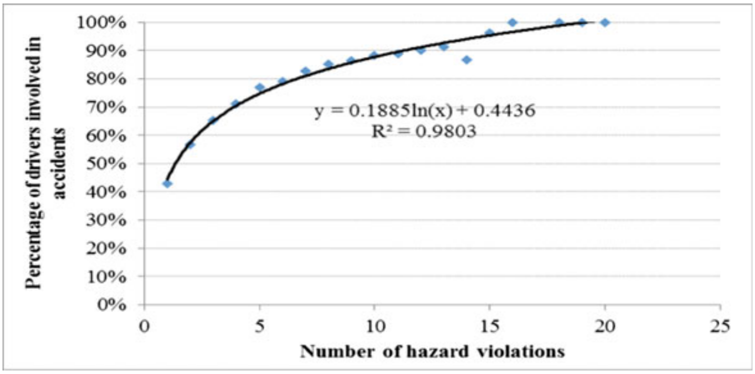


Fig. 8 Percentage of drivers involved in accidents with their hazard violation records

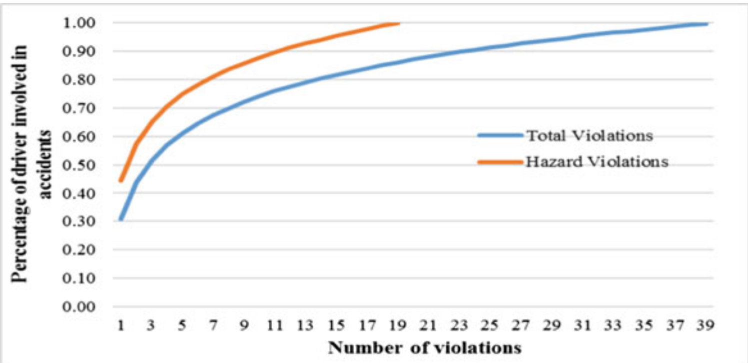


Fig. 9 Estimated percentage of driver involved in accidents

violations. It shows, at a given number of violations, the drivers who have hazard violations are most likely to be involved in accidents in the future compared to the other violation types.

4 Conclusion

The purpose of this paper was to investigate the relationship between the drivers' behavior in terms of their historical traffic rule violation records during the past five years (2010–2014) and their accident involvement records at the same period. Accordingly, the drivers who are mostly to be involved in the accident in the future can be identified and countermeasures against these drivers should be taken. Data from different sources were integrated to create a comprehensive database system which was required in the analysis process. About 713,584 drivers record were used.

The results showed strong relationships between accident rate per driver and the total number his/her violation records. In addition, the likelihood to be involved in an accident significantly increases with increasing the total number of historical violation records. These results are consistent with the results of prior studies that were addressed the same to by using data from other counties.

In conclusion, the results proved that the violation records can be used as a good variable to predict the future accident involvements of drivers, especially the violations that are considered as hazard violations in AD traffic police.

The authors are currently working to extend the data analysis for a future research in order to test more factors that may affect the accident rate per driver; such as demographic characteristics of the drivers (age, gender, nationality, experience), severe accident involvements, type of the violation (face to face or absent), etc. The logistic regression modeling or other advanced modeling approaches will be applied to find the most appropriate variables that can be used to predict the driver accident risk.

References

1. Peck, R.C., Kuan, J.A.: Statistical model of individual accident risk prediction using driver record, territory and other biographical factors. Accid. Anal. Prev. **15**, 371–393 (1983)
2. Hauer, E., Persaud, B.N., Smiley, E., Duncan, D.: Estimating the accident potential of Ontario drivers. Accid. Anal. Prev. **23**, 133–152 (1991)
3. Chen, W., Cooper, P., Pinili, M.: Driver accident risk in relation to the penalty point system in British Columbia. J. Saf. Res. **26**(1), 9–18 (1995)
4. Gerbers, A.M.: Strategy for Estimating Drives Accident Risk in Relation to California's Negligent-Operation Point System. Report No. CAL-DMV-RSS-99-183, California Department of Motor Vehicles, Research and Development Section (1999)

5. Michael, A.G., Raymond, C.P.: Using Traffic Conviction Correlates to Identify High Accident-Risk Drivers. Report No. CAL-DMV-RSS-00-187, California Department of Motor Vehicles, Research and Development Section (2000)
6. Daigneault, G., Joly, P., Frigon, J.: Previous convictions or accidents and the risk of subsequent accidents of older drivers. Accid. Anal. Prev. **34**, 257–261 (2002)
7. Chandraratna, S., Stamatiadis, N., Stromberg, A.: Accident involvement of drivers with multi accidents. Accid. Anal. Prev. **38**, 532–541 (2006)
8. Zhang, Guangnan, Yau, K.W., Chen, G.: Risk factors associated with traffic violations and accident severity in China. Accid. Anal. Prev. **59**, 18–25 (2013)
9. Cooper, P.J.: The relationship between speeding behavior (as measure by violation convictions) and accident involvement. J. Saf. Res. **28**(2), 83–95 (1997)
10. Dissanayake, S., Lu, J.: Analysis of severity of young driver accidents: sequential binary logistic regression modeling. Transp. Res. Rec. **1784**, 108–114 (2002)
11. Baker, S.P., O'Neill, B., Ginsburg, M.J., Li, G.: The Injury Fact Book. Oxford University Press, New York (1992)
12. Alver, Y., Demirel, M.C., Mutlu, M.M.: Interaction between socio-demographic characteristics: traffic rule violations and traffic accident history for young drivers. Accid. Anal. Prev. **72**, 95–104 (2014)
13. Winter, J.C.F., Dodou, D.: The driver behavior questionnaire as predictors: a meta-analysis. J. Saf. Res. **41**, 463–470 (2010)
14. Winter, J.C.F., Kyriakidis, M., Dodou, D, Happee, R.: Using crowd flower to study the relationship between self-reported violations and traffic accidents. In: Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics (AHFE) (2015)
15. Obeid, S., Gitelman, V., Baron-Epel, O.: The relationship between social capital and traffic violations: Israeli Arabs as a case study. Accid. Anal. Prev. **71**, 273–285 (2014)
16. Mohammadi, G.: Prevalence of seat belt and mobile phone use and road accident injuries amongst college students in Kerman. Iran. Chin. J. Traumatol. **14**(3), 165–169 (2011)
17. Biçakisia, P., Özakan, T.: Impulsivity and driver behavior, offences and accident involvement: a systematic review. Transp. Res. Part F (2015)

Acquisition System Improvement and Results Analysis in Small Overlap and Oblique Tests

Núria Parera and Alba Fornells

Abstract In this paper two crash tests are studied: the small overlap test introduced by IIHS (Insurance Institute for Highway Safety) and the moving deformable barrier (RMDB) introduced by NHTSA (National Highway Traffic Safety Administration). Several modifications have been made to both barriers in order to obtain better data from the acquisition system and be able to improve in-depth analysis of the crash tests and vehicle structures. Both barriers have been instrumented with load cells systems, and for the validation of the acquisition system two tests were carried out, one with the IIHS small overlap barrier and a second one with the RMDB barrier in an oblique test. These two tests were performed with the same car and following the test procedure specifications. The results obtained were processed, analysed, compared and some conclusions were made.

Keywords Impact test · Data analysis · Barrier · New configuration impacts

1 Introduction

In 2012 IIHS started evaluating vehicle crashworthiness with the small overlap tests due to a recent crashworthiness study with new vehicles in the EU and USA that found a higher severity in frontal crashes. The result showed that occupant injuries and vehicle structure deformation were severe when the vehicle was loaded outboard (small overlap). The introduction of the IIHS small overlap test configuration barrier implies a more severe and critical load case at the vehicle's structural parts that need to be improved in order to improve safety in those specific crash cases. The small overlap test consists of a crash between the vehicle against the edge of the static and rigid barrier at 64.4 km/h with a frontal overlap of 25 %. In order to deal with problems of cage intrusion, two strategies were followed by manufacturers to dissipate impact energy: make the vehicle slide away from the barrier with

N. Parera (✉) · A. Fornells
Applus IDIADA, L'Albornar, PO Box 20, E-43710 Santa Oliva (Tarragona), Spain
e-mail: nuria.parera@idiada.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_63

lateral translation or make it rotate around the corner of the barrier. For the purpose of obtaining better information of the crash test and improving the vehicle response and energy absorption strategies, the barrier corner was instrumented with load cells.

NHTSA (National Highway Traffic Safety Administration) has introduced a new oblique test configuration presenting a critical new load case that manufacturers are on the way to solving. This test consists of an oblique crash of a moving deformable barrier (RMDB) against a static vehicle with an overlap of 35 %. In order to provide the best devices for passive safety development and enable the analysis of the loads transmitted to the barrier, the barrier has been instrumented with load cells. The registered data will enable the identification of the vehicle's elements that take part in the absorption of energy during the crash and are a valuable tool to improving vehicle safety by comparing the loads transmitted to the barrier in oblique tests.

2 Small Overlap Test

The first barrier built at IDIADA's Passive safety laboratory, had a matrix of load cells of 8 × 4. As can be seen in Fig. 1, the barrier was divided by columns (A, B, C, D) and lines (from 1 to 8). This division was made in order to know the exact position of each load cell and to have better knowledge of their interaction with structural parts of the vehicle during the crash test.

The vehicle's parts that have more interaction with the barrier are as follows: The A-pillar (Upper and Lower) is represented in Fig. 2 in red. The A-pillar together with the firewall delimits the zone of the engine with the occupant's one. It is important not to have big structural deformations in this zone in order to accomplish a better occupant protection. The Shotgun in Fig. 2 is represented in

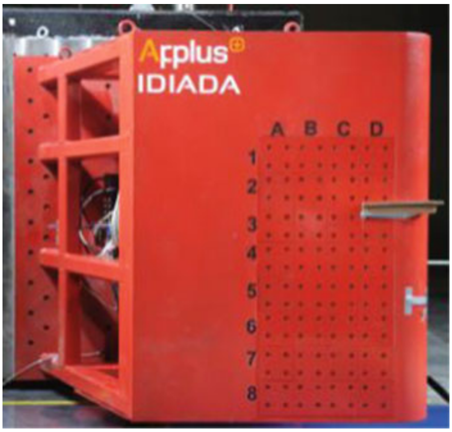


Fig. 1 IDIADA's small overlap barrier

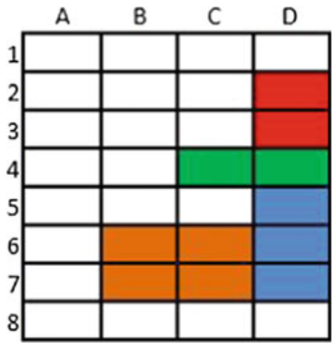


Fig. 2 Vehicle's structural part colour association with the load cells

Green and has been designed to have deformation in order to absorb part of the crash energy. The Sill beam is represented in Fig. 2 in Blue. It is a horizontal structure at the lower part of the vehicle that adds stiffness to the vehicle's structure and the Wheel is represented in Fig. 2 in Orange. The wheel is very important in the small overlap test as its dynamics can force the vehicle to change its interaction with the barrier. Also, the intrusion of the wheel at the occupant's zone can cause severe injuries to the lower extremities.

2.1 *Test with the Barrier and Results Analysis*

A test was carried out at IDIADA's laboratory in order to validate the acquisition system. The vehicle used was a Jeep Grand Cherokee Second generation, with a mass of 1856 kg and a height of 1762 mm. The vehicle crashed against a static and rigid barrier at 64.4 km/h with a frontal overlap of 25 % (Fig. 3).

Fig. 3 Vehicle tested for the small overlap test

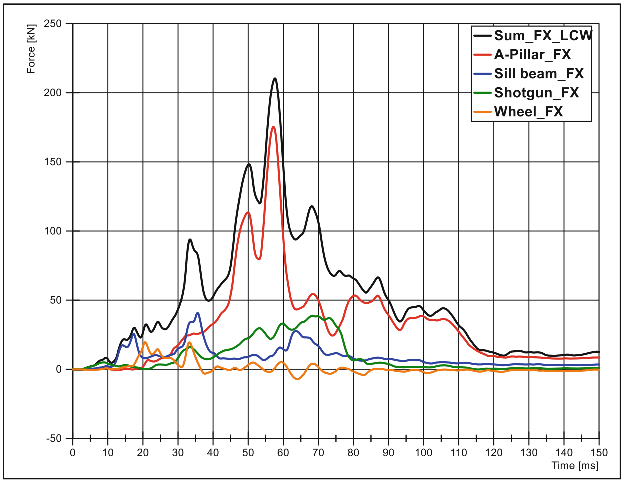


Fig. 4 Forces registered by the barrier

As shown in Fig. 4, the results from the barrier were charted by structural parts and the total sum of the forces was calculated in order to know the total force registered by the acquisition system during the entire test.

The forces registered by the barrier were compared in Fig. 5 with the forces measured by the vehicle's accelerometers (F = m * a). The results showed that the vehicle's force was much greater than the one registered by the barrier, especially at 60 ms where the differences are extremely large.

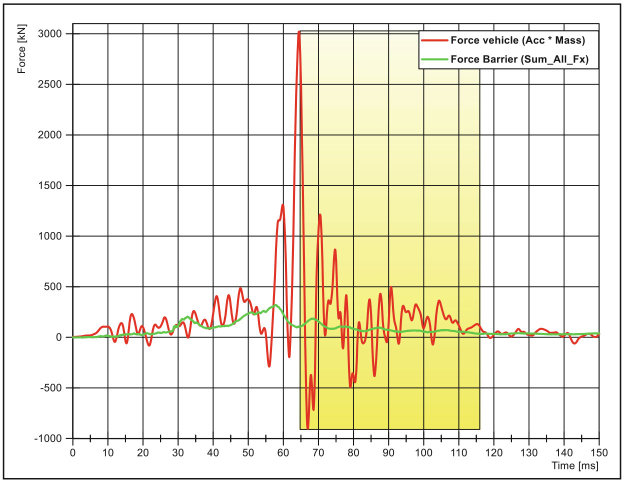


Fig. 5 Comparison between the barrier forces registered and the vehicles' ones

Table 1 Numerical results from energy calculations

Energy	Result (kJ)
Maximum vehicle's energy deformation	286.16
Maximum barrier's energy registered	142.04
Gap between the two energies	144.12

In Fig. 5, a difference of forces can be seen at 60 ms where the barrier forces are small compared with the vehicle. At 60 ms the vehicle is interacting with the corner of the barrier, which was not instrumented, in order to verify this assumption more energy calculations were carried out. In order to know the energy that was registered by the barrier several calculations were done. First of all the initial energy at the moment the vehicle collides with the barrier was calculated. Assuming that the moments of the inertia will be a really small value compared with the kinetic energy of the vehicle (mass * velocity vehicle), the final resultant formula was as follows:

$$E_c = \frac{1}{2} m V^2_{coG} + \cancel{\frac{1}{2} I_{wheel} \omega^2_{wheel} + \frac{1}{2} I_{other} \omega^2_{other}} \quad (1)$$

The numerical results are shown in Table 1 as follows:

The energy gap (144.12 kJ) proves that there is a problem with the acquisition system of the barrier. The gap should not be so big, but also, the two energies will not ever be equal due to other mechanisms that dissipate energy that are not reflected by the data collected at the crash test. Those mechanisms are for example heat [1, 2] and friction, among others. The results from the formula are shown in Fig. 6, where the energy calculated by the barrier acquisition system is compared with the energy registered by the vehicle's accelerometers.

The difference between the forces registered in the barrier and the forces registered by the vehicle sensors can be explained by the fact that the corner of the

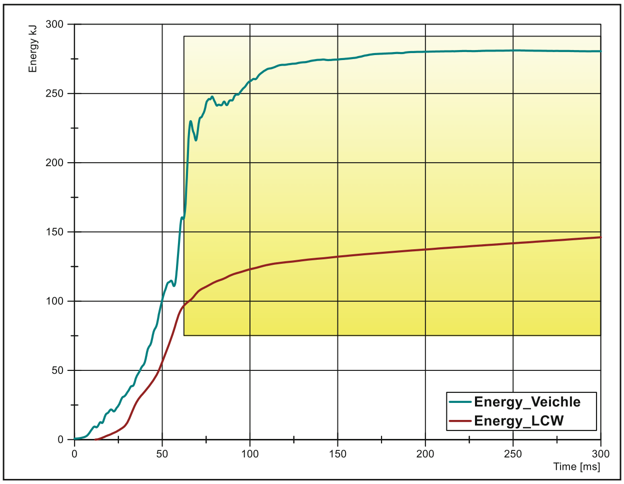


Fig. 6 Energy comparison between the barrier and the vehicle

barrier was not instrumented. In the small overlap test, the vehicles have a high interaction with the corner of the barrier after the first few milliseconds and consequently an important percentage of the crash loads were not registered by the barrier.

2.2 New Barrier with the Corner Instrumented

In order to solve the issue mentioned above related with the acquisition system, a new barrier was developed which also had the corner instrumented with triaxle load cells. This will make it possible to obtain more data from the crash test during the phase where the vehicle impacts the corner, approximately at 60 ms of the test. The gap between the energy lines should be reduced with this new barrier. The new barrier has a load cell matrix of 8 × 5, as shown in Fig. 7.

Following the same procedure with the new barrier, a validation test was carried out in order to assess the new acquisition system. The vehicle was a different model, so the results were not the same and the comparison was made by a force diagram at the same millisecond as shown in Figs. 8 and 9.

Fig. 7 New small overlap barrier with the corner instrumented

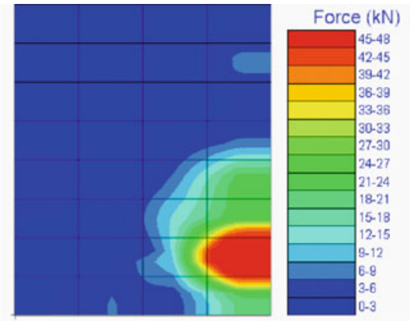


Fig. 8 Old barrier force diagram at 69 ms

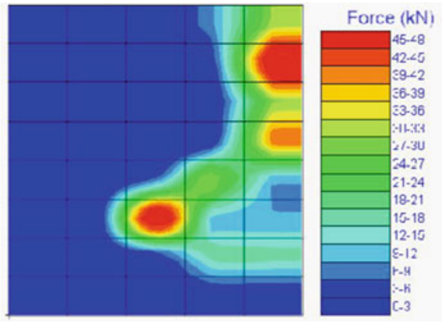


Fig. 9 New barrier force diagram at 69 ms

Although the vehicles were different and each vehicle structure interacts in a different way with the barrier, it can be seen in Fig. 9 that the corner load cell column shows forces that would not be noticed if the test had been done with the old barrier.

Calculations were made with the results of the barrier validation test in order to obtain the energy of the new barrier and the vehicle. Figure 10 shows the energy comparison between the new barrier and the vehicle.

Figure 10 shows that the gap between the two lines is smaller. These results corroborate that the instrumentation of the barrier's corner was the solution for the data loss with the old barrier. Whereas the 100 % correlation of the lines will not be possible due to other dissipation mechanisms that take place at the crash test, this result is given as correct and the results of the acquisition system validation accomplish the aim of improving the barrier.

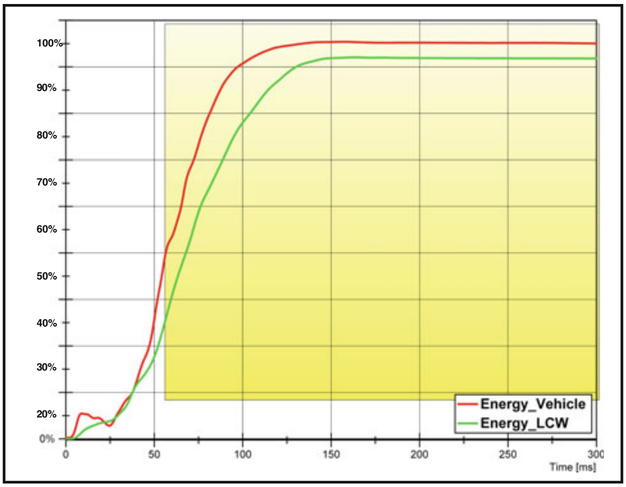


Fig. 10 Energy comparison between the new barrier and the vehicle tested

3 RMDB Acquisition System

For the study a full RMDB was developed following the specifications used by Saunders et al. [3] and the RMDB Manual and Drawing Package publicly available in [4]. The acquisition system for the RMDB barrier consisted of the installation of a load cell wall, Fig. 11. Thirty two triaxle load cells are located between the aluminium back plate and the barrier's frame, however only 18 were installed in the impact area (blue zone of Fig. 11).

3.1 Test with the Barrier and Results Analysis

The validation test was carried out with a second generation Jeep Grand Cherokee with a mass of 1856 kg and a height of 1762 mm. A Hybrid III 50 % was used, but no data analysis was made. The test speed recorded was 90.15 km/h (Fig. 12).

Figure 13 shows how the velocity during the crash test develops. It can be seen that the velocity of the barrier is transmitted at the stopped vehicle at the moment of the crash. At the end of the crash test the vehicle reaches a velocity of 14 m/s and the barrier' velocity is reduced to 13 m/s.

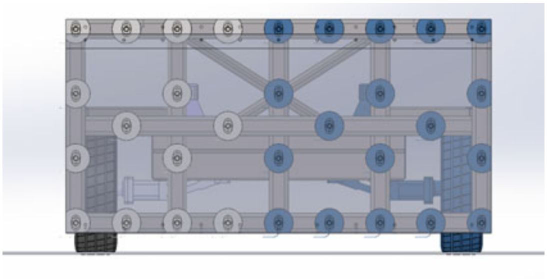


Fig. 11 IDIDA RMDB load cell wall

Fig. 12 Image of the RMDB validation test at 77 ms

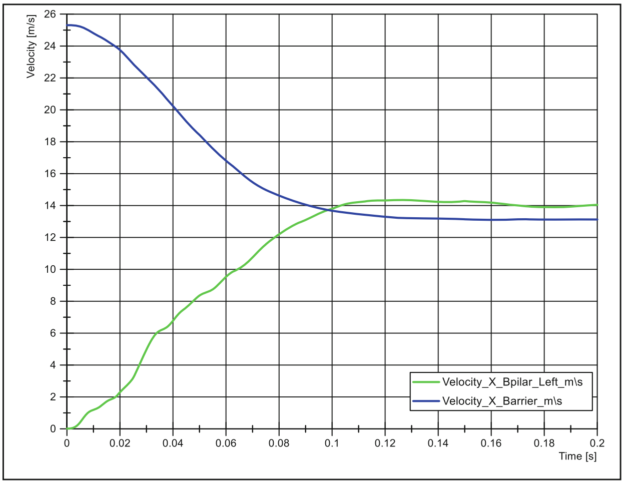


Fig. 13 Development of Velocity during the crash test

The load cells were numbered and named in order to be able to correlate the data with the structure of the vehicle (Fig. 14).

The force calculation is shown by sections (E, F, G, H, I) in Fig. 15. The values of the G and H sections are higher than the others as load cells G1, H1 and I1 are in contact with lower structural parts of the vehicle such as the sill beam. These load cells will register the role of the wheel during the crash and its interaction with stiffer structural elements.

The load cells G2, G3, H2 and I3 will show results from the contact of the barrier and the vehicle's elements such as the bumper beam. In general, it can be observed that the inner part of the barrier is loaded later due to the deformations and impact angle.

The calculation of the total longitudinal (barrier's X-direction) force applied to the barrier is shown in Fig. 15 and the force peak is 350 kN.

In order to know the energy that was registered by the barrier several calculations were done. First of all, the initial kinetic energy of the barrier and the vehicle

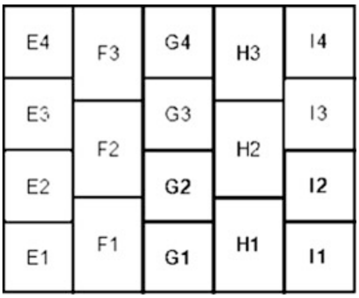


Fig. 14 Distribution of the load cells in the LCW

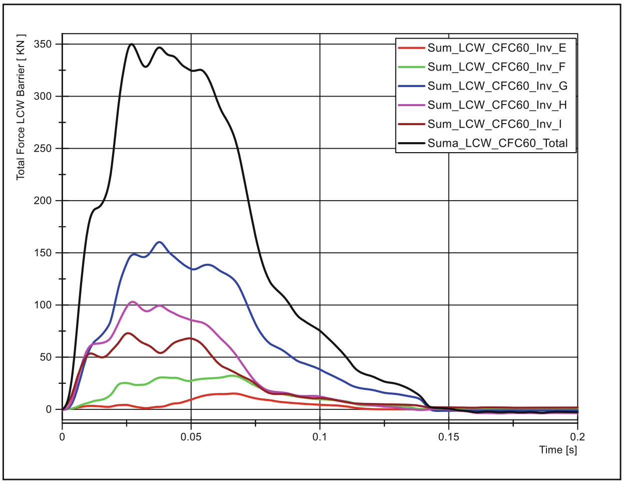


Fig. 15 Load cell wall force by section

was calculated and then the final kinetic energy of the system. The moments of the inertia were really small values compared with mass * velocity vehicle and due to that they were not taken into account, the final resultant formula was as follows:

$$E_{cinitial} = \tfrac{1}{2}\, m_{barrier} V^2_{barrier\ coG} + \cancel{\tfrac{1}{2} I_{wheel}\, \omega^2_{wheel}} + \cancel{\tfrac{1}{2} I_{other}\, \omega^2_{other}} = 778.125\ \mathrm{kJ} \tag{2}$$

The final kinetic energy was calculated by the following formula:

$$E_{cfinal} = 1/2\, m_{barrier} V^2_{barriercoG} + 1/2\, m_{vehicle} V^2_{vehiclecoG} = 391.705\ \mathrm{kJ} \tag{3}$$

The difference between the initial and final kinetic energy was calculated in order to know which amount of the crash energy was applied at the deformation of the vehicle and the barrier.

$$Ec_{initial} = Ec_{final} \tag{4}$$

$$1/2\, m_{barrier} V^2_{barrier\ coG} = 1/2\, m_{barrier} V^2_{barrier\ coG} + 1/2\, m_{vehicle} V^2_{vehicle\ coG} + E_{deformation} \tag{5}$$

The results from the formula are shown in Table 2.

Table 2 Numerical results from energy calculations

Energy	Result (kJ)
Energy deformation	386.42
Initial kinetic energy barrier	778.125
Final Kinetic energy barrier	210.450
Final kinetic energy vehicle	181.30
Barrier's kinetic energy difference	567.72

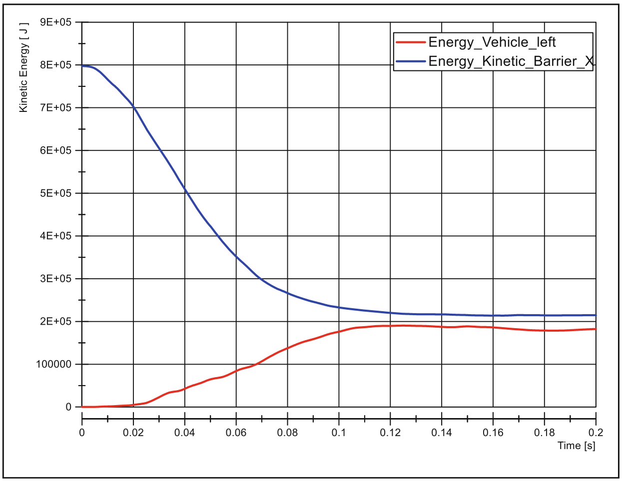


Fig. 16 Graphic of the test's kinetic energies

The results of the kinetic energy calculations are shown in Fig. 16.

Figure 17 shows the results of the amount of energy consumed per unit of time. A maximum load cell wall energy peak of 447 kJ, a maximum peak for the barrier of 580 kJ and a maximum vehicle's energy peak of 190 kJ can be seen.

The 447 kJ registered by the load cell wall of the barrier in comparison with the kinetic energy calculation of deformation (386.47 kJ), should be equal, but as mentioned before, the inertia moments and rotations of the vehicle were not taken

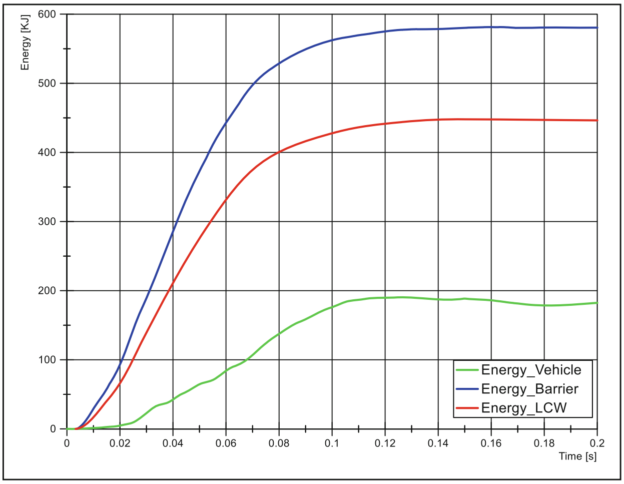


Fig. 17 Graphic of the amount of energy consumed per unit of time

Fig. 18 New barrier acquisition system distribution

into account. Also, other mechanisms of energy dissipation were not present in the calculations, such as energy dissipated in the form of heat and friction. The 60.58 kJ that is the difference between both energies could be associated with those mechanisms of energy, moments and inertias not taken into account in the calculation. Due to that, the validation of the barrier was given as good.

3.2 New Acquisition System Design

The results obtained in this study showed that the developed barrier is suitable for performing the oblique tests according to the procedure developed by NHTSA. However, some potential improvements were identified.

The main areas of improvement were the distribution of the load cells and the post-crash braking manoeuvres. The current non-uniform distribution of the load cells, adapted to the proposed barrier structure, increases the complexity of the analysis of the data. For this reason, a new design was done in order to have a more uniform distribution of the load cells, as shown in at Fig. 18.

The next step will be a validation test with a Thor dummy, in order to be able to make a comparison with the results of the Hybrid III crash test that had been carried out in the laboratory. Furthermore, the test will validate the new load cell distribution of the barrier.

References

1. Kostina, A., Iziumova, A., Plekhov, O.: Energy dissipation and storage in iron under plastic deformation (experimental study and numerical simulation). Fract. Struct. Integrity Ann 2014 **8** (2014) (Frattura ed Integrità Strutturale: Annals)
2. Kapoor, R., Nemat-Nasser, S.: Determination of temperature rise during high strain rate deformation. Mech. Mater. **27**(1), 1–12 (1998)

3. Saunders, J., Craig, M., Parent, D.: Moving deformable barrier test procedure for evaluating small overlap/oblique crashes. SAE Technical Paper 2012-01-0577 (2012) doi:10.4271/2012-01-0577
4. User's Manual: Small Overlap/Oblique Moving Deformable Barrier U.S. Government Property: DOT-NHTSA V22129

Part IX
Road and Rail—Warning Systems/Public Transport

Effects of Intersection Collision Warning Systems and Traffic Calming Measures on Driver's Behavior at Intersections

Manuel Silvestri and Francesco Bella

Abstract The objective of this study was to examine the effect of intersection collision warning systems (ICWSs) and traffic calming measure on drivers' behavior, in response to a potential conflict event at the intersections, which constitute a crucial point with respect of the road safety. The drivers' behavior was analyzed by means of a multivariate variance analysis (MANOVA) procedure. ICWSs were the auditory speech message and the visual warning. Both ICWSs provided to the driver the direction of the violator vehicle. The traffic calming measure was the dragon teeth. Results show that ICWSs help the drivers' to detect earlier the violator vehicle and act a safer braking maneuver to avoid the conflict at the intersections. For the traffic calming measure no statistically significant effects were found. However, for this condition, a decrease of about 1.3 km/h of the minimum speed value reached by the driver to avoid the collision was recorded.

Keywords Intersection collision warning system · In-vehicle devices · Driving simulation

1 Introduction

The intersections are essential elements of the road network but constitute hazardous locations, because imply opportunities for conflicts among vehicles. Although intersections are a slight part of the road system, they emerge as the road sections where a remarkable portion of the accidents occurs [1–4].

There is agreement to believe that this situation is linked to the fact that driving at intersection is one of the most dynamic and difficult task of drivers e.g. [5]. It

M. Silvestri (✉) · F. Bella
Roma TRE University, Vito Volterra N. 62, 00146 Rome, Italy
e-mail: manuel.silvestri@uniroma3.it

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_64

requires large cognitive efforts by the driver to perceive and process the amount of information related to the specific intersection configuration (type di intersection, traffic signs), the traffic condition (crossing vehicles, vehicles driving ahead and oncoming) and the maneuver to act (crossing, turn on the left or right). A such complexity often implies inadequate drivers' behavior and then occurrence of accidents.

Understanding the main factors that influence the occurring of the intersection accidents and developing systems that encourage proper drivers' behaviors and help him in the complex task of drive at the intersections, are deemed to be the keys to improve the road safety at intersections.

For this reason, a lot of research were and continue to be aimed on the factors contributing to crashes at these hazardous locations e.g. [4] and on the development of effective driving assistance systems, such as intersection collision warning systems (ICWS), and on the development of effective traffic calming measures.

Among traffic calming measures, the perceptual measure are particularly interesting. These measures (such as peripheral pavement markings—e.g. dragon teeth, combs—transversal pavement markings) attempt to increase the workload or the perceived risk by the driver through subtle (non-intrusive) change to the road environment while approaching the intersection.

The intersection collision warning systems (ICWSs) are in—vehicular warning systems, which detect obstacles with sensors in vehicles, and devices located at intersection, such as detecting radar and alert the driver of an imminent collision. These systems have an important impact on driving safety because making the potential collision at intersection predictable, by allowing the decrease of the probability and severity of accidents [6–8].

Among the several types of alarms (auditory warnings, visual warnings, vibrotactile warnings and haptic warnings), those most used concerning the auditory and the visual stimulus. The first type of alarm consists in audio signals as beep sounds, auditory icons (i.e. car horn, skidding tires) or speech message, that are sent to the driver through a vehicle on board audio system e.g. [9, 10]. The second type consists in a visual warning signal such as a car symbol, flashing orange warning circle, triangular warning that appears on the vehicle dashboard e.g. [5, 6, 11].

Several studies [5, 11, 12] were oriented to the comparison of the effects on driver's behavior at the intersections due to different types of auditory warnings and different types of visual warnings. Such studies highlight the importance of a directional auditory warning information in an intersection collision warning system: a clear warning with directional information about an urgent hazard event allows the drivers to advance the braking maneuver and, thus, avoid the potential collision. However, it is unclear whether is more effective an audio or a visual warning. In addition, no study compared the effects on drivers' behavior induced by an acoustic and a visual directional warning.

The main objective of the present study was to assess, in response to a potential conflict event at the intersections, the effects of directional auditory and visual warnings on driving performance and if the effectiveness of these systems is affected by a perceptual measure.

A driving simulator experiment was carried out to analyze the effect of directional auditory and visual warnings drivers' behavior in response to a potential conflict event, represented by a vehicle that failed to stop at the intersection (violator vehicle).

2 Methodology

2.1 Driving Simulator Experiment

The study was conducted using the advanced driving simulator of the Department of Engineering—Roma Tre University.

A multi-factorial experiment was designed to analyze the effects of the ICWSs (auditory, visual and no warning signal) on drivers' behavior in response to a vehicle that failed to stop at the intersection, both from test vehicle's right and left. The following section describes the road scenarios ICWSs and the traffic calming measure that were implemented in the driving simulator.

2.2 Road Scenarios, ICWSs and Traffic Calming Measure

A two-lane rural road approximately 38 km long was implemented in the driving simulator. According with the Italian road design guidelines [13], the road cross-section was 9.50 m wide formed by two 3.50 m wide lanes and two 1.25 m wide paved shoulders. The design speed ranged from 60 km/h (on curves with a radius equal to 118 m) to 100 km/h (on tangent), and the posted limit was 90 km/h. The radii changed from 118 to 930 m and the lengths of the tangent ranged from 100 to 1650 m. The vertical alignment had null longitudinal grade, to avoid conditionings on the dynamic variables, like speed or acceleration.

Along to the alignment were designed several stop—controlled intersections (four-leg intersections and three-way intersections). In 12 four-leg intersections were simulated the 12 combinations of the factors ICWS (auditory speech message, visual warning and no warning), traffic calming measure (traffic calming measure and no measure) and direction of the violator vehicle (test vehicle's right and left). In all of these 12 intersections a violator vehicle was implemented to fail the stop sign and cross the road (6 from test vehicle's right and 6 from test vehicle's left) at the speed of 70 km/h.

To ensure the same approach conditions, the approach geometry was the same for all the 6 intersections; the driver, after a curve with a radius of 450 m, traveled an approach tangent to the intersection 600 m long. During this approaching phase, the drivers also encountered a vehicle in the opposite direction.

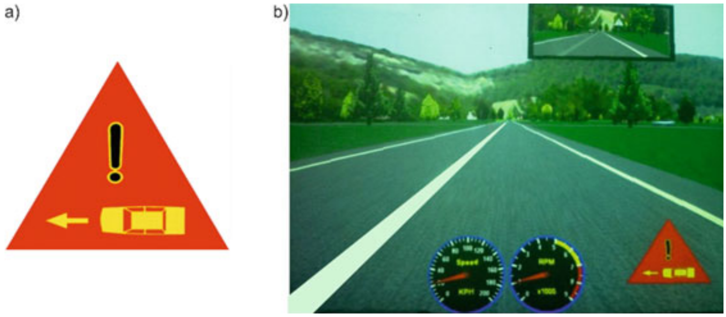


Fig. 1 **a** The *red* car icon of the visual warning and **b** the visualization of the visual warring to the driver during the simulation

To avoid predictability and order effect 6 road scenarios with different sequences of intersections were simulated. Two types of ICWS were implemented in the scenarios.

The first ICWS was the auditory speech message, where the direction of the violator vehicle was specified: "attention, vehicle from right" or "attention, vehicle from left". These speech messages were digitally prerecorded and saved as .wav files. Then they were reproduced into the vehicle through the audio system of the driving simulator at around 70 dB loudness level. They were fully consistent with similar auditory warnings used in literature [14, 15].

The second ICWS was a visual warning, which consisted in a red car icon (an icon of car into a red triangle). It was similar to visual warning used in previous studies in literature and appeared in the right corner of the central display, near the speedometer, to simulate its appearance on a device inside the vehicle. The visual warning provided the direction of the violator vehicle through the icon of car oriented in the direction of arrival of the violator vehicle (Fig. 1). When activated, the visual warning remained in the screen for 7 s.

The triggering point of the ICWS (both auditory and visual) was when the test vehicle reached a point 100 m in advance (i.e. 100 m before) of the intersection. In the same moment the violator vehicle, with the speed equal to 70 km/h, was at 77.7 m from the collision point with the test vehicle. In these conditions and with the hypothesis that the test vehicle is travelling at the posted speed limit (90 km/h), the time to collision (TTC) is equal to 4 s This value, however, is theoretical because it depends on the actual approaching speed of the driver at the intersection during the simulated drive. In other words, if the driver reaches the triggering point at 100 m from the intersection with a higher or a lower speed of 90 km/h, the values of TTC will be lower or higher, respectively, than 4 s.

Concerning the perceptual measure, the dragon teeth were selected. This countermeasure is believed to be particularly effective to give the impression of a narrowing lane and lead the driver to a reduction of the speed. The dragon teeth were placed in the last 50 m before the intersection and were couple of triangles

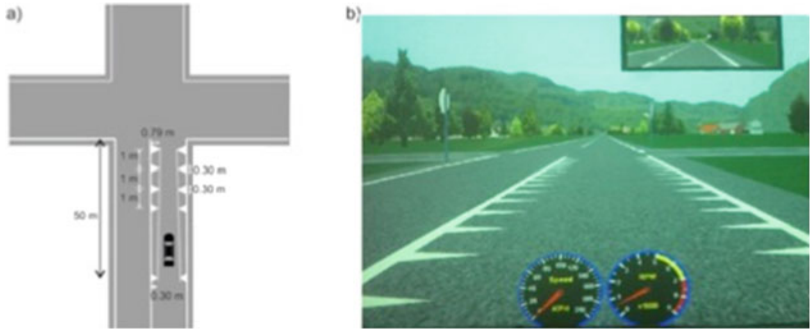


Fig. 2 **a** Dimension and placement of the dragon teeth **b** the visualization of the dragon teeth during the simulation

1 m spaced each other, with a progressively height increasing in order to reduce the lane width from 2.90 to 1.90 m (Fig. 2) [16, 17].

2.3 Apparatus

The driving simulator of the Department of Engineering—Roma Tre University used for this study is an interactive fixed-base driving simulator. It was previously validated [18, 19] and largely used as a reliable tool for the study of the driver's speed behavior e.g., [20–28]. The hardware interfaces (wheel, pedals and gear lever) are installed on a real vehicle. The driving scene is projected onto three screens: one in front of the vehicle and one on either side, which provide a 135° field of view. The resolution of the visual scene is 1024 × 768 pixels with a refresh rate of 30–60 Hz. The system is also equipped with a sound system that reproduces the sounds of the engine and of the auditory warning during the simulation. The simulator provides many parameters for describing the travel conditions (e.g., vehicle barycenter, relative position in relation to the road axis, local speed and acceleration, steering wheel rotation angle, pitching angle, and rolling angle). The data recording system acquired all of the parameters at spatial intervals of 2 m.

2.4 Participants

Forty-two drivers (32 men and 10 women), whose ages ranged from 23 to 70 (average 31) and who had regular European driving licenses for at least three years were selected to perform the driving in the simulator.

The participants were divided into 6 groups; the 6 groups drove the different 6 scenarios, which were each characterized by a specific sequence of intersections where a violator vehicle failed to stop. According to the questionnaire on perceived discomfort (see next Sect. 2.5), 41 of 42 participants experienced null or light levels of discomfort; only one participant was not able to finish the experiment. Thus, the sample used for the analysis consisted of 41 drivers.

2.5 Procedure

The experiment was conducted with the free vehicle in its own driving lane. In the other driving lane, a slight amount of traffic was distributed to induce the driver to avoid driving into that lane. The simulated vehicle was a standard medium-class car with automatic gears. The participants were first briefed about the use of the hardware interface (i.e., wheel and pedals and automatic gear) and then invited to start a training drive at the driving simulator on a specific alignment for approximately 8 min, to become familiar with the driving simulator. After the training, participants came out of the driving simulator for about 5–10 min to restore their initial condition; in this phase, also some information about the experiment were provided. In particular, the drives were informed about the presence of some intersections along the alignment that he had to cross. Drivers were instructed to drive as they normally would in the real world and informed that the vehicle was equipped with an alarm system that advised him of a potential critical situation through an auditory or visual warning. In addition, participants were told that the ICWS system would generate not signal to simulate the condition of driving without an ICWS.

In order to limit the duration of the drive and, thus, reduce the probability of sickness for driver, the experiment was divided in two steps. In the first, the participant drove the first part of one of the six road scenarios and after that, he filled in a questionnaire about his personal data and his driving experience. In the second step, the participant drove the second part of the scenario and then he filled in another questionnaire. This questionnaire consisted in two parts: perceived discomfort, effectiveness of the ICWSs and effectiveness of the perceptual measure. For the first, there were 4 types of discomfort: nausea, giddiness, fatigue and other; each question could be answered by a score of 1–4 in proportion to the level of the discomfort experienced: null, light, medium and high. The null and light level for all 4 types of discomfort is considered to be acceptable for driving. For the effectiveness both of the ICWSs and the perceptual measure, it was asked to the participants if they perceived an effect during the drive. For those who perceived an effect, it was asked to indicate the type of influence (increasing or decreasing the speed, increasing or decreasing the level of attention) and the level of the perceived effect by a score of 1–10.

3 Data Processing

The speed profile of each driver was plotted 150 m in advance of each one of the 12 intersections. As mentioned above, such intersections were 12, equal to the number of the combinations of the following factors: ICWS (3 levels: auditory, visual and no warning); perceptual measure (dragon teeth and no measure), direction of the violator vehicle (2 levels: right and left). Overall, 492 speed profiles (12 intersections × 41 drivers) were analyzed. From each speed profile the following variables of the driver's behavior while approaching the intersection were determined:

- V_i: driver's initial speed value, identified at the moment when the driver starts to decrease his speed, releasing the accelerator pedal or pressing the braking pedal, in response to the violator vehicle;
- V_f: minimum speed value reached by the driver to avoid the collision;
- L_i: beginning distance of the maneuver, the distance from the triggering point and where Vi is located;
- L_f: ending distance of the maneuver, the distance from the triggering point and where Vf is located;
- d_m: the average deceleration rate during the speed reduction phase from V_i to V_f;
- RT: driver's reaction time, which is the elapsed time between the activation of the warning signal (when the test vehicle was at 100 m from the intersection) and the moment in which the driver starts to decrease his speed.

In the intersections where no warning was provided to the driver, the reaction time was assumed equal to the elapsed time between the moment when the test vehicle was at 100 m from the intersection and the moment in which the driver starts to decrease his speed, in response to the violator vehicle.

From the sample were excluded the data of the following cases:

- the driver adopted a too much low (8 data) or to much high speed (13 data) and, thus, the violator vehicle did not affect the driver's behavior (the driver crossed the intersection much late and well in advance compared with the violator vehicle, respectively);
- the driver collided with the violator vehicle 18 (data).

Thus, 453 observations were used for the analysis. It should be noted that the collision events were 12 for the condition of No ICWSs and 6 for the dragon teeth.

4 Data Analysis and Results

The analysis was conducted by means of a multivariate variance analysis (MANOVA) procedure, to investigate all of the interaction and main effects on the dependent variables of the driver's behavior (V_i, V_f, L_i, L_f, d_m, RT) due to the three factors: ICWS (with 3 levels: auditory, video and No ICWS) perceptual measure

Table 1 Descriptive statistics

Variables and factors		Mean Value	SD
Dynamic variable	Vi	83.11 km/h	11.67 km/h
	Vf	33.38 km/h	15.17 km/h
	dm	4.78 m/s^2	1.33 m/s^2
	Li	22.69 m	11.97
	Lf	67.00 m	14.49 m
	RT	0.99 s	0.61 s
ICWS Condition	Auditory speech message	0.34	0.47
	Visual warning	0.35	0.48
	No ICWS	0.31	0.46
Perceptual measure	Dragon teeth	0.49	0.50
	No measure	0.49	0.50
Direction of the violator	Right	0.51	0.50
	Left	0.49	0.50

(with 2 levels: dragon teeth and No measure) and direction of the violator vehicle (with 2 levels: violator from right and left). Table 1 reports the descriptive statistics.

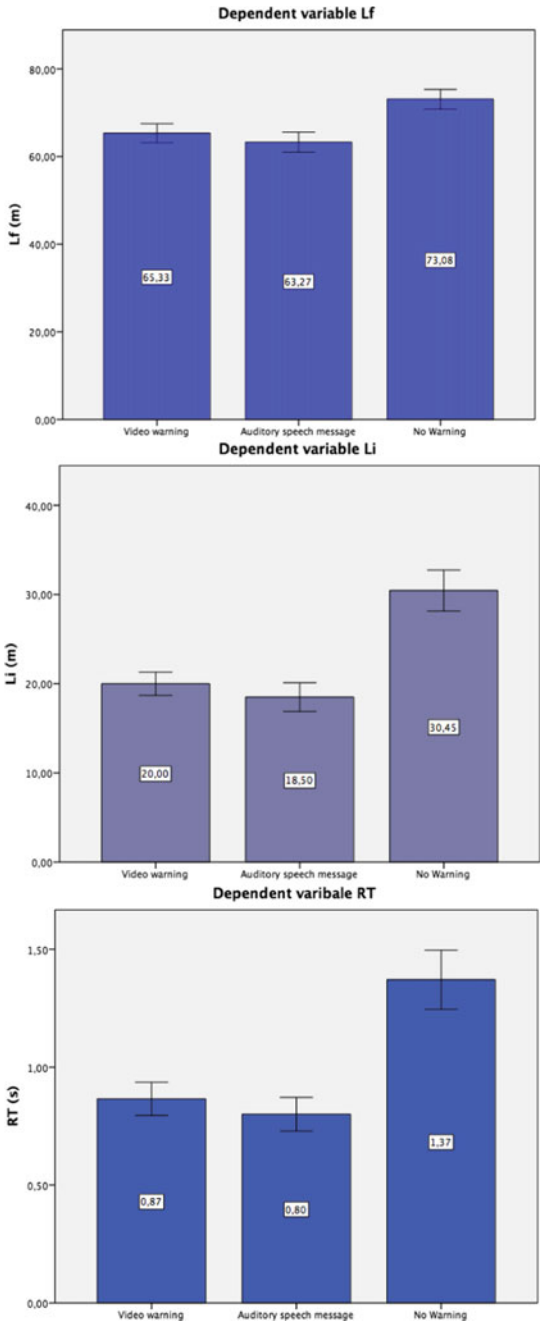
4.1 Drivers' Behavior

The drivers' behavior were compared across the factors ICWS, perceptual measure and direction of the violator vehicle by using the MANOVA test, to asses if these factors affect the driver's behavior in response to a vehicle that failed to stop at the intersections. Bonferroni correction was used for multiple comparisons.

MANOVA revealed that there was a significant main effect for the warning conditions ($F_{(12,872)} = 9.07$, $P = 0.000$ Wilk's $\Lambda = 0.790$, partial Eta squared = 0.111, observed power = 1), while for perceptual measure and the direction of the violator no significant effects were found ($F_{(6436)} = 1.98$, $P = 0.066$ Wilk's $\Lambda = 0.973$, partial Eta squared = 0.027, observed power = 0.728; $F_{(6436)} = 1.29$, $P = 0.301$ Wilk's $\Lambda = 0.962$, partial Eta squared = 0.038, observed power = 0.886, respectively). No interaction effects were found. The test between subjects revealed that the beginning distance of the maneuver L_i, the ending distance of the maneuver L_f and the reaction time RT were affected by the ICWS condition in a significant way ($F_{(2441)} = 54.24$, $P = 0.000$; $F_{(2441)} = 20.69$, $P = 0.000$, $F_{(2441)} = 46.13$, $P = 0.000$, respectively).

The lower L_i was reached for the auditory speech message warning (18.44 m), which was statistically significantly lower than that for the no warning condition (mean difference = 10.45 m, P = 0.000). Also the L_i for the visual warning system (19.99 m) was statistically significantly lower than that for the no warning condition (mean difference = 11.95 m, P = 0.000). The L_i for speech message was not

Fig. 3 Effect of ICWS on drivers' behavior



significantly different from that for the visual warning system (mean difference = 1.49 m, P = 0.658).

The lower L_f was reached for the auditory speech message warning (63.28 m), which was statistically significantly lower than that for the no warning condition (mean difference = 9.80 m, P = 0.000); also the L_f for the visual warning was statistically significantly lower than that for the no warning condition (mean difference = 7.75 m, P = 0.000, respectively). No other statistically differences were found.

For the drivers' reaction time, the lowest value was recorded for the auditory speech message (0.79 s), which was statistically significantly lower than that for the no warning condition (mean difference = 0.57 s, P = 0.000). Also the RT for the video warning was statistically significantly lower than that for the no warning condition (mean difference = 0.50 s, P = 0.000). No other difference was statistically significant. The effect of the ICWSs on drivers' behavior are reported in Fig. 3.

Even if not statistically significant, an interesting effect of the perceptual measure was recorded; compared with the condition of No measure, for the dragon teeth a lower value of the speed V_f was recorded (mean difference = 1.33 km/h; P = 0.337).

4.2 Outcome of the Questionnaire

The results of the questionnaire about the perceived effectiveness of the warning signals and traffic calming measure showed that the entire sample indicated that both the visual warning and the auditory speech message were effective, while no effect of the perceptual measure was perceived. Auditory speech message obtained the highest score for both the speed reduction effect (mean = 8.3, SD = 1.3) and the increase of the level of attention during the drive (mean = 7.3, SD = 2.3).

These results indicate that the participants believed to have been more influenced when the warning signals were present; moreover, the auditory speech message was believed more effective than the visual warning.

5 Discussion

As expected, the drivers' behavior were affected by the ICWS, while no significant effect was recorded for the perceptual measure and the direction of the violator. When the ICWS was present the driver started earlier the maneuver (L_i) to avoid the conflict; moreover, for the auditory speech message the driver started from the triggering point 1.50 in advance (not statistically significant) with respect of the video warning. This result is connected to the outcome of the variable L_f; starting earlier the maneuver, the driver is able to complete the maneuver further from the collision point at the intersection. Also in this case, the lowest value was recorded for the auditory speech message (63.27 m).

Consistently with the results of L_i and L_f, when the ICWS was present also the RT was affected in a significant way. The lowest values (statistically significant) were recorded for the auditory speech message (0.80 s) and the audio warning (0.87 s). However this difference was not statistically significant.

These findings suggest that with the improvement of the ability to react early in response to a critical situation, the driver advances the braking maneuver. This means that he has more time to reduce the speed in order to avoid the collision, adopting a safer braking maneuver. The improvement of the drivers' reaction time due to the presence the ICWSs (both auditory speech message and visual warning) was consistent with the previous studies e.g. [11].

Concerning the effects of the two different types of directional warnings on driving behavior, the results on RT showed that, despite the two warning signals provided the same information to the driver (i.e. the direction of the vehicle), the auditory speech message was better than the visual warning. The higher reaction time (not statistically significant) for the visual warning (0.87 s) compared with that for the auditory speech message (0.80 s) can be due to the fact that for the first, the driver had to focus his attention, and thus his glance, before on the visual signal to identify the direction by the red car icon, and then on the intersection to detect the violator vehicle. For the auditory speech message the driver, instead, could directly detect the position of the violator after he heard the audio signals with the directional information and, thus, advance the beginning of the braking. This result is consistent with previously studies e.g. [11, 29] where the fastest reaction times were found for the auditory signals and remarks the nature of the physical stimuli that are solicited by the warning signals.

In addition, most of the driving activity requires the visual task; this implies that the comprehension of the video signal (i.e. an additional visual task) could disturb the driving activity and, thus, delay the reaction of the driver.

For the perceptual measure, no statistically significant effects were found. However, for this condition, a decrease of about 1.3 km/h of the V_f was recorded. This result is consistent with the aim of the perceptual measure.

For the direction of the violator, no statistical difference was recorded and this result highlight that the effectiveness of the different ICWSs was the same with respect of the direction of arrival of the violator vehicle.

6 Conclusion

The statistical analysis showed that the variables L_i, L_f and RT were statistically affected by the ICWSs, while the perceptual measure and the direction of the violator vehicle did not affected the drivers' behavior.

The initial distance of the maneuver (L_i), which provide information about the capability of the driver to advance the braking maneuver, was lower for the ICWSs with respect of the condition of No ICWS. This means that in presence of ICWS the driver was more able to start earlier the braking maneuver to avoid the conflict at the intersections. In addition, the results showed that for the auditory warning the driver were able to advance the maneuver more than that for the video warning (the difference was not statistically significant). The final distance of the maneuver L_f was also statistically affected by the ICWSs. As consequence of advancing the braking maneuver, the driver reached the minimum speed further from the collision point at the intersection with respect of the condition of No ICWS. This means that the driver execute a safer maneuver to avoid the conflict with the violator vehicle. Also in this case, the final distance was lower (not statistically significant) for the auditory speech message with respect of the video warning.

The driver's reaction time (RT) (the time needed for the driver to react in response to a warning signal) was statistically significantly lower for the condition of presence of ICWS.

These results highlight that the driver, when the vehicle was equipped with one of the warning signals, were able to advance the braking maneuver due to the earlier reaction in response to the violator vehicle; moreover, this effectiveness was slightly more highlighted for the auditory speech message.

The benefits of advance the braking maneuver resulted in no collision event for the warning signals, while for the No ICWS condition, on average, 14.3 % of the drivers collided with the violator vehicle. These findings were also confirmed by the outcomes of the questionnaire on the perceived effectiveness of the warning signals; the entire sample reported that the warning signals were effective. Moreover, the drivers reported that the auditory speech message was more effective of the visual warning in terms of speed reduction and improvement in the level of attention.

The present study was conducted using the advanced driving simulator of the Department of Engineering—Roma Tre, which ensured a full control of the experimental condition and no risk to the participants. In addition, it was previously validated for the analysis of drivers' behavior on rural roads [18]. It should be recognized that the validation of the driver simulator equipped with the ICWSs to has not yet been developed. However, considering the reliability and the method used in previous studies e.g. [5, 11], in which the driving simulators had the same characteristics as the driving simulator used in this study, there are sufficient guarantees of the goodness of the obtained results.

References

1. DaCoTa: Road safety project (2012). http://www.dacota-project.eu/
2. ACI-ISTAT: Rapporto ACI—ISTAT sugli incidenti stradali, anno 2013 (2014) (in Italian)
3. CARE: Traffic safety basic facts 2015. Junctions (2015). http://ec.europa.eu/transport/road_safety/pdf/statistics/dacota/bfs2015_junctions.pdf

4. Tay, R.: A random parameters probit model of urban and rural intersection crashes. Accid. Anal. Prev. **84**, 38–40 (2015). doi:10.1016/j.aap.2015.07.013
5. Werneke, J., Vollrath, M.: How to present collision warnings at intersections?—A comparison of different approaches. Accid. Anal. Prev. **52**, 91–99 (2013). doi:10.1016/j.aap.2012.12.001
6. Penney, T.: Intersection collision warning system. Pub. No. FHWA-RD-99-103 (1999)
7. Lee, J.D., McGehee, D.V., Brown, T.L., Reyes, M.L.: Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. Hum. Factors **44**, 314–334 (2002). doi:10.1518/0018720024497844
8. Atev, S., Masoud, O., Janardan, R., Papanikolopoulos, N.: Real-time collision warning and avoidance at intersections. Report No. Mn/DOT 2004-45 (2004)
9. Gray, R.: Looming auditory collision warnings for driving. Hum. Factors **53**, 63–74 (2011). doi:10.1177/0018720810397833
10. Haas, E.C., Van Erp, J.B.F.: Multimodal warnings to enhance risk communication and safety. Saf. Sci. **61**, 29–35 (2014). doi:10.1016/j.ssci.2013.07.011
11. Chen, H., Cao, L., Logan, D.B.: Investigation into the effect of an intersection crash warning system on driving performance in a simulator. Traffic Inj. Prev. **12**, 529–537 (2011). doi:10.1080/15389588.2011.603169
12. Zhang, Y., Yan, X., Yang, Z.: Discrimination of effects between directional and nondirectional information of auditory warning on driving behavior. Discret. Dyn. Nat. Soc. (2015). doi:10.1155/2015/980281
13. Ministry of Infrastructures and Transports: Decreto Ministeriale del 5/11/2001 Norme funzionali e geometriche per la costruzione delle strade. Istituto Poligrafico dello Stato, Roma (2001). (in Italian)
14. Baldwin, C.L., May, J.F.: Loudness interacts with semantics in auditory warnings to impact rear-end collisions. Transp. Res. Part F Traffic Psychol. Behav. **14**, 36–42 (2011). doi:10.1016/j.trf.2010.09.004
15. Yan, X., Zhang, Y., Ma, L.: The influence of in-vehicle speech warning timing on drivers' collision avoidance performance at signalized intersections. Transp. Res. Part C Emerg. Technol. **51**, 231–242 (2015). doi:10.1016/j.trc.2014.12.003
16. Department for Transport: Traffic calming in villages on major roads. TA Leaflet 1/00 (2000)
17. Highway Agency: design manual or roads and bridges (2004). Traffic Calming on Trunk Roads a Practical Guide, vol. 6 Sect 3 Part 5 TA87/04 (2004)
18. 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
19. 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
20. Bella, F.: New model to estimate speed differential in tangent–curve transition. Adv. Transp. Stud. Int. J. **15**(2008), 27–36 (2008). doi:10.4399/97888548186823
21. Bella, F., Calvi, A.: Effects of simulated day and night driving on the speed differential in tangent-curve transition: a pilot study using driving simulator. Traffic Inj. Prev. **14**, 413–423 (2013). doi:10.1080/15389588.2012.716880
22. 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
23. Bella, F., Silvestri, M.: Driver's braking behavior approaching pedestrian crossings: a parametric duration model of the speed reduction times. J. Adv. Transp. (2016). doi:10.1002/atr.1366
24. Bella, F.: Driver performance approaching and departing curves: driving simulator study. Traffic Inj. Prev. **15**, 310–318 (2014). doi:10.1080/15389588.2013.813022
25. 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
26. Bella, F.: Driver perception hypothesis: driving simulator study. Transp. Res. Part F Traffic Psychol. Behav. **24**, 183–196 (2014). doi:10.1016/j.trf.2014.04.007

27. 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
28. Bella, F., Calvi, A., D'Amico, F.: Analysis of driver speeds under night driving conditions using a driving simulator. J. Saf. Res. **49**, 45–52 (2014). doi:10.1016/j.jsr.2014.02.007
29. Green, M.: How long does it take to stop? Methodological analysis of driver perception-brake times. Transp. Hum. Factors **2**, 195–216 (2000). doi:10.1207/STHF0203_1

Directional Identification of Sirens and Warnings in a Simulated Driving Task: Comparison of Two Loudspeaker Technologies

Christine Mégard, Valentin Le Guelvouit, Steven Strachan, Cendrine Mercier and Christian Bolzmacher

Abstract The hearing perturbation affecting the elderly called presbycusis is characterized by the degradation of the perception of high tones and impacts the capacity of drivers in localizing warnings and sirens while driving. We present the results of a user study performed to compare the capacity of traditional and car glass loudspeakers in providing directional alarms inside the car. The experiment compares the number of directional errors and subjective perception of alarms. It was performed in lab, in a real stationary car while the participant was engaged in a driving task on a driving simulator inside the car. Both technologies provide directional sounds with difficulties in perceiving sounds from the back mainly due to the anatomy of the ear as well as the wrap-around shape of the seats. Glass panes performed slightly better than traditional loudspeakers with regard to left/right confusion due to the position next to the ear.

C. Mégard · V. Le Guelvouit · S. Strachan · C. Bolzmacher (✉)
Sensorial and Ambient Interfaces Laboratory, CEA, LIST,
91191 Gif-sur-Yvette Cedex, France
e-mail: christian.bolzmacher@cea.fr

C. Mégard
e-mail: christine.megard@cea.fr

V. Le Guelvouit
e-mail: valentin.leguelvouit@cea.fr

S. Strachan
e-mail: steven.strachan@cea.fr

C. Mercier
Centre d'Expertise National des Technologies de l'Information et de la Communication pour l'autonomie (CENTICH), 49000 Angers, France
e-mail: cendrine.mercier@hotmail.fr

C. Mercier
Centre Recherche en Education de Nantes (CREN), Université de Nantes, EA 2661, 44312 Nantes Cedex 3, France

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_65

Keywords User-centered study · Loudspeaker comparison · Acoustic glass · Sound localization · Car simulator

1 Introduction

Security in transportation is a worldwide crucial societal problem. Security of the driving elderly is even more worrisome. Statistics in France indicate that in 2014 elderly drivers aged 75 and over have 1.6 times more chances to be killed on the road than any other car driver. Many factors may lead to these statistics. Reflexes deteriorate with age, medicines may impact vigilance, psychological, sensorial and motor performances deteriorate with age. Visual factors may also contribute to these bad statistics but it is now recognized that a bad hearing can be the cause of accidents by decreasing the capacity of the driver to react to outer events (horns, arrival of firetrucks or Police) in critical situations [1].

The hearing perturbation that affects the elderly is characterized by the progressive bilateral degradation of the audition in the perception and in the processing of high tones [2]. This phenomenon is called presbycusis [3].

Preliminary user studies [4] indicated that presbycusic drivers declared to have difficulties in the identification of the direction of warnings and sirens while driving. This degradation may impact severely the sense of safety and self-confidence and of course the safety of elderly drivers. The majority of users with or without hearing aids often meet difficulties, in perceiving the different types of alarms coming from outside the car and in recognizing if an alarm is coming from the front or the back particularly at intersections and during lane change. The front-back localization is harder than left-right and particularly the back localization is difficult even for normal hearing users.

Our study investigates the most appropriate sound system that could be used for alerting the driver in personal cars. Additional traditional loudspeaker can be installed in the car to provide directional auditory information. However innovative technologies developed in our laboratory allow instrumenting the glass of the car to be used as loudspeaker. As a car is equipped with at least 4 surrounding glasses they could be used as directional loudspeakers. The advantage of this so-called "car glass" technology is that it uses the existing infrastructure of the car and requires no visible additional set-up that may impair the visual field of view of the driver. In this paper we present the protocol and the results of the study performed to answer the question about the most appropriate additional sound system that could be used for alerting the driver in personal cars. The protocol is based on the comparison of the capacity of the two technologies (traditional loudspeakers and car-glass technology) in providing directional sounds inside the car. The experiment investigates and compares the number of directional errors and subjective perception of alarms by the participants between the two systems.

2 Description of the Experimental Setup

2.1 Description of the Technologies

In general, cars are not using the glass panes as loudspeakers to spatialize sound. The first goal is therefore to install a loudspeaker system based on acoustic glass. Four piezoelectric transducers have been integrated in a Renault Zoé test car. The glass panes have been equipped with large 50 mm outer diameter piezoelectric Pz26 rings from Ferroperm, hidden in the lower part of the door structure and the lower part of the windshield and the rear window in order to be invisible to the user and protected from environmental influences. Figure 1 shows the exact position of the transducers. It has to be noted that the transducer position on the side window can change when opening the windows. Blinded wires have been thoroughly hidden behind the car's plastic covers and guided towards the center console. Additionally, four traditional loudspeakers have been placed at the center of the dashboard and the back shelf as well as at the lower parts of the front side doors. Both loudspeaker systems (traditional and glass pane loudspeakers) are coupled to a digital signal processor (DSP), which allows for individual addressing of each.

2.2 Description of the Driving Simulator and Primary and Secondary Tasks

In a real setting the occurrence and the direction of alarms cannot be controlled. Laboratory testing provides therefore controlled conditions that are mandatory to compare the efficiency of different technologies in providing directional sounds.

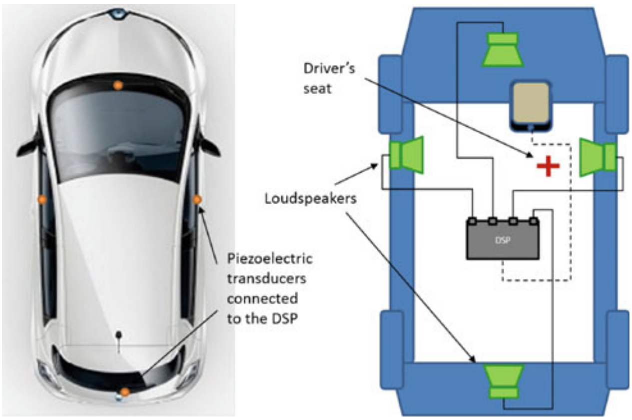


Fig. 1 Transducer and loudspeaker placement in the test car

Fig. 2 Experimental set-up used for the simulated driving task (*left*) and the non-simulated driving task (*right*)

However, the evaluation of the technologies to be used in a car must take into account the attentional processes which are particularly important during the primary task of driving. A primary task is defined as the priority task over all other tasks or any other events.

One solution is to use a simulated driving task as primary task. Realistic driving simulators are proved efficient to be used in in-lab experimental set-up. However their cost in hardware, software and personnel costs for deployment were out of the budget and the timing of the study. The evaluation of the direction of sounds is performed in a real stationary car on the driving simulator "Safety Driving Simulator Auto" from Anuman Interactive [5]. The participant is asked to drive a simulated car in an extra-urban environment. The primary task of the participant was to drive safely the car along the track and indicate as soon as possible the direction of the alarm as soon as they occur during the ride. The participants indicated the direction of the alarm on a Tablet PC located on their left. The answers of the participants were used to collect the number of errors.

The participant was seated on the "passenger" seat, i.e. the right front seat (in the case of French cars) to position a specific steering wheel to control the simulated car (Fig. 2)

3 Description of Protocol

The protocol is composed of three steps. At first, the participant is introduced with the aim of the study and is asked to fill up the form consent. The first step of the experimental protocol is performed in a real stationary car on a driving simulator.

The participant is asked to drive a simulated car in an extra-urban environment inside the physical car. Some alarm sounds occur during the ride. The task of the participant is to drive safely the car along the track and indicate as soon as possible the direction of the alarms on a Tablet PC.

Once the driving step is finished, each participant is asked to fill up a subjective questionnaire concerning the sound performed either by the car glass or by traditional loudspeakers [6, 7] and the Osgood semantic scale [8].

Participants. 23 Participants took part in the study (8 females, 15 males), 9 were over 50 years old, and 14 were young participants (less than 35 years old). All of them were regular drivers.

Within-Subject Protocol. As individual differences will necessarily occur, a within-subject protocol is used. Each participant tested the two conditions, car glass and the standard loudspeakers, randomly.

Test Signals. Three alarms are tested to avoid as much as possible any response strategy applied by the participant. The alarms are played in a random manner either by the traditional loudspeakers or by the acoustic glass on all possible directions. This approach makes it almost impossible to predict where the next alarm signal might come from. It is better to use regular alarms stored in long term memory not to disrupt the participants with new signals [9]. For this reason we used the French Fireman and the French Police sirens as the experiment was performed in France.

We also tested an indoor alarm, the Lane Departure Warning (LDW) because it was the only directional warning that was not triggered by a deliberate action of the driver. This warning was only used in two possible directions (left or right).

Bliss and al. [10] have shown the importance of the duration of a signal for the identification of an alarm signal and as an indicator of the validity of the alarms. Short alarms are often considered as fake alarms. In their experiment, Kuwano et al. [8] used a duration of 500 ms. This duration was used for the Lane Departure Warning. However the two other alarms, the French fireman and the French police have a longer sound pattern. Therefore the duration was 2 s for these two alarms. The amplitude of the warnings with the two technologies was equalized using a soundmeter to provide a perceptively equivalent sound level for the 3 sirens of 65 dB. The motor sound from the simulation game could unfortunately not be provided by our experimental set-up. Each siren or warning was presented 5 times in each possible direction.

4 User Experience Assessment

The efficiency of the car-glass loudspeaker and traditional additional loudspeakers was assessed by objective measures. However the assessment of the user experience was as much important as the objective measures.

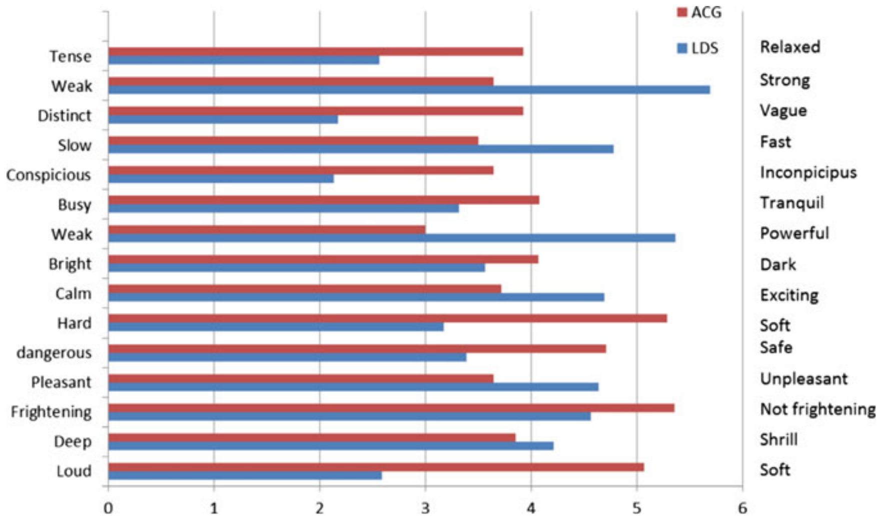


Fig. 3 Osgood semantic differential for the acoustic glass (ACG) and for the traditional loudspeaker (LDS)

Subjective Assessment of the Alarms. Efficient alarms need also to be acceptable for the driver and other possible passengers. The assessment of alarms has been studied by Fastl et al. [6] on the following items: Perceived ease of use, Intention of use, Perceived usefulness and Perceived usability. The following criteria from Hellier and Edworthy [7] were added to the evaluation like: Perceived quality, Perception of annoyance (pleasant/unpleasant; undesirable/desirable; irritating/likable; annoying/nice), Usefulness/worthless/assisting; useless/useful; superfluous/effective; sleep inducing/raising alertness; Bad/good. The subjective answers are gathered on a seven levels Lickert Scale, as used by Fastl et al. [6].

Semantic Differential. The subjective scale provided by Fastl et al. [6] is particularly accurate to gather large subjective judgments relative to user experience, but this scale is not adequate to gather personal descriptions of sounds. We chose to enrich the subjective assessment of the alarms with the semantic scale developed by Kuwano [8]. The semantic scale is used by each participant to characterize their perception of an alarm provided by traditional loudspeaker and car glass loudspeakers. This scale is presented in Fig. 3. Each participant described the perception for one very well-known alarm (fireman siren) delivered by the two technologies tested: traditional loudspeakers and car glass loudspeaker using sixteen pairs of adjectives: Loud/Soft, Deep/Shrill, Frightening/Not frightening, Dangerous/Safe, Hard/Soft, Calm/Exciting, Bright/Dark, Weak/Powerful, Busy/Tranquil, Conspicuous/Inconspicuous, Slow/Fast, Distinct/Vague, Weak/Strong. A French translation of the semantic label has been provided by the authors.

5 Results and Analysis

5.1 Comparison Between the Acoustic Glass and the Traditional Loudspeaker Concerning the Directional Errors

Considering all alarms, the total number of errors in the perception of the direction of alarms is not much different between the acoustic glass and the traditional loudspeaker technology. The mean number of the sum of errors with the acoustic glass is 12.59 (SD = 7.22) when the participant is driving, and 10.17 (SD = 8.50) when is not driving. The mean number of the sum of errors with the traditional loudspeaker is 12.93 (SD = 6.94) when the participant is driving, and 10.52 (SD = 7.37) when is not driving. The mean difference is neither significant between the two technologies when the participants are implied in a simulated driving task ($t(22) = 0.38$; $p > 0.05$), nor when the participants are only dedicated to the direction task ($t(22) = 0.45$; $p > 0.05$).

Directional Errors for Alarms Coming from the Back of the Car. The mean number of errors is quite low for the two technologies when delivering an alarm from the back of the car. The mean number of Back/Left confusions is 0.95 (SD = 1.13) for the acoustic glass and 1.09 (SD = 1.37) for the traditional loudspeaker $t(22) = 0.51$; $p > 0.05$). The mean Back/Front confusion is 1.18 (SD = 1.62) for the acoustic glass and 1.36 (SD = 1.86) for the traditional loudspeaker ($t(22) = 0.23$; $t > 0.05$).

Directional Errors for Alarms Coming from the Front of the Car When the Participants are Driving. The mean number of confusions Front/Right is 0.36 (SD = 0.56) for the acoustic glass and 0.36 (SD = 0.49) for the traditional loudspeaker ($t(22) = 0.00$; $p > 0.05$). The mean number of confusions Front/Left is 0.54 (SD = 0.74) for the acoustic glass and 0.45 (SD = 0.66) for the traditional loudspeaker ($t(22) = 0.438$; $p > 0.05$).

The mean Front/Back confusion is 3.72 (SD = 2.47) for the acoustic glass and 0.72 (SD = 0.85)) for the traditional loudspeaker ($t(22) = 5.04$; $p < 0.05$)*.

5.2 External Alarms (French Firemen and French Police)

Directional Errors for Alarms Coming from the Left of the Car When the Participants Are Driving. The mean number of Left/Front confusions is 0.95 (SD = 1.32) for the acoustic glass and 1.41 (SD = 1.68) for the traditional loudspeaker ($t(22) = 2.88$; $p < 0.05$)*. The mean number of confusions Left/Right is 0.45 (SD = 0.67) for the acoustic glass and 0.36 (SD = 0.58) for the traditional loudspeaker ($t(22) = 1.00$; $p > 0.05$). The mean Left/Back confusion is 0.54 (SD = 0.91) for the acoustic glass and 0.63 (SD = 1.00) for the traditional loudspeaker ($t(22) = 1.00$; $p > 0.05$).

Directional Errors for Alarms Coming from the Right of the Car When the Participants Are Driving. The mean number of Right/Back confusions is 2.13 (SD = 1.58) for the acoustic glass and 1.68 (SD = 1.42) for the traditional loudspeaker ($t(22) = 2.88$; $p < 0.05$)*. The mean number of confusions Right/Left is 0.54 (SD = 0.91) for the acoustic glass and 0.59 (SD = 0.91) for the traditional loudspeaker ($t(22) = 0.37$; $p > 0.05$). The mean Right/Front confusion is 0.22 (SD = 0.53) for the acoustic glass and 0.22 (SD = 0.52) for the traditional loudspeaker ($t(22) = 0.00$; $p > 0.05$).

Number of no Response (Timeout). The total number of timeouts for the external alarms is quite the same for the Front (0 timeout for the acoustic glass and a total of 2 for the traditional loudspeaker). Back alarms generate much more "no responses", but without much difference between the two technologies (24 timeouts for the acoustic glass and 28 for the traditional loudspeaker). For the Left alarm, the number of timeouts is much more important for the traditional loudspeaker (15 timeouts) than for the acoustic glass. The same effect is observed with the Right alarm (35 timeouts for the traditional loudspeaker and 8 for the acoustic glass).

5.3 Internal Alarm (Lane Departure Warning)

Directional Errors for Alarms Coming from the Right of the Car When the Participants Are Driving. The mean number of Right/Back confusions is 0.09 (SD = 0.29) for the acoustic glass and 0.27 (SD = 0.55) for the traditional loudspeaker ($t(22) = 2.16$; $p < 0.05$)*. The mean number of confusions Right/Left is 0.36 (SD = 0.72) for the acoustic glass and 0.45 (SD = 0.85) for the traditional loudspeaker ($t(22) = 0.698$; ($p > 0.05$). The mean Right/Front confusion is 0.04 (SD = 0.21) for the acoustic glass and 0.0 (SD = 0) for the traditional loudspeaker ($t(22) = 1.00$; $p > 0.05$).

Directional Errors for Alarms Coming from the Left of the Car When the Participants Are Driving. The mean number of Left/Front confusions is 0.27 (SD = 0.45) for the acoustic glass and 0.22 (SD = 0.52) for the traditional loudspeaker. The mean number of confusions Left/Right is 0.45 (SD = 1.10) for the acoustic glass and 0.36 (SD = 0.78) for the traditional loudspeaker. The mean Left/Back confusion is 0.04 (SD = 0.23) for the acoustic glass and 0.04 (SD = 0.21) for the traditional loudspeaker.

Number of no Response (Timeout). The timeouts (no response from the participant) are more numerous when the alarm is emitted by the acoustic glass (5 when driving, and 5 without driving) than with the traditional loudspeaker (3 when driving and 1 without driving) for left alarms.

For right alarms, the number of timeout (no response) is higher for the traditional loudspeaker (5 when driving and 4 without driving) than for the acoustic glass (3 when driving and only 1 without driving).

5.4 Semantic Differential

After the whole experimental session, each participant was asked to describe the sound produced with the two technologies (traditional loudspeaker and car glass loudspeaker) using the Osgood semantic differential [8]. Figure 3 shows the mean answer to the Osgood semantic differential for the acoustic glass and for the traditional loudspeaker. The semantic differential shows a clear tendency of the acoustic glass to provide a softer, calmer and weaker sound than the sound provided by the traditional loudspeaker. This difference may be explained by the difference in the quality and the characteristics of the sound emitted by the two technologies.

6 Analysis and Discussion

Internal alarms lead to few errors in general. This result can be explained by the fact that these alarms were provided only on the left and on the right side of the driver, thus reducing the possible ambiguities between the two directions. However, more errors occur on the left side for the acoustic glass than for the traditional loudspeaker. This may be due to the fact that the acoustic glass seems to produce a more diffused sound that is more critical for the sound coming from the left side of the participant as it is farther away (participant was sitting on the passenger seat during tests) than the sound coming from right side.

More errors occur on the right side for traditional loudspeaker. This result is probably due by the fact that the sound coming from the traditional loudspeaker was emitted below the ear level, thus impairing directional perception for right side sounds, compared to the sound clearly emitted on the right side with the acoustic glass. We can hypothesize that in situations in which the driver is seated on the left side of the car we would find opposite results, with a better directional perception with the acoustic glass on the left side and a better directional perception on the right side with the traditional loudspeaker.

Back external alarms are difficult to perceive with both technologies. Few Back/Front errors are detected but numerous timeout responses occur indicating that the participants could not identify the direction of these alarms. This result was found for both technologies. Beside the fact that the anatomy of the ear does not favor the perception of sounds coming from the back, we can hypothesize that the wrap-around shape of the seats of the car hinders furthermore the perception of back sounds.

Subjective scales and the Osgood semantic differential indicate an overall better judged sound quality between the sound emitted with the car glass than with the traditional loudspeaker. This may come from a softer, less sharp sound emitted by the acoustic glass, and thus, perceived less aggressive than the sound from a traditional loudspeaker.

In view with the results of the study, we can suggest that the acoustic glass can be considered as a possible interface to provide directional lateral sounds in cars. The acoustic glass provides good results for those lateral directions due to its position. Delivering a sound is technically more feasible than using loudspeaker at-ear level.

7 Conclusion

This paper presents the protocol and the results of the study performed to evaluate two loudspeaker technologies for alerting a driver of incoming alarms. The protocol is based on the comparison of the capacity of traditional and car-glass loudspeakers in providing directional sounds inside the car. The experiment investigates and compares the user experience and subjective perception of the sounds by the participants. During the experiment, the participant was sitting at the front right seat for experimental purposes.

Internal alarms lead to few errors in general. This result can be explained by the fact that this alarm was provided only on the left and on the right side of the driver, thus reducing the possible ambiguities between the two directions. However, more errors occurred on the left side for the acoustic glass than for the traditional loudspeaker. It is due to the fact that the acoustic glass seems to produce a more diffused sound that is more critical for the sound coming from the left side of the participant as it is farther away than the sound coming from right side. More errors occur on the right side for traditional loudspeakers. This result is probably due to the fact that the sound coming from the traditional loudspeaker was emitted below the ear level, thus impairing directional perception for right side sounds, compared to the sound clearly emitted on the right side with the acoustic glass. We can hypothesize that in situations in which the driver is seated on the left side of the car we would find opposite results, with a better directional perception with the acoustic glass on the left side and a better directional perception on the right side with the traditional loudspeaker technology.

For external alarms, both technologies provide directional sounds with difficulties in perceiving sounds from the back. Few Back/Front errors are detected but participants rather chose the strategy not to answer (numerous timeout responses) indicating that they could not identify the direction of the alarms. This result was found for both technologies. Beside the fact that the anatomy of the ear does not favor the perception of sounds coming from the back, we can hypothesize that the wrap-around shape of the seats of the car hinders furthermore the perception of back sounds.

The car-glass technology is a promising candidate to deliver alarms inside a vehicle. However further technical research should be engaged to provide more directional sounds with the “car-glass” technology in the main directions.

Acknowledgments The work reported here was performed as part of the ICityForAll project (Contract No. AAL 2011-4-056) and is funded by the Active and Assisted Living Programme.

References

1. Accidents seniors, données 2014 pour la France métropolitaine, DSCR/fichiers BAAC (2014). http://www.preventionroutiere.asso.fr/Nos-publications/Statistiques-d-accidents/Accidents-seniors
2. Abel, S., Paik, M., Shelly, J.E.: The benefit of practice for sound localization without sight. Appl. Acoust. **65**, 229–241 (2004)
3. Miller-Keane, B.F., O'Toole, M.T.: Encyclopedia and dictionary of medicine, nursing, and allied health. Saunders, Elsevier, Inc. (2005)
4. Brescia, A., Longo, M., Le Coz, K., Mercier, C., Belouard, Y., Hajjam, J.: Users' requirements —evaluation Report (2013). http://www.icityforall.eu/publications/deliverable_1.1_escoop-centich.pdf
5. Safety driving simulator auto, anuman interactive (2013). http://jeuxvideo.jeu.orange.fr/acheter-telecharger-jeux-pc/safety-driving-simulator-auto-768776.html
6. Fastl, H., Menzel, D., Müllner, M., Wenzelewski, K., Wigro, H.: Perceptive aspects of emergency signals. In: 39th International Congress and Exposition on Noise Control Engineering, Lisbon, Portugal (2010)
7. Hellier, E., Edworthy, J.: On using psychophysical techniques to achieve urgency mapping in auditory warnings. Appl. Ergon. **30**(2), 167–171 (1999)
8. Kuwamo, S., Namba, S., Schick, A., Holger, H., Fastl, H., Filippou, T., Florentine, M.: Subjective impression of auditory danger signals in different countries. Acous. Sci. Technol. **28**(5), 360–362 (2007)
9. Guillaume, A., Pellieux, L., Chastres, V., Drake, C.: Judging the urgency of nonvocal auditory warning signals: perceptual and cognitive processes. J. Exp. Psychol. **9**(3), 196–212 (2003)
10. Bliss, J.P., Fallon, C.K., Nica, N.: The role of alarm signal duration as a cue for alarm validity. Appl. Ergon. **38**(2), 191–199 (2007)

Bus Rapid Transit (BRTS): A Case Study to Investigate Its Use by the Persons with Disability

Devarshi Chaurasia, Sandeep Sankat and Sushil Kumar Solanki

Abstract In India, increasing urbanisation, modernization, commercialisation, changing demographics, extensively increasing number of personal vehicles in the urban areas of India resulted in congestion on the roads. Cities around the world, at some point of time have faced problems associated with passenger mobility and connecting the city periphery with central part, in urban areas and found few innovative solutions to overcome the problems. Urban Planners, Engineers and Urban Administrator around the world have found Bus Rapid Transit (BRT) System as efficient, cost effective and simple as compare to other Light Rail Transit (LRT) and Metro Rail solution to provide 'life line' to city. Cities around the world operate BRTS and got positive results including so many Indian cities. The ultimate aim of this study is to develop an understanding regarding urban public transport issues for persons with disability. To review usability of Bus rapid transport system for persons with disability as commuters. To understand the impact of 'Mass transit System' on Sustainable social change.

Keywords BRT system · Urban transport · City · Passenger mobility · Disability

1 Indian Scenario

The scenario of Indian cities is particularly different from that in the western countries because of their dense and mixed land use. These dense Indian cities do not follow a simplistic structure of CBD and suburban development but are unique multi-nuclei structures with organically evolved road network patterns. In most

D. Chaurasia (✉) · S. Sankat · S.K. Solanki
School of Planning and Architecture (SPA), Bhopal 462030, M.P., India
e-mail: dchaurasia@spabhopal.ac.in

S. Sankat
e-mail: sandeepsankat@spabhopal.ac.in

S.K. Solanki
e-mail: sushil@spabhopal.ac.in

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_66

Indian non-metro cities, the average trip lengths are less than 5 km. Thus, the traditional city form in the Indian context has been sustainable, something that is now changing on account of a high cost, energy-intensive and skewed urban development paradigm that supports only economic growth as against overall well-being. To support such a development paradigm, cities are moving towards a physical sprawl, which is low-density housing development. Providing mobility in this situation requires establishing capital-intensive transport infrastructure, which is pricing out the poor from the transport system as well as from the land market, and in short reducing the accessibility of the poor to urban opportunities. Women of low-income households are the worst affected in this paradigm change [1].

Inclusive sustainable cities, argues Mahadevia (2001), are the ones which have an 'Inclusive Approach', with the vision of the urban poor and marginal sections at the centre of urban policymaking and which is multidimensional in nature. This is also necessary with regards to looking at transport systems.

Another approach to sustainable transport is high-speed public transport, which also forces the poor to spend large amounts on transit. Thus, the poor tend to depend on the NMT modes. The BRT system emerges therefore in this context of developing countries. The BRTS has the potential to provide equitable and sustainable urban public transit on account of its low cost relative to the metro [1].

2 Concept of Accessibility

Accessibility is a key element in ensuring the social sustainability of the transport sector. Accessibility (or just access) refers to the ease of reaching goods, services, activities and destinations, which together are called opportunities. It can be defined as the potential for interaction and exchange. Litman (2011: 1) defines accessibility as people's ability to reach goods, services and activities. "Accessibility can be defined as 'ease of reaching'.

Any transport system is to achieve equity or social inclusion by increasing its accessibility universally, and then it needs to meet the following objectives [1]:

1. Cost effectiveness—affordable pricing of the systems
2. Safety—against physical accidents, against personal violence
3. Minimised environmental impacts
4. Improve well-being through reducing drudgery, increasing leisure time and comfort during travel through low or no crowding
5. Linkage with livelihood
6. Route frequency and passenger stops to board or alight within certain distance or time of walking
7. Reliability in times of bad weather, for example monsoons in India
8. Disabled-friendly

3 Purpose of the Study

People with disabilities have traditionally been excluded from public life and public services, such as transportation. Through the turbulent social movements of the 1960s and 1970s, disability civil rights emerged in the developed countries of the United States and the United Kingdom to challenge the notions of inclusion and human capabilities. Access to public transport was an important demand of those nascent disability movements. While countries were beginning to work on fully recognizing the rights of people with disabilities, improvements in bus systems were being implemented that would shift the paradigm of mass transit [2].

In developing cities prior to 1972, buses plied shared roadways with mixed traffic. Starting in Curitiba, Brazil, a new manner for buses to traverse the city was envisioned. Curitiba's architect cum Mayor, Jamie Lerner, implemented an above ground subway system that was run using buses instead of rail. Key aspects of Lerner's designed system include dedicated right of way for buses so that they were not slowed by other vehicle traffic and bus station entrances that were restricted to passengers who had paid their fare.

The success of BRT is related to the implementation of a comprehensive package of physical infrastructure and operational systems that together provide a high level of customer service. The package includes the following:

1. Dedicated lanes for the exclusive use of BRT buses,
2. Level boarding to ensure that passengers can get in and out of a bus without having to climb steps,
3. Specially designed high quality buses with wide doors,
4. Off-board fare collection,
5. Enhanced intelligent transportation systems (ITS) including automatic vehicle tracking system, and
6. Service level agreements that stipulate penalties for poor performance.

4 BRT Projects in India

The National Urban Transport Policy (NUTP) of 2006 has came up as an outcome of debates. Nationally, the JnNURM (Jwaharlal Nehru National Urban Renewal Mission) has been the key trigger for cities to start preparing the BRTS proposals for funding under the JnNURM. A number of cities have proposed BRTS and are in various stages of development. At least 50 % of the financial assistance has been provided by the Government of India, and the rest has to be managed by the cities and the states.

The JnNURM. Launched in December 2005, the JnNURM covered 63 cities with funds worth USD 11 billion of central assistance with matching grants coming from the state and city governments. The JnNURM was followed by the National

Table 1 Approved BRTS projects under Jawaharlal Nehru National Urban Renewal Mission (JnNURM)

S. No.	City	Approved km	Sanctioned cost (in Rs. billions)	ACA (in Rs. billions)
1	Ahmedabad	88.50	9.82	3.44
2	Rajkot	29.00	1.10	0.56
3	Surat	29.90	4.69	2.34
4	Bhopal	21.71	2.37	1.18
5	Indore	11.45	0.98	0.49
6	Pune and Pimpri Chinchwad	124.77	13.64	6.82
7	Vijaywada	15.50	1.53	0.76
8	Vizag	42.80	4.53	2.26
9	Jaipur	26.10	2.19	1.10
	Total	389.73	40.84	18.95

Source Ministry of Urban Development (2008)

Urban Transport Policy (NUTP), 2006, which emphasised safe, affordable, quick, comfortable, reliable, and sustainable accessibility for India's growing number of city residents to jobs, education, and recreation within cities. Since the beginning, about ten Indian cities have shown interest in the idea of the Bus Rapid Transit (BRT) and many of these cities are in various phases of either planning or implementing the system (Table 1).

5 BRTS Challenges and Opportunities

The rapid spread of Bus Rapid Transit (BRT) systems presents an historic opportunity to create models of accessible transport for passengers with disabilities and for older passengers, often in cities with little previous experience in this field. BRT trunk line corridors and their feeder lines can enable new categories of passengers, including more women and children, to benefit from an improved level of safe, accessible, and reliable public transport.

In order to forecast demand for BRT service by persons with disabilities it is important to be able to count passengers with hidden disabilities, including those who are frail or have a vision impairment or have arthritis, a heart condition, or are deaf, deafened, or hard-of hearing [3].

6 Social Concern—Accessibility and Equity

There is another concept of 'social inclusion' which largely discusses the equity issues in transport. Social inclusion issues in transport are concerned primarily with accessibility or lack of it for those without a car in the Western context. In a wider context, social inclusion in transport has to be pegged with broader social inclusion policies which deal with inclusion of specific social groups in the process of development. Hence, social inclusion discussions in general, and now even in the transport sector in particular, focus on the processes of exclusion of certain urban population segments from the benefits of urban growth. However, the Western concept of social inclusion may not explain the Indian situation fully as the nature of barriers are different in Indian cities, which is discussed further in the following section.

To summarise, if any transport system is to achieve equity or social inclusion by increasing its accessibility universally, then it needs to meet the following objectives [2]:

1. Cost effectiveness—affordable pricing of the systems
2. Safety—against physical accidents, against personal violence
3. Minimised environmental impacts
4. Improve well-being through reducing drudgery, increasing leisure time and comfort during travel through low or no crowding
5. Linkage with livelihood
6. Route frequency and passenger stops to board or alight within certain distance or time of walking
7. Reliability in times of bad weather, for example monsoons in India
8. Disabled-friendly.

7 BRT Features

The scope and scale of BRT interventions offer planners an important opportunity to make public transport infrastructure easier to use for seniors and people with disabilities.

7.1 Pedestrian Movement

BRT systems are far more than dedicated bus lanes and special buses. Typically, implementing a BRT system involves significant improvements to the pedestrian environment along the right of way. These external corridor improvements offer the greatest benefits for people with disabilities. While an accessible and thoughtfully aligned BRT system offers increased connections and mobility, public transport users with disabilities often face great challenges getting to stops and stations.

The pedestrian environment should thus be compliant with local accessibility design guidelines including consistent pavement surfaces, appropriately sloped ramps, and an unobstructed path of travel. Tactile indicator tiles should be positioned at level changes such as curb ramps, stairs or any other locations where the pedestrian path of travel intersects with a potential hazard (i.e., moving vehicle or bicycle traffic).

For people who are blind or experience low vision, tactile way finding paths are often recommended to guide users through pedestrian environments to key destinations such as transport station entrances.

Accessible pedestrian bridges built with ramps instead of stairs can result in fatigue for many passengers and difficulty for use by older persons and others with mobility concerns.

From the standpoint of access by persons with limited mobility, the solutions in ranking order are [3]:

- 1st choice: At-grade crossings controlled by traffic lights
- 2nd choice: Pedestrian bridges or tunnels equipped with elevators
- 3rd choice: Pedestrian tunnels with inclined ramps built to international access standards
- 4th choice: Pedestrian bridges with inclined ramps built to international access standards.

7.2 Boarding Stations

The access route to the platform also must be free of obstacles, including curbs, bollards, turnstiles or other abrupt changes in level. All vertical height transitions must be ramped in a manner compliant to disability access standards. Within the station itself, it is important that signage for vehicle arrival and transport system information is appropriately large type size and high contrast so that it is easily legible. Auditory announcements for vehicle arrival times and other essential public transport information may also be provided for customers who are unable to see or read visual signage [2].

7.3 Transport Vehicle

Reducing the horizontal and vertical gaps between the transport vehicle and the platform assists boarding for able-bodied customers and improves overall system performance Kantor et al. (2006). It also allows boarding for transport customers who are unable to climb stairs (such as wheelchair users, or seniors with limited mobility).

Training, skill certification (and recertification) as well as disciplinary penalties for poor performance should be utilized to ensure proper BRT docking and vehicle accessibility [2].

People with mobility disabilities, such as seniors who are unable to stand for long periods of time, require seating; however, those who must utilize mobility devices and assistive technologies such as canes, crutches, walkers or wheelchairs may require additional interior space for positioning and navigating their devices. The minimum clear space requirements for positioning mobility devices should be determined by following local disability access standards. The vehicle should also have provisions to strap and secure the mobility device.

For transport customers who experience sensory disabilities such as low vision, blindness or hearing impairment, etc. it is important to provide vehicle location and system information (i.e., next stop announcements, updates on system delays/detours, or emergency egress instructions) in alternative formats. Thus, auditory announcements as well as dynamic visual text systems are recommended.

8 Bhopal City BRTS Brief

BRTS stands for Bus Rapid Transit System; Here I want to emphasis on the word ‘System’ that represents an integrated approach to develop not only dedicated lanes for buses but to provide safe and comfortable corridors for pedestrians, cyclists, motor vehicles etc. The important elements of BRT system are bus stops, Foot Over Bridges, Pedestrian Subways, platform, curbs, railings, Public Information System, Pedestrian Crossing Signals, Signages and road markings should be passenger/user friendly of all age groups (Old age, Children), gender and people with varied physical conditions (Pregnant Woman, Wheel Chair Bound Person, Vision Impaired) etc. In, other words BRT system provides us an opportunity to develop our cities in a holistically manner, so that anyone can use it with pride.

Bhopal BRT system is funded by central government under JNNURM scheme. Constructed by Bhopal Municipal Corporation (BMC) funds allocated are around 237 crores. Bhopal BRT (MyBus) operating agency is Bhopal City Link Limited (BCLL). Most of the operational BRT systems in India are primarily meant to connect sub-urban parts of the city, Bhopal BRT system is passing through the main city and market areas supported by Trunk, Standard, Complimentary and Intermediate Para Transit (IPT) routes. The existing ‘Mini Buses’ and ‘Magic’ is going to use the complementary and IPT routes to provide transport services for passengers from inner residential area to main trunk route (TR) and standard routes (SR) i.e., BRT routes.

Fig. 1 Above photographs showing manually operated pedestrian movement on zebra crossing and condition of newly constructed curb and ramp showing poor quality (near *Habibganj* railway crossing, Bhopal). Images showing *encircled* issues like installed pedestrian crossing light (not working) at *Misrod* area. *Source Author*

8.1 Bhopal BRTS Safety Concerns

During the construction phase of four years more than 1600 road accidents occurred on this BRT stretch and around 115 people died. As per Bhopal municipal corporation (BMC) because of the dedicated corridor the accident rates and causalities has come down. Local Bhopal city expert has also expressed concern about the safety of the passenger who wants to access the bus stops, BMC has plan to built pedestrian signal and foot over bridge (FoB), so we need to spread the awareness amongst people to follow the traffic rules to reach the BRT bus stops safely. Foot Over Bridges (FOBs) are an excellent safe means for commuters to across movement of roads but it requires investment and time to build it. Bhopal BRTS is a good initiative to handle the traffic situation for future of the city. Bhopal BRT will be definitely a life line for city for next 20 years and the system is sufficient to cater the need of public transport demand for next two decades. For that we require good quality construction and BRT system design should follow the 'Indian Road Congress (IRC)' recommendations and internationally accepted good practices [4] (Fig. 1).

8.2 Survey Results of the Bhopal BRTS

To analyze BRT system of Bhopal, I have conducted an observational study in the month of October 2013 (after the formal inauguration of BRTS) to assess the quality, condition and availability of various elements and services in all stretches (Routes between two BRT bus stops) of 24 km long BRT corridor starting from Misrod Area (Hoshangabad road) to Bairagarh Area there are 41 Nos. of stretches in BRT corridor. I have prepared a 'matrix' between elements, infrastructure, services (Placed Horizontally) and 41 Nos. of route stretches (Placed Vertically). I have taken into consideration the 'Qualitative Aspects' of following 15 indicators to observe BRT type, BRT corridor, carriage way, bicycle track, pedestrian pathway, bus stop, ramp, railing, Public Information System (PIS) at Bus stops, pedestrian crossing signals, zebra crossing, signage, road markings, tactile paving at bus stops and to approach bus stops [4].

In 24 km long BRT corridors i.e. 'Misrod' and 'Bairagarh' at both the end stretches are 'satisfactory' in terms of development because of the availability of road spaces. When we enter in 'New Bhopal city' from hoshangabad road the satisfactory level of services and infrastructure is getting low because of the major central business districts (CBDs), traffic is more and most of the BRT stretches are not having dedicated central lanes so buses are shearing the road spaces with other vehicular traffic. I found satisfactory level of services and infrastructure is 'Zero' in 'Old Bhopal' because of unplanned organic growth, reduced road spaces, heavy traffic and only shortest way to connect new Bhopal to other parts of Bhopal through old Bhopal.

8.3 Problems with BRTS Infrastructure

We have conducted a hand on exercise with the students of School of Planning and Architecture, Bhopal to analyse the problems associated with disabled persons. Below photographs showing the condition of BRTS with reference to disabled persons (Fig. 2).

Large gap between bus and BRT station without boarding ramp creates problem to person on wheelchair, persons on crutches, old age, children and women with Indian dress (Saari). Footpaths without ramps and height make it difficult to use by vision impaired persons, the absence of tactile tiles as a guide which leads to other spaces for vision impaired are missing or discontinued in various areas makes it of no use. Bollards are placed on the way, in between the end of ramp and zebra crossing with a concrete curb makes a wheel chair bound person unable to move from BRT bus stand to other side of the road through uncontrolled zebra crossing with no traffic stop signal [5].

Fig. 2 Photographs showing problems associated with BRTS users with disability. *Source* Author

9 The Way Forward

Provisions through codes like Indian Road Congress (IRC: 86-1983 and 103-1998), Guidelines and Space Standards for Barrier Free Built Environment for people with disabilities and elderly persons, 1998 have aided transportation and design professionals with empirical ways of incorporating universal design (UD) principles. A multilevel approach at five levels viz. Policy, Accessibility Codes, Technology, Implementation and public participation holds the key to a comprehensive application of accessibility concepts in future. Development of inclusive paratransit models and travel chain, accessible tourism and walkability of pedestrians with diverse abilities remain some of the core areas for future UD research in India [6].

10 Conclusion

The presentation investigates how our policies are friendly with user of all age groups and persons with disabilities and the study is focused to investigate, indentify and suggest the solution for existing BRTS. Agencies planning new BRT systems should seek out empirical studies on those elements, the accessibility benefits of which are not certain. These agencies also should involve the disability community, as well as agencies that work with people with disabilities at every step of the way to help determine what features are in demand and what enhancements would work best. The planning and public involvement processes that are involved with BRT system implementation should be inclusive. Provisions through codes like Indian Road Congress (IRC: 86-1983 and 103-1998) and Guidelines and Space Standards for Barrier Free Built Environment for people with disabilities and elderly persons, 1998 have aided transportation and design professionals with empirical ways of incorporating UD principles.

References

1. Mahadevia, D., Joshi, R., Dutey, A.: Low Carbon Mobility in India and the Challenges of Social Inclusion. Bus Rapid Transit (BRT) case studies of India. CEPT University, Ahmedabad (2013)
2. Herlekar, V., Kost, C., Osborn, J.: The people's way: assessing Janmarg BRT for inclusive design. In: Proceedings of Conference on Transed, New Delhi, India (2012)
3. Tom, R.: Technical and Operational Challenges to Inclusive Bus Rapid Trasit: A Guide for Practitioners (2010)
4. Chaurasia, D.: Bus rapid transit system (BRTS): a sustainable way of city transport (case study of Bhopal BRTS). IJEAT **3**(4) (2014). ISSN: 2249-8958

5. Report on Road Safety Audit of Pilot BRTS Corridor at Pune, India, Prepared by Transportation Systems Engineering Group, Department of Civil Engg., IIT Bombay (2014)
6. Raheja, G., Jaiswal, S.: Universal design for Indian public transportation system. In: Proceedings of International Conference UD 2014, Lund, Sweden (2014)

Ergonomic Assessment and Evaluation of Philippine Buses for Filipinos: A Case Study on Metro Manila City Buses

Dyan Veronica Bombais, Janina Angeli Ferrer, Arrenzo Paul Perea and Alyssa Jean Portus

Abstract Public Utility Buses (PUBs)—both provincial and city buses—are highly used mode of transportation in the Philippines. However, people are put in risks due to (1) problems in the transportation system such as intense traffic, inefficient system and overcrowded vehicles, and (2) both buses are imported from neighboring Asian countries, which could imply that those are not fit for Filipino use. The study aimed to conduct an ergonomic assessment and evaluation of Metro Manila buses and create an ideal bus layout with dimensions generated from anthropometric measurements of Filipinos. A passenger satisfaction survey was conducted among commuters to understand their concerns in the current operations and it was found that majority have issues with regard to bus spaces and environment. Another survey was conducted based on the results of the first survey, asking Filipinos on their preferred bus layout—one that would give them an optimal passenger experience.

Keywords City buses · Metro manila · Bus layout · Public utility buses

1 Introduction

Metro Manila, being a central business district in the Philippines, has continuously contributed to the worsening of transportation condition year after year. Four main sources of the problems in the Philippines identified by the international development magazine, *Development and Cooperation*, includes "the lack of good public transport". This does not only include the transportation system that prolongs traffic but also the passengers' experience while riding the vehicles while stuck in traffic, thus, putting the passengers at risk due to stress and uneasiness. From this, it could

D.V. Bombais (✉) · J.A. Ferrer · A.P. Perea · A.J. Portus (✉)
Department of Industrial Engineering and Operations Research, University of the Philippines, 1101 Diliman, Quezon City, Philippines
e-mail: aaportus@up.edu.ph

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_67

be implied that ergonomics can be the key in solving the said problem. Moreover, the Filipino culture and their attitudes towards public transportation should also be examined since Filipinos have been accustomed to bus drivers and conductors accommodating as many passengers as possible which increases the risk even more.

According to 2015 data from the Land Transportation Franchising and Regulatory Board (LTFRB) [1], there are a total of 156 bus operators handling 5057 PUBs that serve the Metro Manila population of nearly 12 million. Since they largely cater to the transport needs of the metropolitan area, buses should be designed to provide a safe, comfortable, and desirable status and image in order to be conducive to ridership. However, a good number of public buses currently being used have only been imported from neighboring Asian countries which although is beneficial in financial terms for the company, may cause the health of the users in the long run due to ergonomic factors since imported buses were not designed based on the measurements of its intended users. This led the researchers into proposing an ergonomics research on Philippine buses, particularly those travelling within Metro Manila.

The study focused mainly on the generation of layout and dimensional redesign of city buses using anthropometric measurements and user inputs; controls placement, special population design, materials, and other bus types were to be subjected to further study due to limited time and resources.

Application of ergonomics in everyday tasks does not only lower stress levels but also reduces the risk of having musculoskeletal disorders (MSDs). According to the research of Saba et al. [2], physical dimensions of any product must be manufactured based on the anthropometric measurements of its target users to ensure that it follows an ergonomic design. In doing so, health improvement and user comfort can be achieved.

According to Saba et al. [2], most manufacturers of vehicles "design the seats to suit the expectations of their customers without due consideration to the comfort and safety of its passengers"; and that they see their business as "art rather than engineering." This indicates that more weight is given to customer wants such as aesthetics and less on customer needs like ergonomics when designing public transportation vehicles.

In designing buses, Mean [3] suggested that ergonomics and safety should be the two key considerations in developing a concept for an ideal city bus; while review of related studies and legal requirements concerning bus operations, along with market studies and passenger input on their preferences with regard to bus design, are the important information needed in order for the study to be valid.

2 Methodology

The following techniques were used in the study for the assessment and evaluation of the ergonomics of City Buses operating in Metro Manila:

2.1 Passenger Satisfaction Survey on City Buses Used Currently in Metro Manila

The goal of the survey was to determine the main concerns of the bus passengers that must be addressed in the study. Factors regarding both space and environment that affect bus rides were evaluated by the respondents through a scale dependent on how they agree or disagree with the statements based on their experience.

2.2 Anthropometric Measurements for Ergonomic Bus Dimensions

Del Prado-Lu [4] anthropometric measurements was used as a basis in evaluating the ideal dimensions of the bus using the passengers' anthropometric measurements. Key anthropometric measurements to be used in the research are the following:

(a) Sitting positions
(b) Standing positions
(c) Foot anthropometric measurement.

2.3 Dimensional Measurements of Current City Buses in Metro Manila

The measurements taken from a bus operating in Metro Manila (from Company X) was used in comparison with recalculated bus dimensions based on anthropometry and was assessed accordingly. This determined whether the buses being utilized are ergonomic and fit for Filipino use. To re-design the bus, the parts illustrated in Fig. 1 were gotten.

2.4 Passenger Preferences on Ideal City Bus Layout

Questions regarding passenger preference on bus layout, aesthetics, and other bus services were based on the data gathered from the first survey. Insights were collected in the second survey for the refinement of the final layout considering the addition of different features. Additional concerns that may be subjected to further studies are also identified.

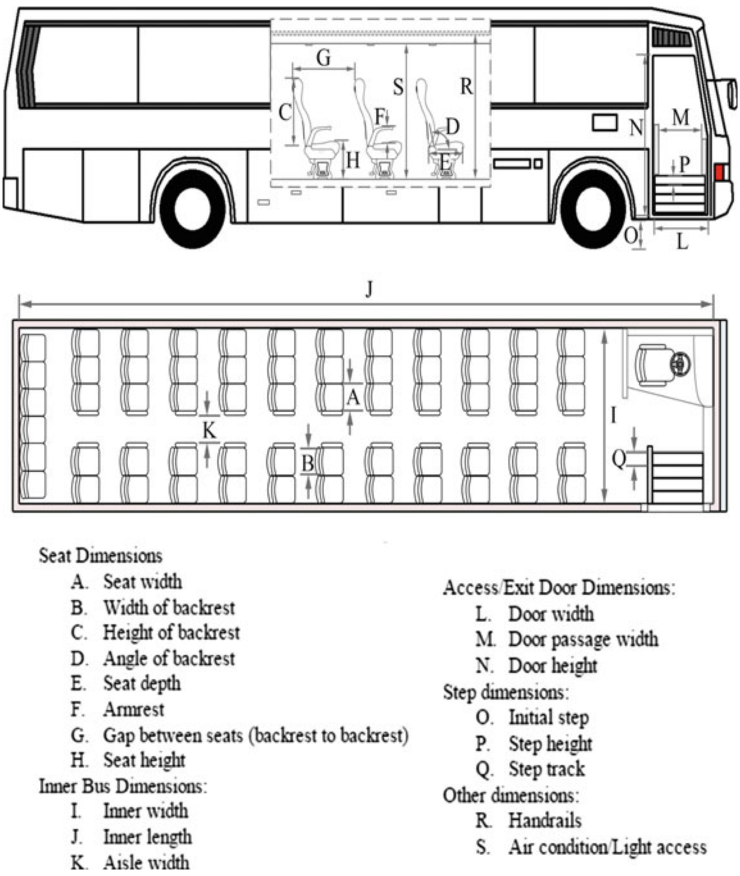


Fig. 1 Bus dimensions

3 Results and Discussion

3.1 Passenger Satisfaction Survey on City Buses in Metro Manila

The researchers gathered 146 responses in the survey, considering both riders and non-riders of city buses. Different sets of survey were given to both groups to account for knowledgeability difference about the topic. The respondents were composed of 101 bus riders (69.2 %) and 45 non-bus riders (30.8 %) aged 15–38. The average number of times they take the bus in a week is 3.45 with 67.63 min per ride.

To compute for the accuracy of the data gathered, the researchers used the formula:

$$n = N * X/(X + N - 1). \quad (1)$$

where

X $Z_{a/2}$ * p * (1 − p)/ME
$Z_{a/2}$ critical value of the normal distribution at a/2
ME margin of error
p sample proportion
N population size.

An error of ±8.5 % is observed when the sample population is set at 50 % and the confidence level is 95 % where a = 0.05 and critical value = 1.96. Since the number of total respondents obtained is greater than the computed sample size (n = 133), the data, therefore, is considered valid. The researchers used information gathered from the survey to understand the demographics' wants and needs. Below were the results of the survey:

Non-bus Riders

a. Different modes of transportation were preferred, specifically cars, jeepneys, and trains, rather than PUBs.
b. Top concerns in riding city buses were: safety, cleanliness, overall bus experience (relaxation inside the bus), buses' passenger capacity, and space for movement inside, including entering and exiting the bus. However, 93.3 % of them said they would consider riding a bus if their concerns were addressed.

Bus Riders

a. Eighty-three (83) riders prefer to ride in an air-conditioned bus because of the cool temperature, comfort, and safety. Almost eight percent (8 %) of the respondents ride ordinary buses because of cheaper fares and its faster travel time. The remaining 9.9 % have no preference when it comes to which kind of bus they ride.
b. Although majority (65.6 %) answered 'no preference' when it comes to bus lines that they usually ride, a bus liner topped the list of the remaining 34.4 % of the respondents. One reason was because of the priority they give to the special population (e.g. elderly, pregnant women).
c. The following table contains the different concerns in riding city buses, divided into two categories: bus space and bus environment. Each were based on bus design guidelines from Napper et al. research [5]. The respondents were given a scale of 1–5 (1 being ideal and 5, otherwise) to rate each bus concern. The average satisfaction rating per concern was computed to determine the concerns that need to be addressed (Table 1; Fig. 2).

Based on the satisfaction rating presented, eight bus concerns garnered an average value greater than or equal to three. These values indicate that passengers

Table 1 Summary of bus satisfaction concerns

Category		Concern
Space	A	There is an ease in entering and exiting the bus
	B	The bus accommodates a reasonable number of passengers
	C	There is sufficient leg room
	D	Handrails are easy to reach
	E	The seats are not cramped and there is enough personal space
Environment	F	The bus smells fresh
	G	The bus temperature is just right
	H	Driver/conductor announcements can be heard clearly
	I	The bus ride does not make you dizzy
	J	The bus is well-maintained
	K	The bus journey is relaxing
	L	The lighting inside the bus is sufficient
	M	There are enough safety measures taken inside the bus

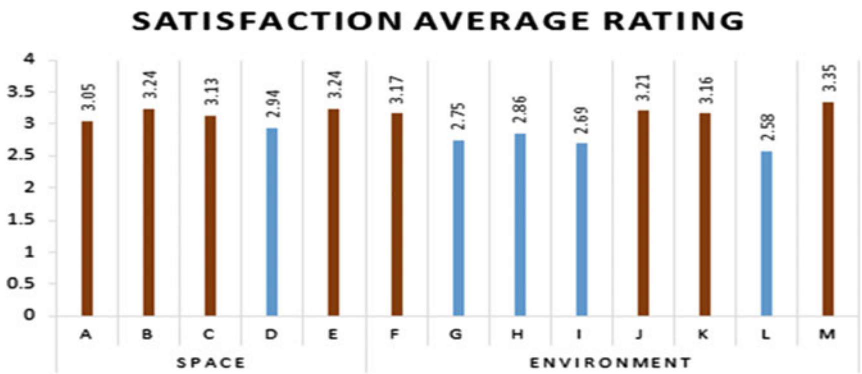


Fig. 2 Satisfaction average rating from survey

were dissatisfied with the respective concerns. Fifty percent (50 %) of the 'bus environment' concerns received a high average while 80 % of the 'bus space' concerns were highly rated. This is an implication of the need for re-evaluation on the ergonomics of the bus, mainly on its dimensions.

3.2 Anthropometric Measurements for Ergonomic Bus Dimensions

Various measurements were taken from Del Prado-Lu's study [1]. Different percentiles were used in determining the appropriate measurement to ensure that the

Table 2 Anthropometric measurement specifications

Bus part	Specific part	Body measurement and population percentile
Individual Seat dimensions[a]	Seat width	Hip breadth, sitting (95th/M) + clothing allowance
	Width of backrest	Hip width (95th/M)
	Height of backrest	Shoulder height, standing—waist height, standing + waist height, sitting (95th/M)
	Seat depth	Buttock popliteal length (50th/M)
	Armrest	Elbow height (50th/F)
	Gap between Seats (backrest to backrest)	Buttock knee length + foot length (95th/M)
	Seat height	Popliteal height (50th/F) + shoe allowance
Inner bus dimension	Aisle width	Hip width (95th/M) + clothing allowance
Access/exit door	Door width	Hip width (95th/M) + allowance + door folding width
	Door passage width	Hip width (95th/M) + clothing allowance
Bus step	Step height	Step height (50th/M) + allowance
Others	Handrails	Upper reach
	Air conditioning unit/light access per seat	Overhead fingertip reach, sitting

[a]Accounted for clothing and shoe allowances (2 and 4 cm, respectively) from Novabos and Po (2012)'s study *The Application of Filipino Anthropometric Data in the Design of House Rooms and Furniture*

dimensions to be proposed are to be correctly accounted for. Table 2 shows the different bus parts to be analyzed and the respective anthropometric measurements used.

3.3 Comparison Between Anthropometric Measurements and Dimensional Measurements of Current Buses

The researcher was able to measure the dimensions of an ordinary (non-air conditioned) city bus. Table 3 contains measurements of Company X were obtained and were compared side-by-side with the ideal measurements computed from the anthropometrics Filipinos. Most measurements incurred a negative difference which indicates that all seat measurements are not ergonomically designed for Filipino use.

Table 3 Existing bus measurements versus proposed bus measurements (in cm)

Bus part	Specific part	Measured dimension	Anthropometric dimensions	Difference
Individual seat dimension	Seat width	31	43	−12
	Width of backrest	31	54.8	−23.8
	Height of backrest	61	67	−6
	Angle of backrest	–	105°[a]	
	Seat depth	44	46.4	−2.4
	Armrest	21	21.89	−0.89
	Gap between seats (backrest to backrest)	71	89.9	−18.9
	Seat height	40	44.5	−4.5
Inner bus dimensions	Inner height	183	–	–
	Inner width	230	–	–
	Inner length	1003	–	–
	Aisle width	70	64.8	5.2
Access/exit doors	Door height	209	–	–
	Door width	73	94.8	−21.8
	Door passage width	48	64.8	−16.8
Bus step	Initial step	44	39	−5
	Step height 1	27	30	3
	Step height 2	25	30	5
	Step height 3	23	–	–
	Step height 4	24	–	–
	Step track	24	28	−4
Others	Floor to handrail	176	177	−1
	Air conditioning unit/light access per seat	–	108	–

[a]Data from “sitting and chair design” by Cornell University ergonomics web

3.4 Passenger Preferences on Ideal City Bus Layout

The researchers developed bus design and layout concepts to be subjected to public opinion. Since it is common for bus operators to maximize the capacity of the vehicle by allowing standing passengers, the researchers considered placing a standing area.

Layouts were the main contents of the second survey, which aimed to determine passenger preferences with respect to different factors found important according to the first survey. Insights were also gathered so that refinement of the design would be possible and input from respondents could be taken into account. This helped the researchers identify the ideal layout of city buses in the Philippines through side by side analysis with anthropometry and passenger concerns.

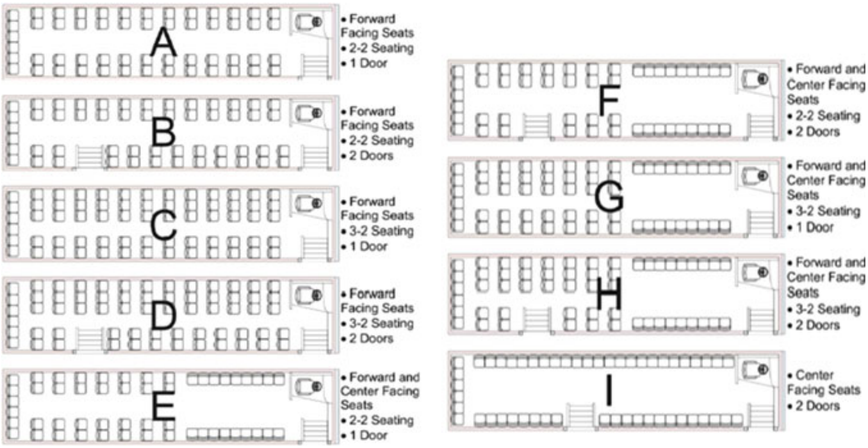


Fig. 3 Bus layouts used in preference survey

The same formula used in the first survey utilized for the second survey, with a recommended sample size of 133. The researchers was able to get 136 respondents, making the survey data valid.

The researchers came up with the layouts as shown in Fig. 3 considering the number of doors, orientation of chairs, seating and standing ratio, and seating and aisle space trade off.

The passengers were also asked of their top three preferred layouts. Layouts B, F, and I topped the list. Weights were given for each rank to account for the ranking of the respondents. After the computations, bus layout F was determined to be the most preferred layout of the survey respondents, with a total weight of 197, followed by layout B and I, with respective weights of 145 and 136, making bus layout F the basis of the final design of the proposed ergonomic city bus.

It was also found from the survey that bus features that passengers would like city buses to have are the following: air conditioning units, buzzers to signal drivers, CCTV cameras, Wi-Fi, curtains, and televisions. These features were chosen by more than half of the respondents. These were recommended by the researchers since these factors contribute to the overall satisfaction of the bus riders. Although some features may not be possible to have in the Philippines due to country policies, it was still listed for future references.

4 Conclusion and Recommendations

After data collection and analysis, it was found that current bus designs of some city buses are not fit for Filipino use. The researchers proposes two layouts: Conventional Layout and Ergonomic Bus Layout. The conventional layout only

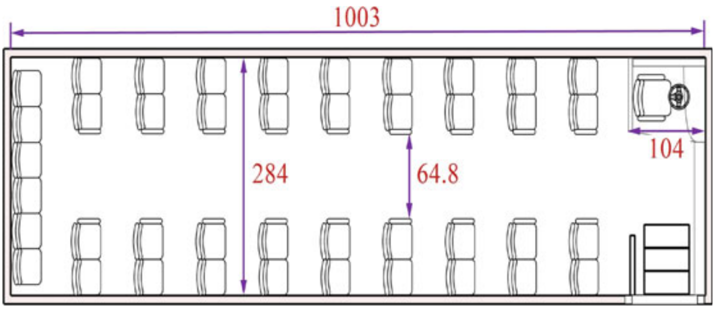


Fig. 4 Proposed conventional layout

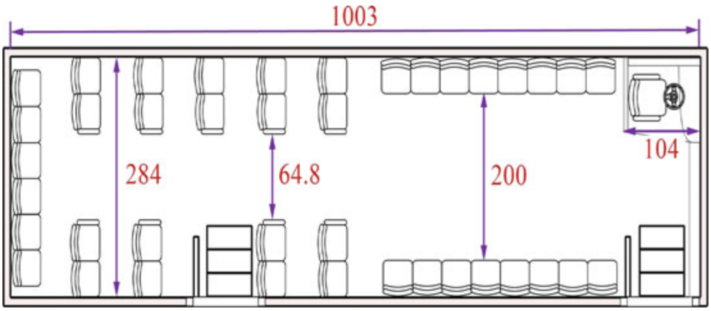


Fig. 5 Proposed ergonomic bus layout

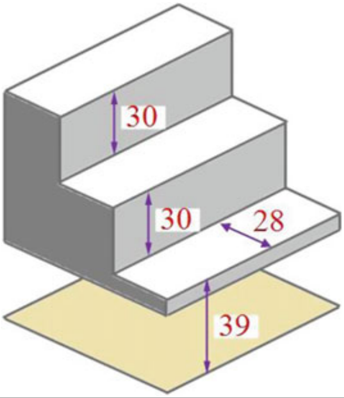
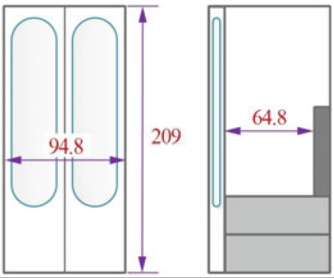
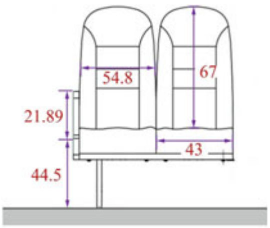
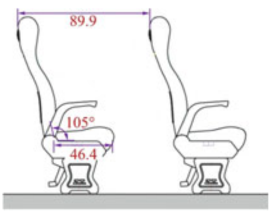
considered the anthropometric measurements while the ergonomic bus layout also considered layout preferences of commuters.

Easier and faster adjustments can be done by the company and the passengers in keeping the conventional bus layouts. However, this would require the adjustment of the body width of buses. Adjusting the bus length was not considered since it is unnecessary to increase it by a few centimetres. The 2 × 2 seating layout was chosen since 3 × 2 seating is not possible given the anthropometric measurements declared. For the Ergonomic Bus Layout, Layout F was the most preferred by the respondents considering ease of access, comfort, and overall presumed satisfaction. The researchers created a layout based on both the anthropometrics and passenger preferred layout. The bus dimensions shown in Figs. 4 and 5 will be followed for the two bus layouts respectively.

For the specifics—bus parts such as seats, doors, stairs, and other features, the researchers standardized the ideal dimensions (in centimeters) which were followed in both layouts. The bus parts are as shown in Table 4.

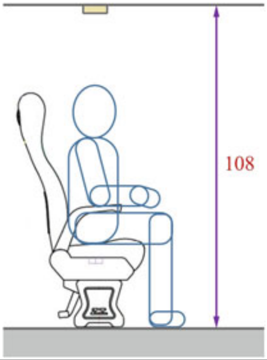
For bus features, it is suggested that existing features be retained (such as air conditioning units and television) while some be added, as mentioned by passengers. A signal for the driver to halt and/or drop off a passenger was recommended for riding efficiency. Presence of Wi-Fi was also ideal to keep the passengers

Table 4 Proposed bus part dimensions

Access/exit steps  	Aside from the initial step, only two steps were added. This was possible since the height of each step was adjusted accordingly to account for the height of the tire without sacrificing the safety and ease in entering and exiting of passengers
Access/exit doors  	The width of the doors has been increased to give an allowance for the folding of the door as it opens for passengers. The recommended width allowed passengers to alight and leave the vehicle comfortably and safely
Seat dimensions   	As for the seats, it is recommended that all dimensions to be used in the seats were mainly based on the anthropometric measurements of Filipinos to ensure comfort. A reasonable gap between seats must be placed in order to have an allowance for legroom for passengers during trips. Cushions and other materials should be carefully selected to provide optimal seating experience for the passengers

(continued)

Table 4 (continued)

Air conditioning/light access 	Air conditioning or light access area is usually directly placed on top of the passengers. The height was adjusted for the comfort of the passengers and ease of use of the features located above

entertained during long trips. The use of CCTV cameras has been passed as a bill in the congress and was recommended as well to promote passenger safety. Curtains, although preferred, cannot be implemented due to the country's safety policies.

Acknowledgments We thank God, for giving us the strength and persistence to do this research; our respective families for their endless support;

To Company X, Hernandez Family, and Sir Alvin Baya for accommodating us and letting us examine their bus units; and to Dr. Hussein Lidasan and Mr. Mark Victor Bautista for sharing their insights on the transportation and design system related to our study.

References

1. Metro Manila bus operators: land transportation franchising and regulatory board (LTFRB) (2015). Retrieved 4 Nov 2015. ltfrb.gov.ph
2. Saba, G., et al.: Ergonomics assessment of passenger seats of mini-buses in Ethiopia. Adama Science and Technology University, Ethiopia (2013)
3. Mean, W.: Development of concept design for an ideal city bus. University Technical Malaysia Melaka, Melaka (2010)
4. Del Prado-Lu, J.: Anthropometric measurement of filipino manufacturing workers. University of the Philippines Manila, Manila (2007)
5. Napper, R., et al.: Bus design guidelines: complementing planning with vehicle design. Monash University, Melbourne (2009)

QZTool—Automatically Generated Origin-Destination Matrices from Cell Phone Trajectories

Christopher Horn, Heimo Gursch, Roman Kern and Michael Cik

Abstract Models describing human travel patterns are indispensable to plan and operate road, rail and public transportation networks. For most kind of analyses in the field of transportation planning, there is a need for origin-destination (OD) matrices, which specify the travel demands between the origin and destination zones in the network. The preparation of OD matrices is traditionally a time consuming and cumbersome task. The presented system, QZTool, reduces the necessary effort as it is capable of generating OD matrices automatically. These matrices are produced starting from floating phone data (FPD) as raw input. This raw input is processed by a Hadoop-based big data system. A graphical user interface allows for an easy usage and hides the complexity from the operator. For evaluation, we compare a FDP-based OD matrix to an OD matrix created by a traffic demand model. Results show that both matrices agree to a high degree, indicating that FPD-based OD matrices can be used to create new, or to validate or amend existing OD matrices.

Keywords Transportation planning · Travel demand · Origin-destination (OD) matrices · Floating phone data (FPD) · Big data · Hadoop

C. Horn (✉) · H. Gursch · R. Kern
Know-Center GmbH, Knowledge Discovery, Inffeldgasse 13/6, 8010 Graz, Austria
e-mail: chorn@know-center.at

H. Gursch
e-mail: hgursch@know-center.at

R. Kern
e-mail: rkern@know-center.at

M. Cik
Institute of Highway Engineering and Transport Planning,
Graz University of Technology, Rechbauerstraße 12/II, 8010 Graz, Austria
e-mail: michael.cik@tugraz.at

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_68

1 Introduction

Since the 1960s urban and transport planning methodologies were developed to forecast future traffic volumes and the expected use of infrastructure and facilities. The purpose of such forecasts is evident: infrastructure and urban planning provide keys to the mitigation and preclusion of transport and environmental problems. The availability of spatial data on demographics, labor and land use has until now been a prerequisite for establishing an origin-destination matrix (OD matrix) for a transport model.

According to the traditional four-step transportation forecasting model the first step is trip generation, followed by trip distribution, before mode choice and route assignment. The trip distribution step matches trip makers' origins and destinations to develop a trip table, a matrix that displays the number of trips going from each origin to each destination [1]. Table 1 shows a small example of an OD matrix consisting of four zones and the trips between them.

The general method for obtaining an origin-destination (OD) trip matrix developed for large scale transport modelling employs a combination of home interviews and roadside surveys. However, the cost of this approach precludes its use for most other applications. Surveys are difficult to conduct and costly, therefore they are conducted infrequently. As a consequence, such data might be even one or two decades old for particular regions. Generally, data in OD matrices often tends to be outdated and unreliable [2].

In the late 1990s and early 2000s cell phone ownership climbed up to levels of close to 100 %. Cell phone networks, as the name implies, are organised in cells. This means that a phone network spanning a region is divided into smaller cells. This cellular organisation theoretically allows for a localisation of individual phones with accuracy proportional to the respective cell size. Cells can be as small as just about 10 m in radius up to even more than 20 km. If a single phone is located multiple times (i.e. tracked over time) the movement of it can be monitored. In order to collect the necessary data for OD matrices, it is not the position of a single phone, which is of interest, but the time-space-trajectories of many phones as they represent different trips in different modes. Therefore, individual phones can be traced retaining their anonymity (i.e. exact position and movement) to circumvent privacy protection issues and this will still yield the necessary movement data.

Table 1 Small example of an origin-destination matrix. The numbers represent trips between zones. In this example zones are exemplified with capital letters

		Destination			
		A	B	C	D
Origin	A	0	2	7	32
	B	54	0	84	65
	C	35	25	0	35
	D	78	1	54	0

Floating phone data (FPD) has a number of advantages over the traditional survey approaches: (i) it allows to monitor traffic around the clock, (ii) there is no need to collect and evaluate questionnaires in the streets, and (iii) it is not prone to false or incomplete statements. These advantages come at a price of a higher technical effort. The floating phone data needs to be collected from the transportation-related operators. Moreover, the cellular network only gives information about the cell a phone is currently logged in, but not the exact location of the phone itself. Therefore, a mapping of the cell's location information to the geographical locations on the road and rail network is needed. This process is called map matching [3].

2 Related Work

A typical usage scenario of location information from cell phone data is traffic management. More specifically, if the domain of transportation research is considered, incident detection, forecast of travel demand as well as the generation and validation of transport models are of interest [4]. Kálmán Filters are capable of calculating the trajectory of individual traffic participants given the potentially noisy cell phone data. Aydos et al. [5] show that linear Kálmán Filters are applicable in this setting. To derive a more accurate description of moving vehicle from cell phone data, a more complex model is necessary. The combination of cell phone network characteristics with knowledge road user behaviour yield in such high quality models [6].

The described models are a way to transform cell phone data into an abstract representation of traffic. Depending on the desired usage scenario (e.g. traffic monitoring, crowd control) specialised and detailed models are needed. Calabrese at el. [7] developed a system for real-time traffic monitoring with special features for taxi and bus usage monitoring. Their system consists of four parts (Localization Engine, Tracking Filter, Mobility State Estimator, Traffic Map Calculator) and is capable of rendering traffic visualisations showing different aspects of people movement like public transport demand, traffic flow and attractive sights. Janecek et al. [8] follow a two-staged approach in their works. In the first step a rough estimation of cell phone users' locations is calculated. This estimate is enough to detect areas with heavy congestion in the road network. In the second step a more accurate position is used to characterise the detected event in more detail. As a whole, the system is capable to predict vehicle travel times from cell phone location data.

The next logical step of transport modelling is the derivation of OD matrices from FPD. The idea of deriving OD matrices from cell phone data was first tested with simulations [4, 9]. After this simulation turned out successfully, further system with real world data have been developed. Calabrese et al. [10] designed a test system working with a static dataset. Their system works in two steps. A trip of a cell phone user is estimated and from this trip the source and the destination is extracted and a count is added to the OD matrix. Many works along those lines have been undertaken. Caceres et al. [3] give an overview about systems realized until 2008. Since then, some other works were presented using more elaborate schemes to derive OD matrices. Neglecting the cell phone data aspect for a minute, Li [11] uses Bayesian Inference, and Toledo and Kolechkina [12] use a generalised Least Squares method to generate OD matrices. Iqbal et al. [13] present an approach where FPD is combined with limited traffic count data.

3 System Design

The QZTool is an integrated system that processes large amounts of passively collected cell phone data with the goal of creating near-real-time OD matrices. Figure 1 illustrates an overview of the system. Each component will be described in the remainder of this chapter.

3.1 External Data Sources

The QZTool relies on the three different external data sources cell phone data, OpenStreetMap data and timetables of public transport.

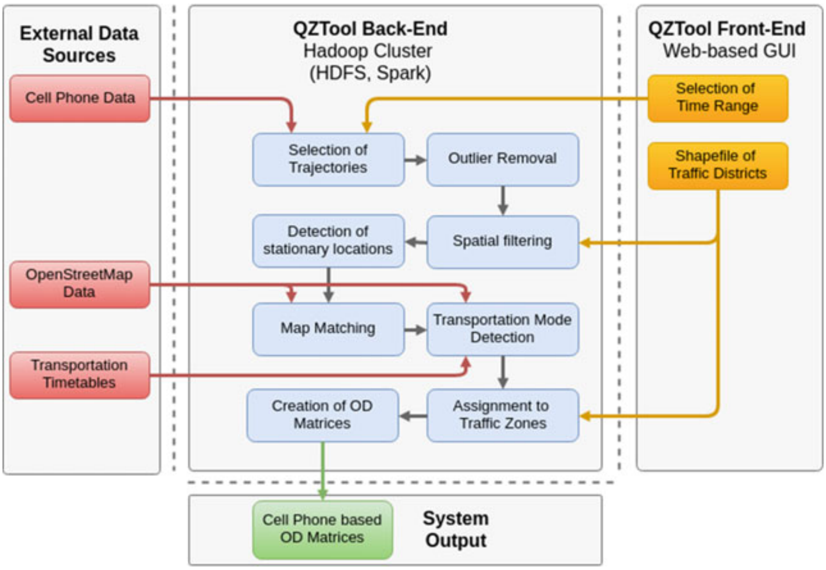


Fig. 1 QZTool system overview. The QZTool is divided into a core component (QZTool back-end) and a web-based graphical user interface (QZTool front-end). The QZTool uses different data sources (external data sources) to generate OD matrices (system output)

Cell Phone Data. This data source is considered to be the main data source of the system. Cell phone data is passively collected at the cell phone provider's infrastructure, meaning that no user interaction (e.g. installing an app) is required in order to collect the data. Cell phone events are triggered when the cell phone changes its area location, or the user initiates or receives a voice call or text message. For this research, we had access to the data of a large Austrian cell phone provider with a market share of more than 30 %. On an average day, more than 30 GB of raw data is generated. The vast amount of data also determines the used technologies. For this system, state-of-the-art Big Data technologies were used. More specifically, the cell phone data was stored in a Hadoop cluster. To process the data, the Apache Spark framework was used. This setup enables us to scale out horizontally if the data size increases even further.

One of the drawbacks of cell phone data is the spatial inaccuracy. Unlike GPS data, the accuracy of cell phone data depends on the type of base station (BS) which the cell phone is currently connected to. Common types are macro cells (range up to 30 km), micro cells (up to 2 km) and pico cells (up to 200 m). In rural areas, most of the base stations are equipped with macro cells. In urban areas, micro and pico cells are more commonly used. However, it depends on various factors which cell type is mounted on a base station. One needs to take this range of uncertainty into account when using cell phone data for transport modelling.

OpenStreetMap Data. The second external data source is OpenStreetMap data. OpenStreetMap (OSM) is a collaborative map created by volunteers. In this system, OSM is taken as source for matching the cell phone trajectories to the underlying road network ("map-matching"). In addition, the data from OSM is used to determine if a cell phone user is located at or near a railway station.

Transportation Timetables. The final external data source are public transportation timetables. They are used as a feature to detect the mode of transportation. This is motivated by the hypothesis that a high overlap of a cell phone trajectory with the public transportation timetable makes it more likely that public transportation is used.

3.2 Input

In order to create a cell phone-based OD matrix, the user has to specify two parameters, namely the desired time range, and the traffic districts for which the OD matrix should be calculated.

Selection of Time Range. This parameter specifies the time span for which the OD matrix will be generated. This enables the generation of matrices for different seasons (e.g. holidays, business days, long weekends).

Shapefile of Traffic Zones. A Shapefile is a commonly used file format to describe geographic entities like polygons. In our system, a Shapefile is used to describe the different traffic zones of the OD matrix. For example, a Shapefile with 5 traffic zones will result in a 5 × 5 matrix.

3.3 Processing

Once the input parameters are defined, the processing is triggered by the click of a button. The computation itself is carried out on a Hadoop cluster infrastructure using the Apache Spark framework. Apache Spark offers fast in-memory computing and parallelization of the work load. The cell phone data itself is stored on Hadoop Distributed File System (HDFS), which provides a seamless integration with Spark.

Selection of Trajectories. In the first processing step, the relevant input data is filtered based on the chosen time span. Giving the huge amount of data, an efficient storage technique should be used, i.e. to avoid full data scans when only a subset of the data is selected. To resolve this issue, we chose to vertically partition the data and to store the data in a directory structure that reflects the date when the data was recorded. This enables us to quickly select the relevant data. For example, the directory */data/2016/03/08/13/* contains all cell phone events that where recorded on March 8th, 2016 from 1 to 2 pm.

Using concatenation and wildcards, one can now analyse arbitrary time ranges. For example, */data/2016/03/01/*,/data/2016/03/02/*,/data/2016/03/03/** would analyse March 1st to March 3rd, 2016.

Once the data is selected, trajectories can be created. This is achieved by grouping the cell phone events by the anonymized user id, yielding to trajectories in the form

$$\mathrm{UserId} = \; > \mathrm{List} < \mathrm{Cell\,Phone\,Events} > . \tag{1}$$

Outlier Removal. In the next step, the created trajectories were cleaned from potential outliers. We define outliers as events which are very unlikely to happen due to large distance and low timespan between two consecutive events. Such events may occur due to measurement errors or may be injected by the network provider to hamper the creation of accurate trajectories due to privacy issues. We applied previously developed algorithms to remove such outliers [14].

Spatial Filtering. After creating and cleaning the trajectories, a spatial filtering is applied. Only those trajectories are kept which intersect at least one traffic zone at some point in time. This is a crucial step since it might heavily reduce the amount of trajectories, depending on the size of the area to analyze.

Detection of Stationary Locations. In this step, stationary locations are determined. In OD matrices, as the name already implies, trips between an origin and a destination are counted. In our case, a certain position counts as origin or destination, if the duration of the stay within that zone exceeded a given threshold. In our setting, we chose a value of 60 min for this threshold.

Map Matching. Cell towers are rarely located near streets, but are mounted on buildings or elevated sites to achieve a better coverage. Hence, cell phone trajectories reflect the movement of the user between the cell towers and are not aligned with the underlying road network. In order to accurately map the trajectories to the traffic zones of an OD matrix, they must be projected to the underlying road network. As opposed to GPS-based trajectories, this process is more challenging for cell phone trajectory due to the low spatial accuracy of cell phone events. There exist several map matching algorithms tailored towards this problem [15, 16]. The QZTool uses the approached presented by Schulze et al. [16].

Transportation Mode Detection. In transportation planning, the OD matrices are usually split into two sub-matrices: one for motorized individual transport, and one for public transportation. However, the transportation mode of a given cell phone trajectory is not known upfront, but needs to be determined before assigning it to one of the matrices. This task is known as Transportation Mode Detection. Several approaches have been introduced for detecting the transportation mode of cell phone trajectories [17, 18]. In the QZTool system, a rule-based classifier is used, which determines the mode based on feature scores. We distinguish the two transportation modes *private transportation* and *public transportation*. The first includes all individual traffic, e.g. cars or bikes. For the second mode, only train passengers were assigned to public transportation. This will be extended to further means of transportation (e.g. buses) in the near future. To distinguish between transportation modes, the following features are used.

- *Distance to next rail and road link*: Every trajectory is matched on two different network types, namely the road network and the rail network. The assumption is that if a certain transportation mean was used, the average distance to its underlying network is shorter than the distance to the other network. The two networks have been imported from OpenStreetMap, using the attributes "motorway, trunk, primary, secondary, motorway_link, trunk_link, primary_link, secondary_link" for the road network, and the attributes "rail" for the rail network. Additionally, the attribute "oneway" was used for the road network.
- *Overlap with public transportation timetable*: It is assumed that if a user takes the train, the cell phone events not only occur on or near the current spatial position of the train, but also correlate temporarily. This feature assigns a higher score if there is a strong overlap of the cell phone trajectory with a train connection, taking both the spatial position and the temporal correlation into account.

Assignment to Traffic Zones and Creation of OD Matrix. In this step, every consecutive pair of stationary locations within a trajectory is added to the corresponding entry in the matrix. For example, if the user visits location A, D and F, the value of the entries (A,D) and (D,F) is increased by one. Finally, after all trajectories have been processed, the OD matrix is created and stored as comma separated value file (CSV).

3.4 Frontend

Users can interact with the QZTool via a Web-based graphical user interface (Fig. 2). This interface is implemented using HTML and JavaScript and runs in every modern browser without the need of further technologies. Hence, the QZTool can be used easily and does not require the installation of additions software on the users' computers. The typical work flow interacting with the QZTool consists of the following steps: The user uploads a Shapefile that contains the traffic zones, selects the desired time range, and starts the calculation on the Hadoop cluster with a single click. Once finished, the OD matrix can be downloaded as a CSV file.

4 Evaluation

In order to evaluate the system, we selected an area in the east region of Austria (including Vienna) and split it into 82 traffic zones in total (Fig. 2). This area has a total population of more than 3 million people. We created two matrices using different techniques, which were then compared to each other: (i) a cell phone-based OD matrix, and (ii) a traffic demand matrix based on a workday transport model. As an analysed time period for cell phone data, we chose three weeks in June 2014 (June 9, 2014 to June 28, 2014) and each day was compared with the result of the workday traffic demand matrix.

The row sums of every traffic zone for both matrices are shown in Fig. 3. The first group of spikes refer to the four traffic zones in Vienna, the largest city in Austria. Naturally, large cities create a lot of traffic, either as origin or as destination. The second group of spikes refer to the traffic zones in the south east of Vienna, which are popular areas for commuters. Again, due to their nature, a lot of traffic is generated from those traffic zones.

Next, we compared both matrices, based on absolute and relative values, using a statistical test. Both matrices are checked for their statistically independence,

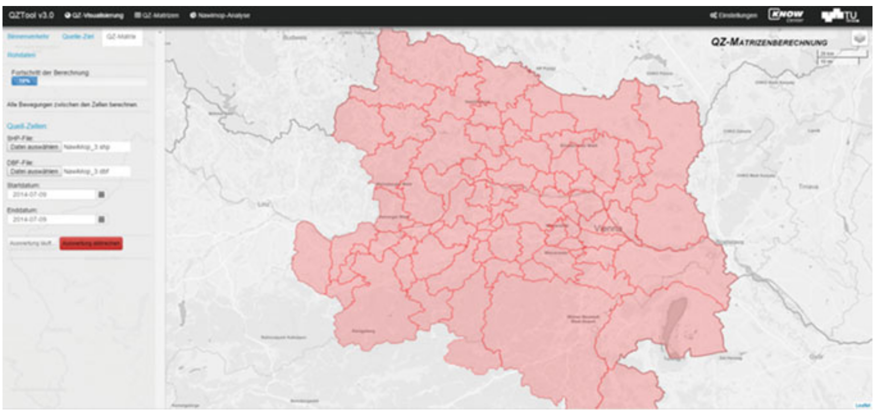


Fig. 2 Screenshot of the QZTool web-based graphical user interface

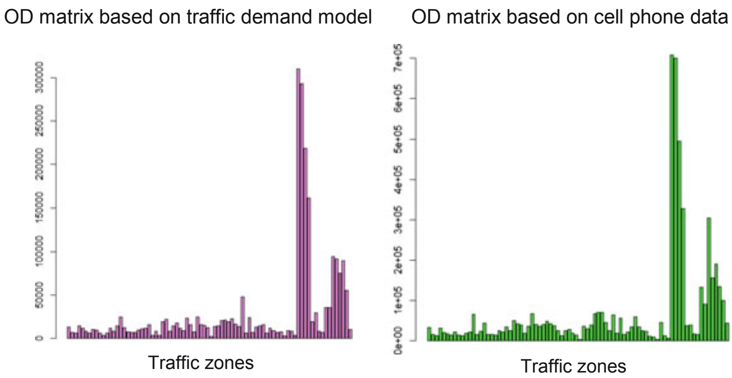


Fig. 3 Comparison of OD matrices, domestic traffic is not considered. Both matrices agree to a high degree

especially every origin and destination zone-vector, Therefore, a Pearson's chi-squared test for independence was applied for absolute and relative values. H_0 and H_1 were formulated as follows:

- H_0: The matrices (vectors) are stochastically independent from each other
- H_1: The matrices (vectors) are stochastically dependent.

The test statistics for each origin and destination zone-vector shows a p-value of <0.05, therefor H_0 can be rejected. It is shown that both OD matrices, based on traffic model and cell phone data, are stochastically dependent from each other.

5 Conclusion

In this paper, a system for generating OD matrices based on Floating Phone Data is presented. We showed that FPD can be a valuable alternative to create, validate or amend traffic demand models.

There are several challenges that need to be addressed when using this data source. First, the cell phone trajectories need to be separated into private (motorized) and public transport. Detecting the transportation mode of a given trajectory is a non-trivial task. The second challenge is the projection of the trajectory to the underlying road and rail network. And last but not least, the huge amount of cell phone data requires Big Data processing techniques in order to ensure a good user experience.

For future research, we plan to further improve the algorithms for transportation mode detection. Besides that, creating real-time OD matrices is one of the challenges we'd like to address.

Acknowledgments The Know-Center is funded within the Austrian COMET Program—Competence Centers for Excellent Technologies—under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

References

1. Ben-Akiva, M., Lerman, S.: Discrete Choice Analysis. MIT Press, Cambridge (1989)
2. Bera, S., Rao, K.V.K.: Estimation of origin-destination matrix from traffic counts: the state of the art. Eur. Transp./Trasporti Europei. **49**, 2–23 (2011)
3. Caceres, N., Wideberg, J.P., Benitez, F.G.: Review of traffic data estimations extracted from cellular networks. IET Intel. Transport Syst. **2**(3), 179–192 (2008)
4. Steenbruggen, J., Borzacchiello, M.T., Nijkamp, P., Scholten, H.: Mobile phone data from GSM networks for traffic parameter and urban spatial pattern assessment: a review of applications and opportunities. GeoJournal **78**(2), 223–243 (2011)
5. Aydos, C., Hengst, B., Uther, W.: Kalman Filter Process Models for Urban Vehicle Tracking. In: 12th International IEEE Conference on Intelligent Transportation Systems (ITSC '09), pp. 1–8. IEEE Press, New York (2009)
6. Caceres, N., Romero, L.M., Benitez, F.G., del Castillo, J.M.: Traffic flow estimation models using cellular phone data. IEEE Trans. Intell. Transp. Syst. **13**(3), 1430–1441 (2012)
7. Calabrese, F., Colonna, M., Lovisolo, P., Parata, D., Ratti, C.: Real-time urban monitoring using cell phones: a case study in Rome. IEEE Trans. Intell. Transp. Syst. **12**(1), 141–151 (2011)
8. Janecek, A., Hummel, K.A., Valerio, D., Ricciato, F., Hlavacs, H.: Cellular data meet vehicular traffic theory: location area updates and cell transitions for travel time estimation. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pp. 361–370. ACM, New York (2012)
9. Caceres, N., Wideberg, J.P., Benitez, F.G.: Deriving origin destination data from a mobile phone network. IET Intel. Transport Syst. **1**(1), 15–26 (2007)
10. Calabrese, F., Di Lorenzo, G., Liu, L., Ratti, C.: Estimating origin-destination flows using mobile phone location data. IEEE Pervasive Comput. **10**(4), 36–44 (2011)
11. Li, B.: Bayesian inference for origin-destination matrices of transport networks using the EM algorithm. Technometrics **47**(4), 399–408 (2005)
12. Toledo, T., Kolechkina, T.: Estimation of dynamic origin-destination matrices using linear assignment matrix approximations. IEEE Trans. Intell. Transp. Syst. **14**(2), 618–626 (2013)
13. Iqbal, M.S., Choudhury, C.F., Wang, P., González, M.C.: Development of origin-destination matrices using mobile phone call data. Transp. Res. Part C: Emerg. Technol. **40**, 63–74 (2014)
14. Horn, C., Klampfl, S., Cik., M., Reiter, T.: Detecting outliers in cell phone data: correcting trajectories to improve traffic modelling. Transp Res Rec.: J. Transp. Res. Board (2014)
15. Lou, Y., Zhang, C., Zheng, Y., Xie, X., Wang, W., Huang, Y.: Map-matching for low-sampling-rate GPS trajectories. In: Proceedings of the 17th International Conference on Advances in Geographic Information Systems (SIGSPATIAL), pp. 352–361. ACM, New York (2009)
16. Schulze, G., Horn, C., Kern, R.: Map-matching cell phone trajectories of low spatial and temporal accuracy. In: IEEE 18th International Conference on Intelligent Transportation Systems, pp. 2707–2714. IEEE Press, New York (2015)

17. Stenneth, L., Wolfson, O., Yu, P.S., Xu B.: Transportation mode detection using mobile phones and GIS information. In: Proceedings of the 19th ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL), pp. 54–63. ACM, New York (2011)
18. Wang, H., Calabrese, F., Di Lorenzo, G., Ratti C.: Transportation Mode Inference from Anonymized and Aggregated Mobile Phone Call Detail Records. In: 13th International IEEE Conference on Intelligent Transportation Systems, pp. 318–323. IEEE Press, New York (2010)

Part X
Aviation—Human Factors in Aviation

Follow-the-Greens: The Controllers' Point of View Results from a SESAR Real Time Simulation with Controllers

Karsten Straube, Marcus Roßbach, Björn D. Vieten and Kerstin Hahn

Abstract Although pilots are often supported by signage, markings and lighting when taxiing on the airport surface, navigation and monitoring remain workload intense tasks even in good weather conditions. Radio communication is near capacity limits on many airports today resulting in waiting times and delay. Apart from well-known safety issues this also constitutes a negative impact on the environment. The European aviation research program SESAR addressed this problem and came up with a solution: A new surface traffic management concept proposes the automated use of Airfield Ground Lighting with individually switched green taxiway centerline lights indicating the path to be followed. This paper presents official validation results indicating that SESAR developed a safer, quicker and greener surface traffic management concept.

Keywords Human factors · Airfield ground lighting · Follow-the-Greens · Safety · Air traffic control · Airport operations

K. Straube (✉)
Deutsches Zentrum für Luft- und Raumfahrt e.V., Lilienthalplatz 7,
38108 Brunswick, Germany
e-mail: Karsten.Straube@dlr.de

M. Roßbach
Munich Airport, Nordallee 25, 85356 München-Flughafen, Germany
e-mail: Marcus.Rossbach@munich-airport.de

B.D. Vieten · K. Hahn
FRAPORT AG, 60547 Frankfurt, Germany
e-mail: B.Vieten@fraport.de

K. Hahn
e-mail: K.Hahn2@fraport.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_69

1 Introduction

On 22nd December 2013, a British Airways Boeing 747 from Johannesburg to London Heathrow was taxiing for departure and had been cleared to taxi to holding point runway 03L via taxiway Bravo but missed the turn towards the holding point runway 03L and continued on the smaller general aviation taxiway Mike until the right hand wing collided with and sliced through the walls of an office building. [1].

This example reveals that taxiing at major airports with a complex aprons and taxiway system constitutes a workload intense navigation and monitoring task for the flight crew which can ultimately result in safety critical situations. In order to mitigate these risks, pilots should be supported by concepts and procedures that increase their situational awareness. One of these concepts and procedures is the Airfield Ground Lighting (AGL) service with the Follow-the-Greens (FtG) procedure. The first step in validating the Airfield Ground Lighting (AGL) service as a cornerstone of a holistic Advanced Surface Movement Guidance and Control System (A-SMGCS) was made between 2007 and 2010 in the German Aviation Research Programme IV. A report in German language on the scope and the outcome of the project can be retrieved from [2]. After transferring parts of the concept to the European SESAR Programme from 2010 on, a first large-scale validation exercise on the FtG concept with the focus on the flight crews' point of view was performed on the Frankfurt Airport's SASIM (SESAR AGL Simulator) platform in 2013. The exercise was very successful and results can be found in [3].

This paper focuses on a second validation trial performed in 2015 on SASIM-2, the successor of SASIM, with a clear emphasis on the controllers' point of view, also performed in the premises of Fraport AG in Frankfurt Key objective of the second validation was to find out, whether partial or full implementation of FtG change controller workload, efficiency, safety and airport performance. Furthermore, the validation was conducted to prove that FtG allows for acceptable working conditions for controllers despite increasing traffic loads.

The FtG concept used for the validation exercise were developed in the context of the SESAR Programme and co-financed by the SESAR Joint Undertaking (SJU). The responsibility for this paper lies exclusively with the authors. The SJU and its founding members are not responsible for any use that may be made of the information contained herein.

2 Follow-the-Greens: The Concept

FtG is an innovative, self-explanatory and performance enhancing guidance method for aircraft and ground vehicles operating on airport taxiways. The objective of FtG is to make the process in all more efficient and more eco-friendly by improving the traffic flow on the aerodrome surface and consequently by reducing taxi times and fuel burn.

FtG provides visual navigation support to flight crews and vehicle drivers along the cleared route by activating a defined stretch of taxiway center line lights in front of the aircraft. All center line lights not needed for the visualization of a cleared path are deactivated. By moving forward, the aircraft literally pushes the lit segment frontward, while all lights below and behind the aircraft are switched-off again. Furthermore, the logic within the fully automatic system safeguards longitudinal separations and wingtip clearances in all kinds of converging traffic situations and assists the controller by acting as a holistic safety net.

The operational procedures related to FtG are a simplification of the current way of working, especially when it comes to radio communication: After the initial instruction to "Follow-the-Greens" followed by and information on the destination of the movement, e.g. "Lufthansa 123, Follow-the-Greens to Stand A28", all further instructions will be provided via AGL and without an accompanying voice instruction.

The regulatory basis for FtG was already laid in ICAO Doc 9830 [4]. The authors of the document state that visual aid instructions can be provided by using taxiway center line lights. In addition, the document clearly defines the general meaning of different colors of AGL: Green lights in front of the mobile represent the instruction to follow. The absence of activated green lights; or the presence of activated red lights; indicate the instruction to stop the mobile. Yellow or flashing lights mean caution.

The first local implementations of FtG, in those days manually or semi-automatically switched, date back to even before ICAO Doc 9830 was written. The pioneers were airports such as Munich Airport and London-Heathrow that can look back at histories of more than two decades or more than ten years of using FtG respectively. The system comprises an integrated Controller Working Position (CWP) and a fully-fledged Surface Management System (SMAN) covering several SESAR solutions related to Routing, Planning and Guidance, and Safety Nets.

Technically speaking, AGL can provide individual guidance information to any cooperative (transponder equipped) or non-cooperative target, e.g. an aircraft, a vehicle, or a combination thereof. While many other guidance services are partly or fully depending on on-board installations, guidance via AGL is a purely ground based service. If the AGL service is implemented at an airport, it will need to be available 24/7 and under all operating conditions, including traffic peaks and low visibility conditions. Automated guidance via AGL has the following features at the current stage in SESAR:

1. Segments of lights: Taxiway centerline lights are switched in segments of lights each containing a minimum of two and a maximum of six lights. Each segment can only be entirely on or entirely off. These segments should be as short as possible close to intersections, turns, or slopes in order to make the visual indication for the flight crew as clear and precise as possible. All segments activated for one individual mobile at a defined point in time are called the individual route indication. The individual route indication comprises a locally configurable number of taxiway centerline lights in the field.

2. Single Lamp Control: Each individual lamp at the taxiway centerline can be controlled by the ground service. This feature is intended to deliver a smoother flow of the activated route indication in the field of vision of the flight crew and the vehicle driver. Furthermore, Single Lamp Control allows for a more efficient sequencing of ground movements.

Each ground controller monitors the traffic in his or her area of responsibility via the CWP Human Machine Interface (CWP HMI). Ground controllers are responsible for assuring that all mobiles comply with the guidance provided from the controller via the AGL. This means that either all guidance decisions are taken by the system, e.g. who crosses an intersection first, or all decisions taken by the controller have to be entered into the CWP HMI.

The SMAN automatically translates all guidance instructions into individual switching commands for the AGL. Ground Controllers are always in the position to modify previous decisions via the CWP HMI whenever necessary. In case of non-compliance by the flight crew, specific procedures will be established such as revert to radio telephony (R/T) instructions. In general, R/T will remain available at any time and e.g. as a backup in case of AGL malfunction or for the clarification of the intent behind an instruction, etc.

2.1 Follow-the-Greens: Expected Results

The purpose of this specific real time simulation within SESAR was to gain reliable figures on the operational feasibility and benefits of the new system. Consequently, the operational feasibility had to be validated according to controllers' and pilots' feedback in terms of a complex airport environment as a first step. Then, secondly, operational benefits in terms of increased safety, increased situational awareness, decreased workload, increased capacity and reduced fuel burn and CO_2 emissions were in the focus of the exercise [5]. The expected results per Key Performance Area were:

Safety: The visual guidance via AGL will lead to a reduction of route deviations and holding position overruns during taxi procedures [5].

Human Performance: FtG will increase the controllers' and flight crews' situational awareness in Low Visibility Conditions operations and to reduce their workload while using FtG [5].

3 Method

The SESAR European Airports Consortium "SEAC" validated the new concept under the leadership of Munich Airport GmbH with its partners Fraport AG, DFS (with DLR acting as sub-contractor), Flugsimulator Frankfurt and ATRiCS between the 20th and 24th April 2015. The validation was executed on the basis of the layout and traffic characteristics of Munich Airport. Scenarios included runs in good weather as well as low visibility conditions, different traffic loads and remote apron control, i.e. the control of aerodrome surface areas without a direct line of sight.

3.1 Participants

Seven controllers (one female, six males, average age = 34.4 years; *SD* = 10.6 years) from Munich Airport and five pilots (all males; average age = 40.6 years; *SD* = 5.4 years) from different countries and airlines participated in the exercise. While all controllers were highly experienced in Munich Airport's apron control, the majority of pilots had no experience with the validated airport layout. The selected Munich Airport Controllers who participated in this validation had several training sessions on the ATRiCS CWP prior to the execution in order to get them familiarized with the system. Each controller received about eight hours of training some weeks before the exercise and one hour of refresh immediately before the exercise.

3.2 Test Environment

The validation was performed by means of real-time simulations on SASIM-2 at Frankfurt Airport with configurable traffic and different weather conditions (good visibility/low visibility) available. The validation platform was configured to simulate Apron 2 and Apron 3 of Munich Airport and was linked to an Airbus A320 cockpit simulator. SASIM-2 comprises four highly integrated CWPs from ATRiCS, acting as the map-based HMI for a Surface Management System All inputs to the system were done by means of multi-finger-touch. The view out of the window onto the aprons was simulated with professional 3D software from ATRiCS. The traffic simulation software was also provided by ATRiCS. The cockpit simulator was provided by Flugsimulator Frankfurt with an external 3D view and a very realistic look and feel. Digital voice communication systems provided the radio telecommunication channels and data storage applications saved all relevant information for future analysis including the voice communication.

3.3 Scenarios

The term scenario means a cased-based, timely limited and contextually defined experiment during the validation exercise with a specific technical/operational feature. Each scenario consisted of two runs—one for each controller group. A run is a combination of missions during a pre-defined timeframe. A mission describes a single movement conducted by one flight crew during a run.

Within this validation, ten different validation scenarios were performed by two different validations groups of different controllers and pilots. In the validation week, 20 runs were performed in total. In contrast, two individuals acted as controller assistants for the entire validation week. The validation consisted of ten different scenarios, two of them designed as reference scenarios and eight as solution scenarios. The flight plan used in this validation was based on a stored traffic sample reflecting 40 min of traffic on an average day in Munich airport. This peak consisted of 46 movements (arrivals and departures) and was used for all validation scenarios except for Solution 7 and Solution 8. In these two scenarios an off-peak traffic sample was used in order to represent early morning operations at Munich airport. In all scenarios, two movements were performed by the Airspace Users sitting in the cockpit simulator. All other movements were performed by pseudo pilots.

In both reference scenarios the runs were performed by two controllers per apron. One of them was the controller and the other one was the controller assistant. The first reference scenario presented the pre-SESAR technologies and procedures according to A-SMGCS level 2. In this reference scenario the guidance of the controller was provided by the guidance means which are already in use today. That means the FtG guidance concept was not applied and guidance was limited to guidance instructions via radio telephony (R/T). The second reference scenario was a specifically designed scenario representing today's guidance procedures at Munich Airport under low visibility conditions (LVC). The difference to the first reference scenario was that Munich Apron Control is already today operating with a semi-automatic AGL guidance system in LVC. Within this scenario the AGL was activated/deactivated in segments. Depending on the length of a segment a variable number of lights (3–16) could be switch on or off. Only one aircraft per segment was allowed. Consequently, the separation between two aircraft increased compared to situations under good visibility conditions (CAVOK) without FtG operations.

Eight solution scenarios were developed in order to allow for evaluation whether the SESAR FtG concept, including new ground controller equipment and AGL guidance, are beneficial compared to the systems and procedures used today. For most of the solution scenarios the staff situation remained identical compared to the reference scenarios with one controller and one assistant per apron. Nevertheless, in some specific scenarios the controller was the only person responsible for an entire apron and in two solution scenarios (7 and 8), Aprons 2 and 3 were merged into one area of responsibility. Table 1 shows all the major scenario characteristics as they were finally used for this validation.

Table 1 Scenario characteristics

	Pre-SESAR World (reference)		Follow-the-Greens World (solution)							
	Refernece VMC	Reference LVC	Solution 1	Solution 2	Solution 3	Solution 4	Solution 5	Solution 6	Solution 7 apron merge 1	Solution 8 apron merge 2
Number of controllers per apron	2	2	2	2	2	1	1	1	1	1
Weather	VMC	LVC	VMC	LVC	LVC	VMC	LVC	LVC	VMC	LVC
Follow-the-Greens navigation support	No	Today's MJC ops	Single lamp switching	Smallest available TCL unit	Single lamp switching	Single lamp switching	Smallest available TCL unit	Single lamp switching	Single lamp switching	Single lamp switching
Spacing concept	No	Manual block	No	Auto Block	Floating	No	Auto Block	Floating	No	Floating

3.4 Measurements

In order to obtain the metrics and indicators, four types of measurements were applied during the validation exercise.

Observations were performed during the simulation session and were supported by an observation grid built on the basis of expected controller and pilot behavior throughout the scenarios and on success criteria related to the objectives.

Different *tailor-made questionnaires* were used throughout the validation to gain the necessary feedback from the participants regarding validation objectives and personal experiences. Before the start of the validation, the flight crew, the apron controllers and the assistants completed a demographic questionnaire.

In order to assess possible operational improvements concerning Human Factors, mainly in terms of Situational Awareness and Workload, the apron controllers and assistants were presented an Inter Run Questionnaire consisting of the standardized NASA TLX and 3D SART approaches. In addition, all validation observers filled in an observer sheet after each run. After each cockpit mission a NASA TLX and a 3D SART were presented to the flight crew.

A *data logging* function focused on several aspects like average and maximum taxi times and taxi speeds and the number of stops and re-starts during taxi of all aircraft movements in the validated airport area. After the completion of the validation week, all recorded data was analyzed with the TRAZER analysis tool provided by ATRiCS. This software tool is able to visualize and analyze surveillance data of airport ground movements.

Debriefings were fulfilled with the controller and flight crew after a specific number of runs and at the end of the entire validation. The debriefing gave controller and pilots the opportunity to jointly discuss any issues which may have surfaced during any of the scenarios.

4 Results

4.1 Operational Improvements in Terms of Safety

The following Fig. 1 taken from the Validation Report [6] illustrates the controllers' perceived safety. In all, each and every controller and pilot indicated in several debriefings that they always fully trusted the FtG system and procedures. Furthermore, neither controller nor flight crew identified any safety critical issues at any time. The rating scale for the following figures is a Likert Scale ranging from one (unlikely/disagree) to 7 (likely/fully agree).

Further safety relevant results are that the total number of ground movement stops as well as the aggregated duration of ground movement stops is reduced by providing guidance with FtG. By using FtG procedures the total amount of communication can be reduced by more than 20 % compared to the pre-SESAR

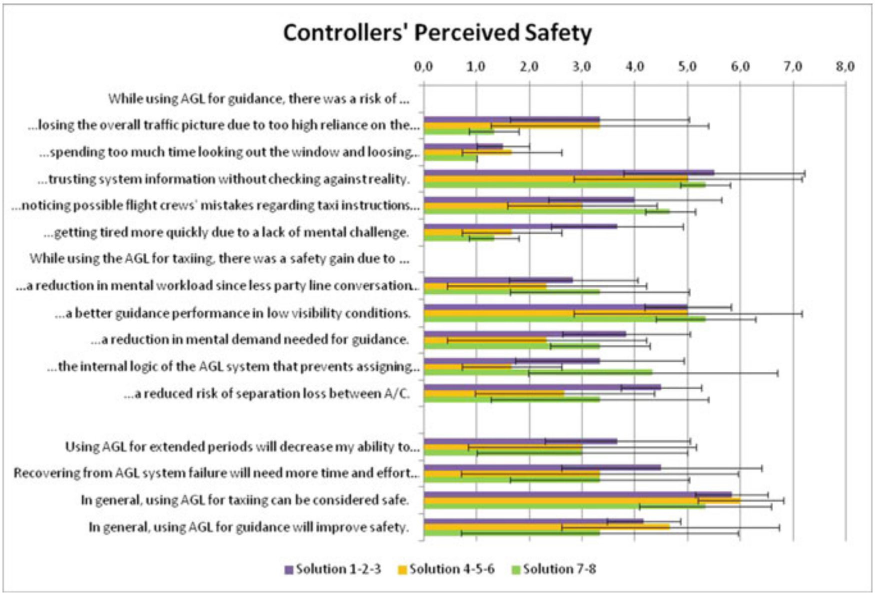


Fig. 1 Controllers' perceived safety

scenarios. Consequently the controller has more time to focus on other safety relevant aspects. When operating FtG, no route deviation has been recorded in all validation scenarios [6].

4.2 *Operational Improvements in Terms of Human Performance*

Due to the reduced amount of R/T communication in FtG operations, the controllers' R/T frequency is less often occupied. Consequently, the controller might be in a better position to give a clearance earlier compared to operations without FtG. Controllers and flight crews had the impression that their mental, physical and temporal demand as well as their effort and frustration level was reduced when using FtG [6].

Situational Awareness. The controllers and flight crews were asked about their perceived situational awareness in the different scenarios by the use of a standard 3D SART. The rating scale for the 3D SART was as follows: Demand on Attentional Resources (−3 minimally demanding/+3 excessively demanding), Supply of Attentional Resources (−3 greatest possible effort/+3 excessively demanding), Understanding of Situation (−3 virtually nothing/+3 almost everything). The following Table 2 from the validation report [6] shows that according

Table 2 Controllers' and flight crews' perceived situational awareness

	Controller		Flight crew	
	PRE-SESAR	POST SESAR	PRE-SESAR	POST SESAR
Demand on attentional resources	0.2	−0.3	−0.4	−0.8
Supply of attentional resources	0.6	0.0	0.4	−0.3
Understanding of situation	2.6	2.8	1.9	2.6

the analysis of the 3D SART the situational awareness of the controllers' and flight crews' is improved compared to the reference scenario values.

Perceived Workload. The controllers and flight crews were asked about their perceived workload in the different scenarios by the use of a standard NASA TLX questionnaire. The following Figs. 2 and 3 illustrate the answers by the controllers and flight crews about their perceived workload when using FtG guidance.

It is obvious that controllers and flight crews had the impression that their mental, physical, temporal demand and their effort and frustration level were reduced by using FtG. However, according to the answers their performance highly declined. This unexpected behavior has been analyzed in a post-validation study. It is now assumed that the rating scale caused this unexpected performance result. The rating scale evaluating the performance category was the only one reversed in the questionnaires compared to the other categories. In some post-validation debriefings with the controllers who participated in the validation they confirmed that they rather felt an improvement of their performance when using FtG instead of the documented declination. Hence, it is highly probable that the unexpected result was caused by misunderstanding.

Usability. Figure 4 gives an impression of the controllers' rating of the usability of FtG. The feedback during the validation of all controllers and all pilots was that

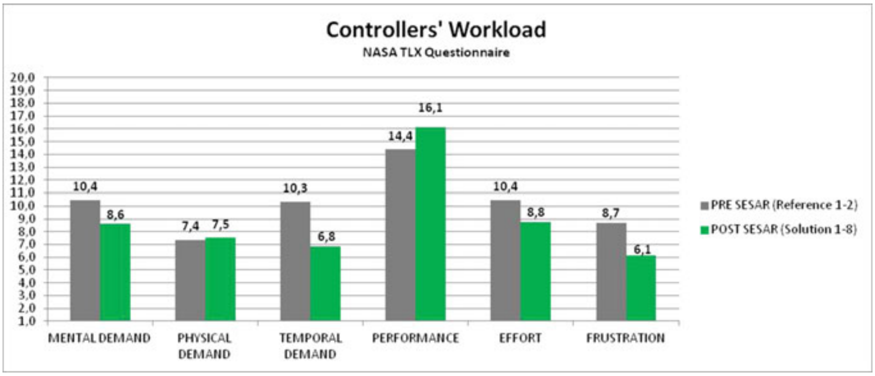


Fig. 2 Controllers' workload

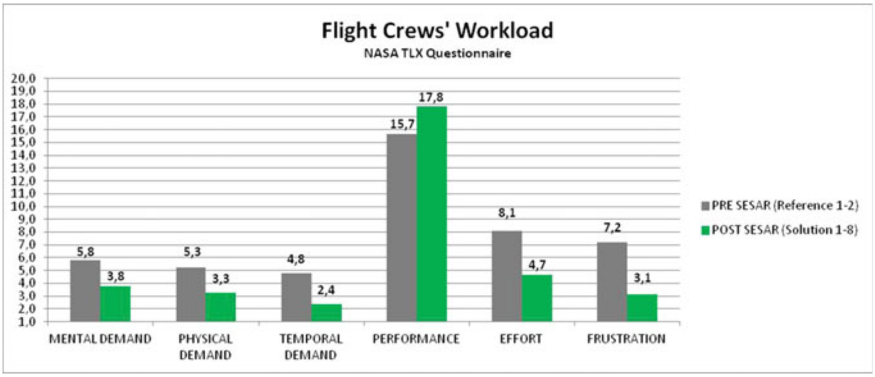


Fig. 3 Flight crews' workload

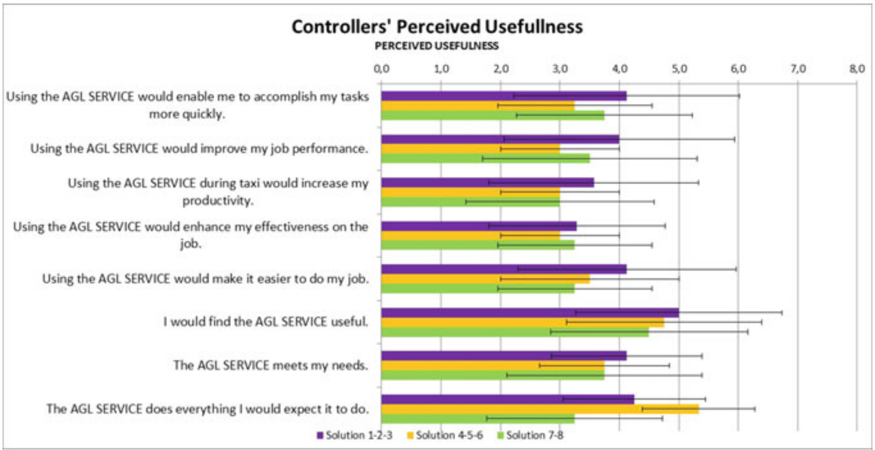


Fig. 4 Controllers' perceived usefulness

they appreciated to work with the FtG system and that it would improve their daily operations. The rating scale for Fig. 4 is a Likert Scale from one (unlikely/disagree) to 7 (likely/agree).

The observations made by a special observer team were also very positive. They revealed that the controller made almost no use of the direct line of sight onto the aerodrome surface and that they were focused on the HMI of the system. A very interesting result is that due to the reduced workload with the FtG concept the controller had the feeling that there was less traffic in the scenarios compared to the reference scenarios. The validation team then clarified that the traffic sample was identical for all scenarios, only the call signs and some types of aircraft were changed in order to avoid learning effects.

A further observation regarding the pilots was that the FtG concept reduced the pilots' head-down times. Detailed information can be found in [6].

5 Discussion/Conclusions

The results of the real-time simulation allow for declaring FtG as successfully validated on a large, complex and congested hub airport. In all, FtG is a clear improvement compared to today's procedures from the controllers' and the flight crew's point of view. It was demonstrated that FtG qualifies for the guidance of future traffic on complex airports with very high traffic loads and also for very bad visibility conditions during fog and darkness. The simulation proved operational feasibility and demonstrated the scale of the operational improvements inherent in the concept. In terms of safety, responses as given by the controllers and pilots clearly indicate that FtG will lead to safer airport operations. These results are supported by increased Situational Awareness and reduced Workload of controllers and flight crews. The validation showed that the merge of aprons 2 and 3 during off-peak situations would be possible.

5.1 Recommendations

The validation showed that the FtG concept and its system environment have successfully reached V3 maturity and that this solution is ready for industrial applications. Furthermore, a significant number of airports are highly interested in implementing the concept. On a European and global scale, an adequate standardization covering all aspects from technical requirements and parameters to procedural standards up to issues like the phraseology is still missing. The lack of standardization is currently seen as a major roadblock in Europe and an advantage in competition for other regions of the world. Therefore, it is recommended to put additional pressure on the initiatives aiming at developing of standards for "Follow-the-Greens" such as the EUROCONTROL A-SMGCS Task Force.

5.2 Outlook

Further research activities should focus on special operational procedures, e.g. in case a part of a taxiway (temporarily) cannot be equipped with centerline lights. In this case the wording "Follow-the-Greens to xyz" seems inappropriate. Furthermore, future research should clarify the latest possible point in time for a route change in front of a mobile. In addition, the visualization of warnings and alerts via AGL should be elaborated as well as instructions to not pass traffic on parallel taxiways in case of wingspan restrictions.

A further topic for future validations should be the question on how to hand over the traffic during shift changes. Nowadays, the decisions and guidance instructions given by the previous controllers can be reproduced by the use of the pre-SESAR technologies. Using the FtG procedures, it might be impossible to reproduce the instructions given by the previous controllers.

Acknowledgments © SESAR JOINT UNDERTAKING, 2016. Created by SEAC, DFS, and ENAV for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL.

References

1. Accident: British Airways B744 at Johannesburg on December 22nd 2013, took wrong taxiway and buried wing in building. http://avherald.com/h?article=46d6e18c&opt=0
2. Verbundvorhaben "Wettbewerbsfähiger Flughafen", Forschungsthemencluster 2: Rollverkehr: Management- und Optimierungssysteme": Schlussbericht; Förderprogramm: Luftfahrtforschungs programm IV (LuFo IV) (2010). https://www.tib.eu/en/search/download/?tx_tibsearch_search[docid]=TIBKAT%3A647436647&cHash=404d169e3ee5c332cecb087c4ab89b22#download-mark
3. SESAR EXE-06.07.03-VP-649 D62: Validation Report
4. ICAO Doc 9830 AN/452: Advanced Surface Movement Guidance and Control Systems (A-SMGCS) Manual, 1st edn. (2004)
5. SESAR EXE-06.03.01-VP-759 D152: Validation Plan
6. SESAR EXE-06.03.01-VP-759 D153: Validation Report

Purple Sky Framework Towards the Flight Deck of the Future Experience: Through Co-design, Rapid UX Prototyping, and User Testing

So Young Kim, Jennifer Cooper, Alexander Carroll and Sundar Murugappan

Abstract Envisioning and conceptualizing the future experience in flight deck present particular challenges. It is constrained by legacy system architecture, platform, and capabilities. Therefore, it typically requires significant investment upfront in terms of resource and time. Also, due to its complex nature, the flight deck systems are rather a combination of different subsystems than an holistic integration of subsystems [1]. A design framework, Purple Sky, is established to consider these challenges. The framework includes a design process based on Human-Centered Design and User Experience (UX) Rapid Prototyping Platform. In this paper, we will introduce the framework using a case study of designing future flight deck concepts for business jet operations.

Keywords Flight deck · Human-centered design · Rapid prototyping

1 Introduction

Flight deck systems are among the most complex and regulated systems. The complexity of the current systems and strict regulations on the design of the current systems hindered revolutionary updates on the flight deck systems. In addition, the

S.Y. Kim (✉) · A. Carroll · S. Murugappan
General Electric Global Research, 2623 Camino Ramon, San Ramon, CA 94583, USA
e-mail: sykim@ge.com

A. Carroll
e-mail: alex.carroll@ge.com

S. Murugappan
e-mail: murugappan@ge.com

J. Cooper
General Electric Digital, 2623 Camino Ramon, San Ramon, CA 94583, USA
e-mail: jennifer.cooper@ge.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_70

effort of prototyping in flight deck systems is typically tied to a product platform. Therefore, it takes long time to prototype a new idea, and exploring multiple ideas become prohibitively costly.

With advancement in technologies, the modern flight deck systems have been improved significantly. However, these improvements are from additions of technologies in an ad hoc manner [2], rather than holistic integration.

1.1 Challenge 1: There Are No Prototyping Platform and Process that Are Flexible and Rapid with Low Cost

Prototyping and testing new and novel ideas on flight deck systems for new concepts of operations has not been properly incorporated in the design process. Existing prototyping and testing processes are often time-consuming and inevitably expensive similar to developing the flight deck systems that can be installed and operational in aircraft.

1.2 Challenge 2: Current Flight Deck Systems Are Unnecessarily Complex

Current flight deck systems as an integrated system for pilots are unnecessarily complex due to its tendency of incremental updates and additions of new system components. Each component is developed somewhat independently of the overall operational context of the whole system, resulting in systems for engineers, not for the user of the flight deck systems, pilots.

1.3 Case Study

We will use an example project as a case study throughout this paper. The goal of the project is to design a flight deck experience for future business jet operations.

1.4 Purple Sky Design Process

The Human-Centered Design approach was applied to the design process. It comprises of Co-Design Workshop, Human Factors and User Experience (UX) Research and Design, Prototyping, and finally, User Testing. When there is a new design idea, technology concept, user challenges, or business problems, one can result in value-tested, refined proof of concepts (POCs) through this process. First, preparatory research and co-design workshop are conducted. Especially, the co-design workshop brings users and stakeholders together in problem definition and/or design ideation process. Different perspectives are combined in a structured way using techniques such as affinity diagram, storyboarding, and journey map, etc. Based on the outcomes from the research and the workshops, Human Factors and UX research and design are conducted. Techniques such as cognitive task analysis, workflow analysis are used. As designs get done, prototypes also get implemented. The fidelity of prototypes is ranged from paper prototypes and simulated prototypes. Some prototypes could also be just scripted instead of simulated to show the newly-design workflow. Once a prototype is ready, regardless of its fidelity, design/concept validation is conducted. The format of the validation depends on the fidelity of the prototype. For example, if a paper prototype is prepared, a concept walk-through can be used to receive qualitative feedback on overall concepts and on potential values to users and stakeholders.

One should note that these four activities are highly iterative and also executed in an overlapped fashion although illustrated as a sequential process.

1.5 Rapid Prototyping Platform

The rapid prototyping platform for Purple Sky was designed as an integral part of the design process. Because the flight deck is an inherently complex environment with many safety critical interactions, we needed the ability to quickly evaluate interaction design concepts in an environment that will be sufficiently engaging for users.

We selected our tool chain with the goal of minimizing the rework between each level of fidelity, enabling designs to progress iteratively as far as was required to demonstrate and evaluate the concepts being explored. Design files were directly imported into the software development environment, eliminating the need for software engineers to spend significant time on layout activities, and helping to ensure strong agreement between the visual design and interactive prototype. The engineers then added interactivity using industry standard patterns and practices, such as the Model-View-View Model (MVVM) design pattern, and a message driven architecture. We included the ability to connect our designs to existing, commercial-off-the-shelf (COTS) simulation engines, enabling us to provide higher levels of fidelity in our prototypes as well as to script scenarios at higher levels of abstraction and reuse. Because we are leveraging a COTS tool chain, we are able to quickly integrate and explore new interaction technologies and capabilities such as speech, gesture, or touch, in addition to proven methods such as cursor control devices (CCD) and keyboards.

2 Purple Sky Framework

2.1 Co-design Workshop

Co-design (also referred as Co-creation) is a process originated from Human-Centered Design approach that "brings diverse stakeholders together to achieve breakthroughs in how they solve problems" [3]. It is a widely used and studied process [4, 5]. The co-design workshop is a setup that compresses this process into multiple-day-long activities. This setup has been used to brainstorm user challenges or business problems and to ideate potential solutions with stakeholders. Particularly, in Purple Sky framework, it has been very successful in gathering the cross-functional team to work closely by defining problems and building solutions together and also engaging and empowering customers and users to be involved in the design process.[1]

A co-design workshop can be formulated based on the "Game Design" method which uses the divergent-emergent-convergent strategy [6]. This strategy provides a structured brainstorming format so that the desired outcomes are achievable within a timeline. First, an initial state (e.g., initial problem statement, user challenges, or an domain of interests) and a target state (e.g., a goal of the workshop) are defined. From that initial state, a workshop starts opening existing ideas and information to the participants. Also, most importantly, a scope of design space is defined. Then, the group explores further ideas, followed by converging the ideas towards the target state. Various activities that facilitate each step are available [6]. Common activities include affinity diagram, persona modeling, storyboarding, paper prototyping, etc. [7].

Case Study. The workshop structure for designing a future flight deck concept is listed in Table 1. As mentioned earlier, the workshop follows the process of Divergent-Emergent-Convergent. During Divergent, we focused on creating an overview of users' work in context to ensure the understanding of the participants on business jet operations from a pilot's perspective. During Emergent, we explored multiple scenarios in which users experience difficult and critical situations. During Convergent, we focused on ideation based on the scenarios we created in the previous activity and prototyped using pen and paper in low fidelity.

Divergent. User journey map is shown Fig. 1. This activity can be used to understand the context of use of information and systems inside the flight decks and

[1]Evidenced by the quotes from the participants. "*I was very skeptical about the process we followed—I thought we could have come up with the answer in the first hour of the first day…. however….after going through the process, I can see the value in it—and we came up with a different answer than I would have provided in the first hour of the first day. I appreciate the opportunity to collaborate with GE, and hope we can continue the relationship.*" "*Creative exercises were very successful at highlighting issues with current systems and consolidating ideas from the team.*"

Table 1 An example of a workshop design for business jet flight deck use case

Phase	Agenda	Outcome
Divergent	Understanding business jet pilot challenges in current flight decks	User journey map, Pattern of issues in operating/flying modern business jet
Emergent	Moments of tension in pilots' lives	Patterns of situations in which pilots are under extreme stress
Convergent	Storyboarding	Critical scenarios that describes issues and challenges in context

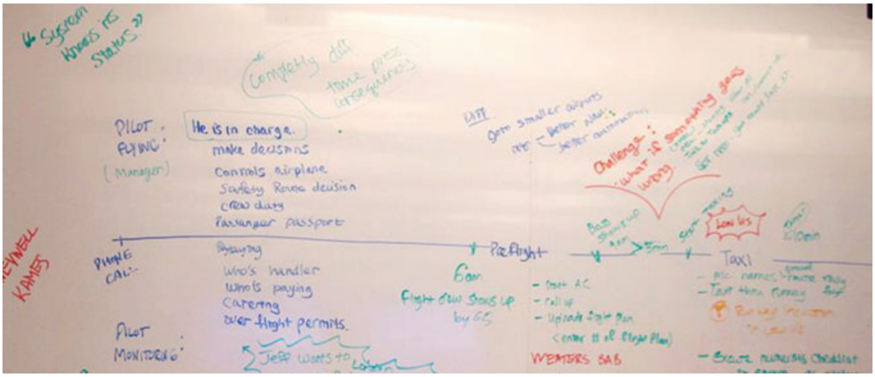


Fig. 1 An example of user journey map created with business jet pilots during the activity

to identify existing and potential issues and challenges of using current flight deck systems. Also, it can be used to identify the gaps and needs in operational practices.

While creating the journey map, discussions about pilots' aspiration and motivation are encouraged. Also, discussions of emergencies that could happen help illustrate "edge case" scenarios. The information captured in this exercise provides a high-level view of the users' work. Organizing the data captured during the activity based on emerging themes helps view the individual issues from a systems perspective. This view supports identifying fundamental issues that lead to a better solution.

Emergent. "Moments of tension" activity is designed for participants to focus on the experience that users go through without bringing technology into the discussion. That is, it allows us not to limit the discussion to existing technologies but to identify opportunities. Figure 2 shows how the activity can be conducted during a workshop.

An example of a narrative resulted from this activity could be following:

> *"Dark and stormy night over Newfoundland with a fully fueled airplane and in cruise flight, the path were secured the first course. The pilots, Boris and Tom were settling in altitude, checking weather and talking. All of a sudden they began to smell a light smell, somewhat toxic, like electrical wires and possibly insulation burning. They both looked at each other in concern. Boris told Tom to go find the source of the fire. […]"*

Fig. 2 During the moments of tension activity. A facilitator is teamed up with an engineer and a business jet pilot and articulates a number of one-paragraph narratives that describes moments of tensions in pilots' lives

These narratives point out the moments in context in which pilots run into emergency situations. With the descriptions of the users' emotional status and stress level, it emphasizes the urgency of the situations. Also, one can identify what supports the pilots would need and how these should be delivered. One can easily imagine that even the best technology will fail to support the users if it requires complex programming to enable the technology in this situation.

Convergent. Among the narratives created during the previous activity, through discussion and voting from participants based on its criticality and impact, we identified patterns of situations that the pilots considered to be problematic and challenging: (1) challenges during en-route and (2) challenges before takeoff. Based on that, storyboarding can be used to illustrate details of situations and, thus, opportunities. Storyboarding is a method to capture issues with technologies, operations, etc. from the users' perspective. Similar to the narratives, the storyboarding describes issues and challenges within context, enabling readers to empathize with users. Compared to the narratives, the storyboard allows capturing more details of the situation. Figure 3 shows the process of creating a storyboard and the resulting storyboard illustrated in a cartoon style. Storyboards can take a variety of different forms, however, having a visual is an effective way to convey the story.

2.2 Research and Design

To create the designs, we first needed to fully understand the scenario, the sequence of events, the data and information required, and action that needed to be performed. We used the storyline created by the users during the workshops, based on experiences of business jet pilots and the starting point and scope for our

Fig. 3 The workshop participants are developing a storyboard (*left*). The resulting storyboard is shown (*right*)

investigation. One of the identified problematic areas was a checklist interaction. The checklists that related to this particular scenario were analyzed, and the contents categorized into 4 parts:

- Key decisions to be made
- Actions to perform
- Warnings and Cautions
- Supplementary Information.

We also captured the external tasks to the checklist that need to be performed, such as communicating with ATC, cross-checks between pilots, responsibility handovers (pilot-not-flying and pilot-flying) and supporting tasks such as checking weather and rerouting the flight plan. The 'current state' of the full workflow was then used as a basis for design, and as a list of requirements with respect to what must be done to ensure the safety of the flight and successful completion of the tasks.

Design Principles. After data from the workshops and procedure research was collated and synthesized, we distilled the recommendations into 2 design principles.

Simpler Interaction. Provide a quicker way of dealing with the checklist. Several pain points to do with wasting time and effort during the checklist procedure were highlighted by the users, indicated by quotes taken for the workshop: "Don't make me search for the checklist" and "Don't make me search for the control switch", along with observed behavior during research sessions regarding checklist confusion and information overload.

Reduce Cognitive Workload. Alleviate the Pilots cognitive load by providing him with a correct and efficient procedure:

- Give advice on what to do and in what order, based on warnings, sensor information and pilot decisions
- Remove unnecessary and distracting steps
- Highlight the pertinent data at the right time.

Workflow Redesign. With these core principles in mind, the current checklist workflow was redesigned. Note that although the flow and tasks were reassigned, the intention of the checklists were maintained, and all information and steps required were still performed in order to adhere to requirements of the procedure for successful completion.

With the intention of simplifying the interaction, several advanced user technologies were introduced to flight deck environment: touch, voice, and gesture control. The interaction and interface design were then wireframed based on the features associated with these technologies. Throughout the wireframing process, a small number of pilots were consulted on a biweekly basis to review the designs and ensure that they remained intelligible, useful, and practical within the flight deck context.

As smart phones and tablets are now prevalent, with 68 % of Americans now owning a smart phone [6] and iPads becoming a common place object within the flight deck,[2] we decided to utilize interaction patterns designed and shared by some of the leading manufacturers such as Google and Apple. The intention for this was that, as users become familiar with the interaction methods for these devices within their day-to-day lives, this could be taken advantage of to simplify and streamline the "onboarding" process for understanding how to use the NUI technology within the flight deck.

Visual Design. In parallel to wireframing the workflow and interaction methods, the visual design system for the interfaces was developed. This began with investigatory research into the design requirements for flight decks as required by the FAA and design principles based on modern day consumer products, and progressed into full visual interface design as the wireframes grew to a stable state.

The resulting visual designs for each individual screen were captured in Adobe Illustrator files that could then be imported for the prototyping process. Also created was a flight deck "Style Guide," used to communicate common themes that stretched across the entire flight deck. Figure 4 shows the integrated visual concept.

2.3 Prototyping

The Purple Sky rapid prototyping platform is composed of both logical patterns and software components that come together to accelerate the prototyping process, enabling more design concepts to be explored. While we are presenting the implementation used in the case study, the same process, architecture, and software components have been used in other domains, sometimes with different tool chains.

Process. The prototyping process begins with a review of the design to be prototyped, and identification of what the prototype needs to convey. The team then

[2]American Airline, United Airlines and Alaska Air have all confirmed the use of iPads within their operational flight decks.

Fig. 4 Integrated visual concept for a flight deck of future

breaks down the designs by capability, identifying new interaction needs, system interactions, data dependencies, and open design areas. This decomposition enables the engineers to start in parallel with the design team, building out capabilities that will be wired up to the designs in the prototype. Once design assets become available, they are imported into the prototyping environment and the engineering team begins to add interactivity and connect the User Interface (UI) elements to the underlying system data. Periodic reviews help clarify design intent, and because the prototyping architecture limits dependencies between, design changes are incorporated in an ongoing basis.

Tool Chain. Commercial off the shelf tools (COTS) provide a plethora of readily available rich functionality for designers, and hence, we elected to leverage them as heavily as possible. Once designs were ready for higher fidelity development, our visual designers developed them in Adobe Illustrator (AI). When ready, the designs were directly imported into Microsoft Expression Blend to convert the AI vectors to XAML (Windows Presentation Foundation). Design work and cleanup of the import could continue in Blend, or be handed off to the development team. By logically separating the two activities, the designers and developers have the freedom to work in their familiar environments and provides for effective collaboration. The developers used Microsoft Visual Studio for most of their work, adding interactivity, and connecting the designs to the system model, simulated data, and workflow. We used COTS simulators to accelerate our work, using XPlane 10 for the aircraft model and out the window view, and Ternion Corporation's FLAMES for scenario scripting and specific component modeling. As with most things in our process, this was an iterative process, and our early out the window views were powered by Google Earth. We were also able to leverage third party Natural User Interface (NUI) devices and Software Development Kits (SDKs) to rapidly prototype and evaluate new interaction modalities such as touch, voice and gesture interactions.

Unifying Architecture. We adopted several modern software engineering architectural patterns that collectively helped reduce the dependencies between components in the prototypes, facilitated re-use of capabilities, and reduced the effort required for each iteration. Two of the most important are presented here.

We adopted the Model-View-View Model (MVVM) pattern for our UI development. The MVVM pattern enabled us to reduce the dependencies between the visual designs and the underlying logical implementation, which facilitated design changes, and enabled us to divide the work, with one developer working on the View and another developing the Model, connecting them with the View Model. It also made it quite easy to replace early, hard-coded models with dynamic models that were connected to simulator data without having to refactor the front end.

To reduce coupling between the components in our platform, we adopted a message-based architecture, which has enabled us to quickly swap system components, such as simulation engines, or NUI devices. In addition, it also has enabled us to explore multi-modal interfaces, and improved our ability to instrument our prototypes for Human In The Loop (HITL) evaluations. This architecture also has the virtue of mapping to the federated architecture of current cockpit designs.

2.4 User and Business Value Testing

With the interactive prototype, we conducted multiple sessions to test designs. The fidelity of the testing was low in that we walked through the concept with the users and had them experienced in first hand. While they were using the prototype, we gathered qualitative feedback on design components. These were critical to improve

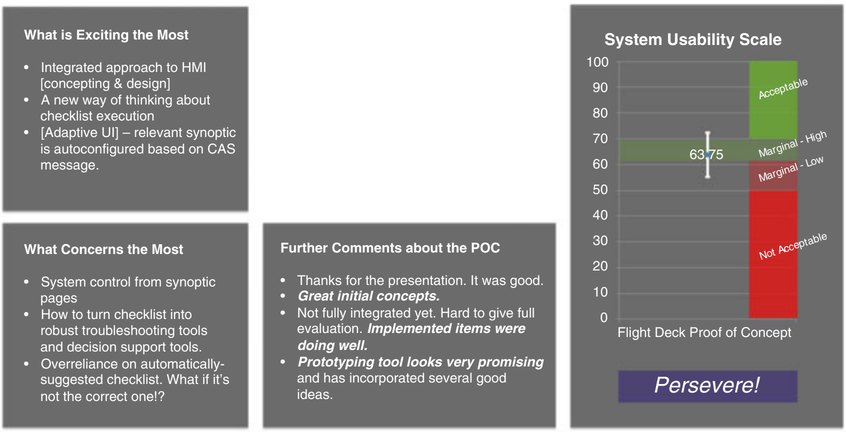


Fig. 5 User testing results: qualitative feedback as well as SUS score

the prototype, however, we also needed a quantifiable measure that we could use as a baseline for future updates and improvements. Due to the low level of fidelity of the prototype, we chose a subjective usability score. Many measures are available such as System Usability Score (SUS) [8], Usefulness, Satisfaction, and Ease of Use (USE) questionnaire [9], etc.

We used the SUS among the available measures. Figure 5 shows the qualitative feedbacks and quantitative data, SUS gathered during the user testing session.

3 Discussion

Purple Sky framework provides a systematic approache to brainstorming, designing, building, and testing. As we followed the process industry setting, we experienced some challenges. First, access to users is limited. The whole framework emphasizes on the continuous involvement of users, but practically, it is not easy to engage users throughout the process as desired due to many reasons such as customer relationship, business restrictions, and so forth. Second, it is still pervasive that once you build one, that is the final outcome. In engineering industries, the human-centered design and rapid prototyping are still new concepts. Lastly, a transition from the prototype to the product requires significant development re-work. The rapidness of the prototyping is based from the fact that we take advantage of rich capability of COTS tools. Therefore, we can quickly ideate, build, and test. The value still stands in that we can spend more time prototyping and validating the values of a product idea before building it into a product platform. However, a clear and economic way to transition a prototype into a product can bring further significant values.

4 Conclusion

We introduced Purple Sky framework for envisioning flight deck of future. We also applied this framework to designing a future flight deck concept for military helicopter operations. Although, this framework was established specifically for the flight deck design domain, the approach and the method can be applied to any complex systems. As a future work, we plan to apply this process in domains such as power plant operations and locomotive operation.

Acknowledgments This work was funded by General Electric Aviation. We would like to acknowledge designers who supported the work, Masaki Akiyama and Boris Soliz and partners from GE Aviation who helped us to engage with customers and participated in the co-design workshop, Norman Ovens, Michael Figard, and Sharon Green.

References

1. Subrahmaniyan, N., et al.: Human-centered design perspective on modern business jet flight deck operation. In: International Conference on Human-Computer Interaction in Aerospace 2016, ACM, Paris, France. Manuscript Submitted
2. 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)
3. Rau, K., Alcorn, K.: Co-creating the Industrial Internet. In: Global Service Design Network Conference 2015, Service Design Network, New York, NY (2015)
4. Roser, T., et al.: Co-creation: New Pathways to Value: An Overview. Promise & LSE Enterprise (2009)
5. Sanders, E.B.-N., Stappers, P.J.: Co-creation and the new landscapes of design. Co-design **4**(1), 5–18 (2008)
6. Gray, D., Brown, S., Macanufo, J.: Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers. O'Reilly Media, Inc. (2010)
7. Goodwin, K.: Designing for the Digital Age: How to Create Human-Centered Products and Services. Wiley (2011)
8. Brooke, J.: SUS-A quick and dirty usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)
9. Lund, A.M.: Measuring usability with the USE questionnaire. Usability Interface **8**(2), 3–6 (2001)

An Important Failure: Lessons from Daedalus and Icarus

Simon Cookson

Abstract The aim of this paper is to reflect on the ancient myth of Daedalus and Icarus, which was probably the first record of a flying accident, in order to see what lessons it holds for modern aviation. The paper provides an analytical description of the original text composed by Ovid, one of the canonical poets of the Roman world, two thousand years ago. This examination of Ovid's version of the myth reveals links to several important concepts in aviation human factors. Moreover, there are interesting divergences between the modern version of the story and that of Ovid. The paper concludes by proposing a new interpretation of the myth, the *Daedalus Dilemma*, which is relevant to contemporary flight operations.

Keywords Anxiety · Communication · Hazardous attitudes · High-risk technology · Human factors · Mythology

1 The Myth of Daedalus and Icarus

As it is usually retold today, the myth of Daedalus and Icarus is quite simple. Daedalus, a brilliant technician, is being held against his will by King Minos on the island of Crete. Desperate to escape, he makes wings for himself and his son, Icarus, and they attempt to fly to freedom. Daedalus warns his son to fly neither too low nor too high, lest the wings get damaged by seawater or the heat of the sun. Alas, neglecting his father's advice, Icarus flies too close to the sun with the result that his wings melt and he falls to his death in the sea below.

Modern retellings of the myth tend to concentrate on Icarus, neglecting the role played by Daedalus [1–3]. This is a relatively recent development; in classical times Daedalus—a fascinating character whose expertise extended to architecture, metalworking, sculpture and engineering—featured prominently in the story.

S. Cookson (✉)
Aviation Management Department, J. F. Oberlin University, 3758 Tokiwa-machi, Machida-shi, Tokyo 194-0294, Japan
e-mail: scookson@obirin.ac.jp

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_71

2 Ovid's *Metamorphoses*

The most complete and influential version of the myth that survives from antiquity is that written by the Roman poet Publius Ovidius Naso in his epic work, *Metamorphoses*.[1] Ovid was a talented writer living in turbulent times, and *Metamorphoses* is considered to be his "most elaborate and ambitious work" [4]. Completed in 8 AD, the 15-book collection of tales of transformation link to form a mythological 'history' of life since the Creation. The myth of Daedalus and Icarus appears in Book 8 of *Metamorphoses*, bracketed by two other stories featuring Daedalus. Preceding it is the story of Theseus and Ariadne escaping from the labyrinth, which Daedalus had designed to imprison the minotaur. Following it is the tale of Daedalus attempting to kill his nephew, Talus, who is transformed into a partridge.

Since classical times Ovid's version of the myth has informed countless works of art, with the period between the mid-16th and mid-17th centuries being particularly noteworthy. The torrent of paintings and etchings produced during this time included iconic images of *Landscape with the Fall of Icarus* from the studio of Pieter Bruegel the Elder. In the modern age the tale has been the theme of works such as the 1958 mural created by Pablo Picasso for the UNESCO headquarters in Paris, and Roger Brown's 1991 mosaic that adorns the entrance façade of a skyscraper in Chicago.

Why this particular myth—out of more than 250 recorded by Ovid in *Metamorphoses*—has been so enduring is not clear. Stephens and McCallum have pointed out that it is "succinct and quotable and lends itself to illustration" [5]. Another reason for the continued popularity is that it is a deceptively simple story that lends itself to multiple interpretations. Furthermore, amidst an abundance of tales of *gods* transforming people into animals, often birds, this myth stands out as the doomed attempt of a *man*, Daedalus, to act in a godlike way by transforming himself and his son into birdlike creatures.

3 The Relevance of Myths to Modern Aviation

An ancient myth dating back almost three millennia that tells of men flying with wings made of feathers and wax might seem to have little relevance to the technological world of modern aviation.

However, there are many connections between myths and aviation. We note the widespread use of mythology as a source of names for aircraft, engines and space

[1] Ovid's *Metamorphoses* has been translated into English many times. This paper makes use of the translation by D.E. Hill, part of which is reproduced in the appendix. It should be noted that Ovid himself had previously written another version of the myth of Daedalus and Icarus in Book 2 of *Ars Amatoria*.

programs: the Bristol *Jupiter* and *Pegasus* radial engines of the early 20th century; the Hughes H-4 *Hercules* and Lockheed C-130 *Hercules* transport aircraft; the pioneering *Mercury* and *Apollo* spaceflight programs; rockets and missiles such as *Atlas*, *Thor*, *Titan*, *Poseidon* and *Trident*; and, returning to the theme of this paper, the 1988 MIT *Daedalus* 88 human-powered aircraft which still holds world records for distance and endurance. Such names served in many cases to highlight the god-like power of the machines, and, in a period of tremendous technological advances, references to mythology imbued them with a talismanic confidence.[2] In fact, the early decades of the 20th century, which witnessed so many flights into the unknown, were called the "age of Icarus" [9].

The particular myth of Daedalus and Icarus has been read as a cautionary tale about the risks of new technology. The mythologist Joseph Campbell interpreted it thus, drawing the lesson that projects which attempt to break new ground or field new technology always carry "the danger of too much enthusiasm, of neglecting certain mechanical details" [10]. A tragic reminder of this warning occurred in 2014 when the Virgin Galactic Scaled Composites SpaceShipTwo experimental spaceplane crashed due to a problem with its feathering system [11].

In addition, the Icarus myth is relevant as a straightforward account of a flying accident. With their narrative structure, stories are powerful tools for conveying information. Indeed, the benefit of using stories of accidents or incidents in aviation training has been noted by Dekker: "Many sources, even within human factors, point to the value of storytelling in preparing operators for complex, dynamic situations in which not everything can be anticipated" [12].

Let us now consider the main events of Ovid's narrative and draw parallels to key concepts in contemporary human factors. It is argued here that ancient myth is a means of reflecting on modern aviation safety.

4 Human Factors in the Myth of Daedalus and Icarus

4.1 Factor 1: Anxiety and Get-Away-Itis

The opening lines of Ovid's story tell of the anxiety felt by Daedalus to leave the island of Crete and return to his homeland. Although he wanted to leave, Daedalus was unable to do so because the land and sea routes had been blocked by Minos, the king of Crete:

> "Daedalus meanwhile detested Crete and his long exile and, though affected by a longing for his native soil, had been shut off from it by the sea. 'He can block off', he said, 'the

[2]Gordon [6] observed that some early planes and engines used mythological names, while other commentators have discussed nuclear weapon systems that were named after mythical characters [7, 8]. This trend continues with modern military aircraft such as the Boeing KC-46 *Pegasus* and P-8 *Poseidon*.

> lands and seas, but the sky at least is open. We shall go that way! Though he may possess everything, Minos does not possess the air.'" [13]

A modern parallel may be drawn with the anxiety experienced by the flight crews involved in the 1977 runway collision at Tenerife. Following a diversion due to a bomb attack, the crew of Pan Am Flight 1736 wanted to depart as soon as possible, but were forced to wait more than two hours because the taxiway was blocked by another aircraft refueling. Meanwhile, the crew of the other aircraft, KLM Flight 4805, were also anxious to leave quickly because they were near their duty time limits for the month, and exceeding those limits would have resulted in serious consequences. In an analysis of this disaster, Weick suggested that anxiety contributed to the premature and unsafe departure of the KLM airplane [14]. The reasoning was as follows: after the original flight plan was disrupted, the KLM captain diverted high-level cognitive attention to resolving the problem of the duty time limits; this used up his limited information processing capacity, reducing his ability to monitor the radio and interpret cues from the outside environment; he was therefore unable to fully process critical radio transmissions to or from the Pan Am aircraft.

In the argot of present-day flight operations, *get-there-itis* is a widely recognised tendency for a pilot to continue towards the intended destination even though conditions have changed and become dangerous [15]. This tendency can be particularly strong during the approach phase at the end of a flight and is a form of *plan continuation bias*.[3] Daedalus was experiencing an analogous phenomenon, which occurs at the start of a flight and may be labeled *get-away-itis*, when a pilot feels undue anxiety to make a quick departure [17].

4.2 Factor 2: High-Risk Technology

Necessity being the mother of invention, the resourceful Daedalus, with all land and sea routes blocked, turned his attention to the air. He devised a revolutionary plan to escape by flying away: "directing his thoughts to skills unknown, he changed nature" [13]. The plan necessitated the use of unproven, high-risk technology, namely straws and feathers bound with thread and wax into the shape of birds' wings. For the reading audience two thousand years ago, this description of flight would have conjured up images of an amazing and almost incomprehensible level of imagination and technical skill.

The use of wings fastened with wax was to have tragic consequences for Icarus, just as the age of modern flight has been punctuated by accidents that followed the introduction of new technologies. From the crashes of the Wright brothers' Model C aircraft to the aforementioned break up of the SpaceShipTwo spaceplane, these accidents are solemn reminders that new technologies are invariably

[3]Plan continuation bias is an unconscious cognitive bias to persist with the original plan even though conditions are changing [16].

accompanied by new risks. In an example from 2015, the recently-introduced Airbus A400 M Atlas transport aircraft suffered its first crash following a problem with the engine control software [18].

4.3 Factor 3: Interruption and Distraction

As Daedalus prepared their flying equipment, Icarus interfered with his work by playing with the feathers and wax:

> "His boy, Icarus, was with him standing there and, unaware that he was handling his own danger, was now with shining face trying to catch the feathers moved by the wandering breeze, and now softening the yellow wax with his thumb, and hindering his father's amazing work with his play." [13][4]

In recent years, considerable attention has been devoted to studying the impact of distractions and interruptions on the ability of pilots to handle multiple concurrent tasks. In one such study, Loukopoulos, Dismukes and Barshi argued that such are the cognitive demands imposed by these "commonplace perturbations" that they "play a central role in pilot's vulnerability to error" [20]. The study cited about 60 incident reports from NASA's ASRS database in which perturbations resulted in flight crew failing to do something or omitting checklist items.

4.4 Factor 4: Poor Crew Communication

Before starting their flight, Daedalus advised Icarus to maintain a safe altitude and avoid being distracted by stars. He finished by telling his son to follow his instructions:

> "And he fitted out his son and, 'I warn you, Icarus,' he said, 'to run on the middle path in case, if you go too low, the water weighs down your wings, and, if too high, the sun's fire burns them. Fly between each of them and do not, I tell you, look at Boötes or Helice or Orion's drawn sword; pick your way where I lead.'" [19][5]

The oral instructions given by Daedalus took the form of a monologue: Daedalus spoke; Icarus did not. Significantly, in the absence of response, Daedalus failed to check his son's understanding. In the context of contemporary airline operations,

---

[4]It has been pointed out by Anderson that this scene clearly foreshadows the accident that later befalls Icarus, "when the easily molded wax melts and the feathers float away from his artificial wings" [19].

[5]Several commentators have noted that, for observers in the northern hemisphere, the constellations of Boötes and Helice (Ursa Major) appear high in the sky whereas Orion is lower and near to the horizon. Hill concluded that this reference to constellations "is an instruction to fly neither too high nor too low" [21].

this constitutes poor communication and is indicative of a steep authority gradient between the two aviators.

One of the goals of crew resource management (CRM) training programs is to enable all crew members to communicate effectively. Critical information must be shared. The aim is to prevent the recurrence of accidents, such as the crash of Korean Air Flight 801 in Guam in 1997. This crash was complex and—as is invariably the case with airline accidents—featured multiple causal factors. Nevertheless, the official accident report found the probable cause to be "the captain's failure to adequately brief and execute the nonprecision approach and the first officer's and flight engineer's failure to effectively monitor and cross-check the captain's execution of the approach" [22]. In other words, prior to the crash the captain's briefing was not effective and the other crew members did not articulate any concerns about the flight.

4.5 Factor 5: Hazardous Pilot Attitudes

After Daedalus fitted the wings to his son, the two aviators began their flight over the Aegean Sea. Icarus grew in confidence and then, recklessly deviating from his father's instructions, flew too high until the wax melted and his wings fell away:

> "…the boy, beginning to enjoy his bold flight, deserted his leader and, drawn by a desire for the heavens, took too high a course. The nearness of the raging sun softened the sweet-smelling wax that bound his wings. The wings had melted; he shook his bare arms…" [19]

Icarus's deviation from the flight plan could indicate any one of several attitudes that have been identified as hazardous to flying: anti-authority (he resented his father's control over the flight); impulsivity (he suddenly decided to climb); invulnerability (in his joyful inexperience he felt that nothing bad would happen); or macho (he wanted to show his father and the onlookers below how high he could fly). As Krause noted, these hazardous pilot attitudes "can negatively affect the judgment and decision-making processes" [21]. An example of an airline flight in which rules and procedures were disregarded was the 1994 crash of Aeroflot Flight 593 which occurred after a pilot brought his children into the cockpit and apparently let them take the controls. As a result the autopilot became disengaged and the flight crew were unable to regain control of the aircraft [23].

5 The Daedalus Dilemma

In the preceding section, the myth of Daedalus and Icarus was linked to five concepts in the domain of human factors. In addition, this short myth encapsulates a dilemma. On the one hand, it describes the godlike competence of the master technician Daedalus who, facing imprisonment on Crete by the powerful King Minos, achieved the following things:

- devised a way of escaping even though the land and sea routes were blocked;
- constructed wings that enabled humans to fly;
- assessed the risks of flying too high or low, or the risk of getting distracted;
- instructed his son, Icarus, with rules for dealing with these risks.

However, for all his knowledge and competence, an unpredictable human element still remained which Daedalus was unable to account for and which ultimately ruined his plan. This unpredictable element manifested itself at three points in the story. Firstly, during the preparation for flight, vulnerability to error was increased by anxiety, distraction and interruption. Secondly, although the contents of the pre-flight briefing were basically sound, Daedalus did not check whether Icarus understood them or whether he had any misgivings or questions. Finally, during the flight itself, recklessness led Icarus to disobey the instructions and fly higher until his wings melted.

We may summarize the *Daedalus Dilemma* as follows. For any technological endeavor, no matter how thorough the planning, design and preparation, there always remains an unpredictable element: the human factor.

6 Conclusion

In the celebrated poem, *Musée des Beaux Arts*, W.H. Auden wrote of a ploughman who was present as Icarus fell to his death and who “may have heard the splash, the forsaken cry, but for him it was not an important failure” [24]. Written on the eve of World War Two, the poem was a comment on the human proclivity for ignoring the suffering of others.[6] While acknowledging the importance of Auden’s message, the contention of this paper is that for aviation the death of Icarus was actually an important failure, which still has the capacity to inform modern flight operations.

The story of Daedalus and Icarus bears more than a passing resemblance to contemporary test flights, in which aviators map out the flight envelope limits of new flying machines. Icarus tragically exceeded those limits and perished in the attempt. However, it is important to note that Ovid’s version of the myth ends with Daedalus completing his flight, a detail often overlooked in modern retellings. In other words, the new technology succeeded, but at great cost.

In antiquity one of the roles of Daedalus, depicted in a cycle of myths, was to act as an embodiment of “translatio studii” and enable the transmission of knowledge to new generations [2]. This role remains relevant to the present day. As we embark

[6]Auden wrote this poem in part as a response to a painting by Pieter Bruegel the Elder (or his circle) titled *Landscape with the Fall of Icarus*, which is housed in the Musées Royaux des Beaux-Arts, in Brussels.

on ever more complex endeavors with the dazzlingly high technology of the 21st century, the myth of Daedalus and Icarus is a salient reminder that we should never lose sight of the risks associated with the unpredictable human factor.

Acknowledgments I am extremely grateful to Professor John Maher at International Christian University, Tokyo, for ideas and comments that made this paper possible and greatly improved the manuscript.

Appendix: Ovid's Myth of Daedalus and Icarus

D.E. Hill's translation of the myth of Daedalus and Icarus from Book 8 of Ovid's *Metamorphoses* is reproduced below [25].

Daedalus meanwhile detested Crete and his long
exile and, though affected by a longing for his native soil,
had been shut off from it by the sea. 'He can block off', he said, 'the lands
and seas, but the sky at least is open. We shall go that way!
Though he may possess everything, Minos does not possess the air.'
He spoke and, directing his thoughts to skills unknown,
he changed nature. For he put feathers in a row
beginning from the smallest, with the shorter following the long
so that you would think they had grown on a slope; so does the rustic
pipe gradually expand with unequal stalks.
Then he bound them together in the middle and at the base with a thread and wax,
and when they had been arranged so, he bent them to a slight curve
to imitate real birds. His boy, Icarus, was with him
standing there and, unaware that he was handling his own danger,
was now with shining face trying to catch the feathers moved
by the wandering breeze, and now softening the yellow wax
with his thumb, and hindering his father's amazing work
with his play. After he had put the last touch
to what he had begun, the craftsman balanced his own body
on a pair of wings and hovered in the air they were moving.
And he fitted out his son and, 'I warn you, Icarus,' he said,
'to run on the middle path in case, if you go too low,
the water weighs down your wings, and, if too high, the sun's fire burns them.
Fly between each of them and do not, I tell you, look at
Boötes or Helice or Orion's drawn sword;
pick your way where I lead.' At the same time he gave him flying
instructions and fitted the unfamiliar wings to his shoulders.
In the midst of his work and warnings, the old man's cheeks grew wet

and the father's hands shook. He gave kisses to his son
that would never be repeated and, lifted on his wings,
he flew in front and feared for his companion, like a bird who has
led her tender offspring from a high nest into the air,
and he urged him to follow and trained him in ruinous skills.
He both moved his own wings and looked back at his son's.
And someone while trying to catch fish with a trembling rod,
or a shepherd leaning on his staff, or a ploughman on his shaft
saw them and was down dumbfounded and, since they could press through the air,
believed they were gods. And now Juno's Samos
was on the left side (Delos and Paros had been put behind them)
and Lebinthos was on the right as was honey-rich Calymne,
when the boy, beginning to enjoy his bold flight,
deserted his leader and, drawn by a desire for the heavens,
took too high a course. The nearness of the raging sun
softened the sweet-smelling wax that bound his wings.
The wings had melted; he shook his bare arms
but, lacking oarage, he did not grip the air at all,
but, with his mouth crying out his father's name, he was received
by the aquamarine water, which took its name from him.
But the unhappy father, not now a father, said, 'Icarus,'
'Icarus,' he said, 'where are you? In what region should I seek you?'
'Icarus,' he was saying; and caught sight of feathers in the waves
and cursed his skills and buried the body
in a tomb; the land was called after the name of the boy in the tomb.

References

1. Campbell, J., Moyers, B.D., Flowers, B.S.: The Power of Myth. Anchor Books, New York (1991)
2. Ide, K.J.: The Daedalus of history and myth: the meaning of creation in literature from Homer to Joyce. Unpublished thesis, Department of Comparative Literature, University of California, Davis (2011)
3. Stephens, J., McCallum, R.: Retelling Stories, Framing Culture: Traditional Story and Metanarratives in Children's Literature. Garland Pub, New York (1998)
4. Gould, H.E., Whiteley, J.L.: Ovid Metamorphoses VIII, p. xiii. Bristol Classical Press, London (2008, original work published 1940)
5. Stephens, J., McCallum, R.: Retelling Stories, Framing Culture: Traditional Story and Metanarratives in Children's Literature, p. 69. Garland Pub, New York (1998)
6. Gordon, A.: Naked Airport: A Cultural History of the World's Most Revolutionary Structure. The University of Chicago Press, Chicago (2008)

7. Montgomery, M.: An Introduction to Language and Society, 3rd edn. Routledge, London (2008)
8. Simpson, P., Mayr, A.: Language and Power: A Resource Book for Students. Routledge, London (2010)
9. Boitani, P.: Winged Words: Flight in Poetry and History, p. 32. The University of Chicago Press, Chicago (2007)
10. Campbell, J., Moyers, B.D., Flowers, B.S.: The Power of Myth, p. 162. Anchor Books, New York (1991)
11. National Transportation Safety Board: Public Meeting of July 28, 2015.: In-Flight breakup during test flight scaled composites SpaceShipTwo, N339SS near Koehn Dry Lake, California October 31, 2014. NTSB, Washington, DC (2015)
12. Dekker, S.: Ten Questions about Human Error, p. 196. Lawrence Erlbaum Associates, New York (2005)
13. Hill, D.E.: Metamorphoses V-VIII, p. 111. Aris & Phillips, Warminster, Wiltshire (1992)
14. Weick, K.E.: The vulnerable system: an analysis of the tenerife air disaster. J. Manag. **16**(3), 571–593 (1990)
15. Aviation Safety Reporting System.: CALLBACK. Issue 389, June 2012. NASA Aviation Safety Reporting System, Moffett Field, California (2012)
16. Dismukes, R.K., Berman, B.A., Loukopoulos, L.D.: The Limits of Expertise: Rethinking Pilot Error and the Causes of Airline Accidents. Ashgate Publishing, Aldershot (2007)
17. Aviation Safety Reporting System.: CALLBACK. Issue 275, July 2002. NASA Aviation Safety Reporting System, Moffett Field, California (2002)
18. Airbus Defence & Space.: Press Release of June 3, 2015: Statement regarding Accident Information Transmission (AIT) to A400M operators as follow up to AOT of 19 May. Airbus Defence & Space, Toulouse (2015)
19. Hill, D.E.: Metamorphoses V-VIII, p. 113. Aris & Phillips, Warminster, Wiltshire (1992)
20. Loukopoulos, L.D., Dismukes, R.K., Barshi, I.: The Multitasking Myth: Handling Complexity in Real-World Operations, pp. 46–47. Ashgate Publishing, Farnham (2009)
21. Krause, S.S.: Aircraft Safety: Accident Investigations, Analyses, and Applications, 2nd edn, p. 9. McGraw-Hill, New York (2003)
22. National Transportation Safety Board.: Aircraft Accident Report: Controlled Flight into Terrain, Korean Air Flight 801, Boeing 747–300, HL7468, Nimitz Hill, Guam, August 6, 1997. NTSB/AAR-00/01. NTSB, Washington, DC, p. 175 (2000)
23. Learmount, D., Velovich, A.: FDR Backs A310 Crash Allegations. Flight International. 27 April–3 May, p. 5 (1994)
24. Auden, W.H.: Collected Poems. Random House, Inc., New York (2007, original work published 1976), p. 179
25. Hill, D.E.: Metamorphoses V-VIII, pp. 111–113. Warminster, Wiltshire, Aris & Phillips (1992)

Improved User Acceptance During Stepwise Air Traffic Control Display Functionality Introduction

Oliver Ohneiser

Abstract This paper addresses migration tolerant human machine interfaces for approach controllers considering shifting air traffic control (ATC) approach from distance- via time- and trajectory- to performance-based. It includes a series of learning steps between prototypically implemented display revisions that were evaluated by ten controllers. Nine iterative steps comprised distance marking, merge point appearances, new symbols, and parallel screen border orthogonal route structures for unidirectional aircraft movement on resulting monitoring display. Study participant group G9 worked with all ten display revisions experiencing nine transition steps; group G2 only used revision zero, five, and final display layout. Participants' ratings on usability, learnability and operational acceptance dropped down for G2 but not for G9 especially when transitioning to last display revision with different step sizes. Thus, introducing new ATC display functionalities in a row of small logical and consecutive instead of very few broad integration steps improves user acceptance.

Keywords Air traffic controller · Situation data display · Human machine interface · Migration tolerance · Transition step · User acceptance

1 Introduction

Today, air traffic controllers (ATCO) mainly use radar screens to guide air traffic safely in the terminal maneuvering area with a distance-based air traffic control (ATC) approach [1]. In the future, new requirements will have to be fulfilled by controllers due to SESAR (Single European Air Traffic Management Research Programme) and NextGen (Next Generation Air Transportation System) [2–4]. These programmes plan three operational timely overlapping steps going forward

O. Ohneiser (✉)
Institute of Flight Guidance, German Aerospace Center (DLR),
Lilienthalplatz 7, 38108 Brunswick, Germany
e-mail: Oliver.Ohneiser@DLR.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_72

from the current distance-based ATC approach via a time- and a trajectory- to a performance-based approach of operation methods [5].

This goes along with greater parts of monitoring tasks for controllers and is difficult to handle with current displays that hardly have any visual assistance functionality [6–8]. Actual operational displays are basically able to show aircraft positions, waypoints, routes, or runways. However, it is difficult to estimate durations of aircraft needed to reach certain waypoints by estimating not only distance but also speed and to derive accurate commands.

In the course of four-dimensional trajectories comprising latitude, longitude, altitude, and time, at least deviations from the planned status have to be displayed for regulation purposes. This is also true for performance parameters in the third ATC step. Thus, two questions about the HMI design after each ATC approach step and the transitions between those steps arise.

Current HMIs in the air traffic domain have life cycles of decades [9]. In contrast to current broad integrations of new controller working position systems and displays at a single point in time, systems will have to be more flexible in the future.

Adaptations in controller displays should be integrated iteratively and user-centered [10, 11]. To avoid extensive system user trainings and to increase user acceptance simultaneously, new functionalities can be integrated in a row of few small consecutive steps instead of great steps [12].

This paper addresses related work in Sect. 2. Section 3 contains the concept of a migration tolerant human machine interface for approach controllers with a series of learning steps between prototypically implemented display revisions. Section 4 outlines the evaluation study design with corresponding results in Sect. 5. Section 6 discusses those results whereas Sect. 7 summarizes and concludes.

2 Related Work

Next to new requirements due to shifted ATC approaches, there is potential by using air-ground data link negotiated four dimensional trajectories, flight management system (FMS) data on ground and more precise navigation capabilities [13]. Accuracy due to time and airline preferred flight profiles [14] regarding whole trajectories or only waypoints can be achieved using advanced flight management systems [15]. These assumptions could be considered in advance for new and enhanced controller display concepts.

In addition, controllers' tasks will change especially with time and trajectory constraints. The role of an active managing controller will more and more be replaced by an actively monitoring controller who intervenes in abnormal situations or at strong deviations [8, 16, 17]. Besides, flows of aircraft may be considered to a greater degree than today if applicable in new air traffic management procedures [18, 19].

Some research has been performed on time-based ATC and related controller displays in the past [20, 21]. There are also ideas about reducing complexity on controller displays by removing lateral information and aircraft moving in only one direction [22, 23]. However hardly any research concerns neither ATC display functionalities corresponding to future requirements nor a user-centered and iterative introduction of them.

3 Concept of Migration Tolerant Display Steps

The current radar-based controller display was modified using several consecutive steps and considering design metaphors. Each of the new nine iterative display revisions has only one logical element that was changed (see Fig. 1) as humans do not like permanent changes [24].

Supporting the controllers in some current straining tasks may relieve them and free cognitive resources to be used for new additional tasks. The first display revision looks pretty much the same as current operational controller displays to have a common start point. The development process followed user centered loops (develoops) [25] and included improvements after prototype feedback from air traffic management (ATM) experts, and controllers.

Step one included marking of distances in nautical miles on standard arrival routes (STAR) for better distance estimation in the outer terminal manoeuvring area (TMA). The second step comprised symbols for cardinal directions of aircraft entering at an initial approach fix (IAF) on their route to the runway.

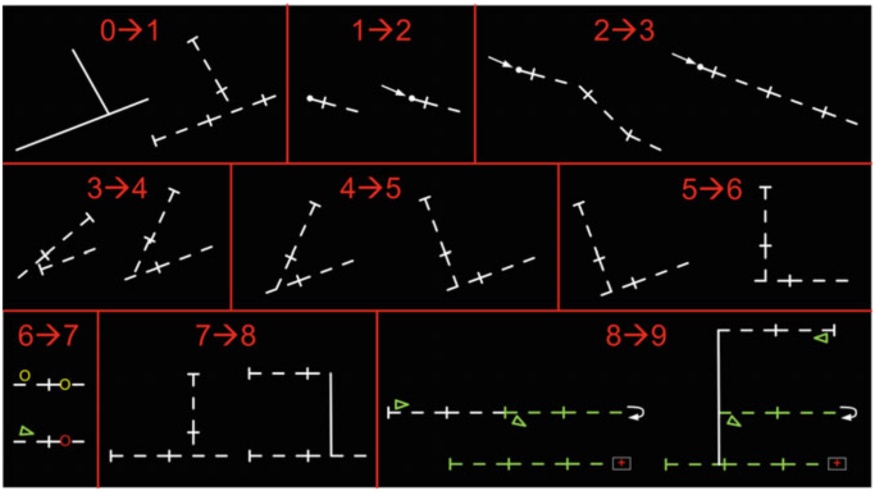


Fig. 1 All changes of controller display elements at the nine transition steps between ten display revisions (*white* and *green lines* are routes; *circles* respectively *triangles* show aircraft) [30]

Other domains such as the rail traffic use simple complexity reduced maps with mainly rectangular routes for passengers or supervisors. It is only of interest how to get from a start to an end point and where intersections are. This concept could also be adapted to the STARs of air traffic that merge to common routes at specific points for the last approach phase. The third step consisted of straightening displayed routes while real aircraft flown routes did not change. Distances of aircraft on different STARs to a single merge point are now more intuitive to supervise. An aircraft path deviation from the STARs on the left or right side is still recognizable in the display.

However, the missing geographic resolution on display makes reasonable heading commands of controllers difficult. Reliable locating of aircraft is even more important in order to design displays that are more abstract. In addition, conformance monitoring tools need to support controllers by visualizing aircraft states regarding their plan in a trajectory-based environment [26]. In the course of adaptations, controller roles therefore also changed. Monitoring controllers use the resulting complexity reduced migration tolerant display. They are assisted by executive controllers with conventional radar displays. Hence, monitoring controllers may delegate conflicts or propose solutions to the executive acting as a team [27, 28].

The straightened routes have different angles at their merge points. These angles were homogenized to multiples of 45° in step four to ease conflict geometry [29]. The fifth display revision led to a rectangular structure of routes. The whole view was rotated into a runway dependent horizontal view in display stage six. This avoids cognitive rotation of the displayed airspace by controllers or tilting one's head. Afterwards, the merge point appearance was adapted to merge lines with STARs approaching from one side to different ends of the merge line. These lines support better and earlier vertical comparison of aircraft on STARs. Earlier adaptations of the controller to avoid that for example more than one aircraft reaches a start point of a downwind at the same time can lead to less flown route and fuel consumption. Hence, this approach focusses on the trombone approach with downwind, base, and final leg, but is but limited to this kind of approach procedure. Performance-coded aircraft symbols were introduced in step eight. These symbols indicate individual optimization potential of aircraft. Advanced FMS equipped aircraft are symbolized by circles as they already should follow an optimized plan. The number of edges of the triangles representing alternative aircraft reveals their individual optimization potential.

The last revision included a parallel and screen border orthogonal route structure for unidirectional movement of aircraft on the resulting monitoring display. Only aircraft on downwind, an area that should be avoided due to its delaying function and thus kerosene wasting characteristic, fly in the opposite direction.

4 Design of Evaluation Study with Controllers

During an evaluation study at DLR Braunschweig ten controllers (one female; age [years]: Ø: 41, SD: 9) of DFS (DFS Deutsche Flugsicherung GmbH), the German air navigation service provider, had to work with the implemented display set.

Participants were split into two groups and had to absolve three simulation runs after training. The air traffic simulation runs consisted of a replay with forced deviations of aircraft and conflicts to be detected.

The first group (G9) had to work with all ten display revisions (revision 0 to 2 in run 1; revision 3–6 in run 2; revision 7–9 in run 3) and therefore experienced nine transition steps. The display view automatically changed after roughly 12 min during the scenario. The second group (G2) only went through two transitions between the state-of-the-art display revision zero (run 1), the intermediate stage number five (run 2), and the final display layout (run 3).

Controllers had to answer a transition questionnaire and rated items on the system usability scale (SUS) [31] on a five point Likert scale [32] after each display step. Questions T01-T05 concerned display transitions and their handling. Questionnaire items T06-T10 included a comparison of different characteristics between each current and preceding display revision. Group G2 was also asked four times in run 2 respectively three times in run 3 but had to evaluate the same display revision multiple times.

5 Evaluation Results

The ratings on questionnaire items T01-T05 (Reasonability of display transition; Clarity of aircraft movement on screen; Non-confusion of display; Learnability of display step; Planning and delegation during simulation task) are shown in Fig. 2 for group G9 and their nine display revisions respectively in Fig. 3 for group G2 regarding display revision 5 and 9 with black lines for standard deviations.

The great majority of rating averages for group G9 has a value of 2 or above which means affirmation to the five item aspects as a trend. Almost the same is true for group G2 when working with display revision 5.

Nevertheless, the acceptance dropped down when experiencing the transition to revision 9 in one step compared to the four intermediate steps of group G9. This shows the disadvantages of too big transition steps. However, the standard deviations became greater in further display steps. This shows the variety of participating controllers that could or could not imagine to work with the migration tolerant displays.

Figures 4 and 5 display the ratings on questionnaire items T06-T10 (Degree of display change; Supervision of air traffic; HMI complexity; Recognition of conflicts; Solution of specific tasks).

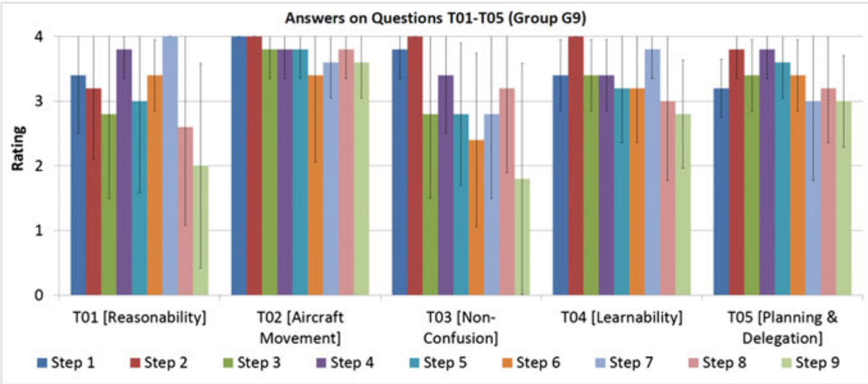


Fig. 2 Transition questionnaire results on items T01-T05 of group G9 (the higher the rating, the more positive the controllers' opinion)

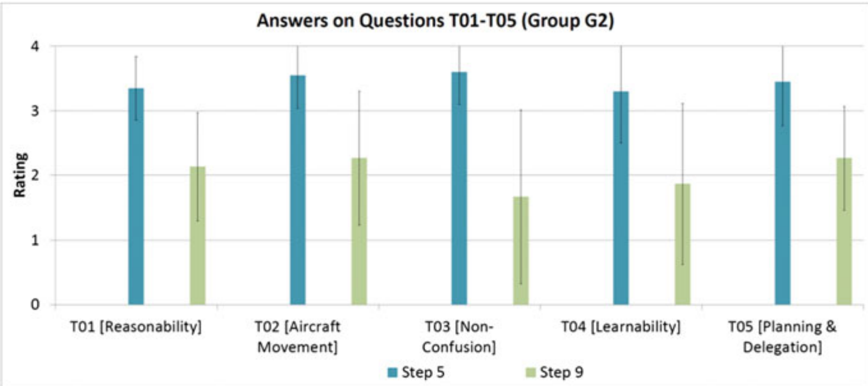


Fig. 3 Transition questionnaire results on items T01-T05 of group G2 (the higher the rating, the more positive the controllers' opinion)

Results of item T06 show that step 2 and 7 are too small, step 9 is slightly too large. In general, the degree of changes between display revisions fits very well. For items supervision of traffic, complexity, task processing and mainly conflict recognition some improvements especially at steps 4 to 7 can be seen. Again, the values of group G2 are quite comparable at step 5 due to the lower degree of change between revisions 0 and 5. The acceptance values fell below 2 in average for items T06, T07, T08, and T10, which means a negative change. The comparison of step 9 values between G9 and G2 repeatedly show that acceptance is higher when introducing new functionalities in a row of steps than just one step. Nevertheless, conflict recognition seems to be even better for group G2 with the last display revision than with current controller displays.

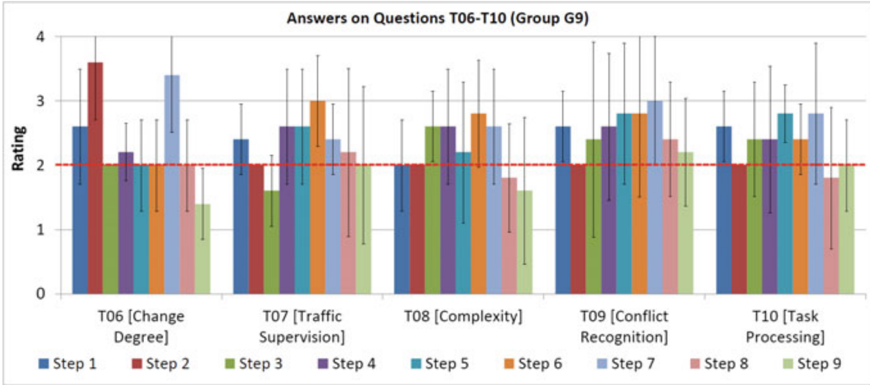


Fig. 4 Transition questionnaire results on items T06-T10 of group G9 (*0* negative change respectively "too big" at T06; *2* no change; *4* positive change respectively "too small" at T06). Most of the ratings of group G9 have the value 2 or above which means no respectively slightly positive change compared to the last display revision

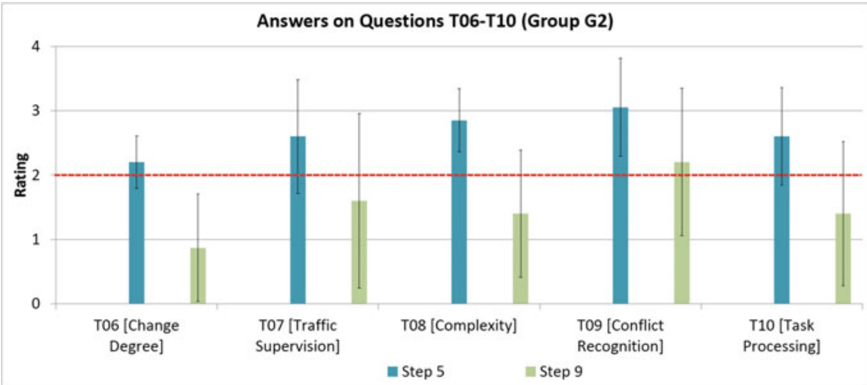


Fig. 5 Transition questionnaire results on items T06-T10 of group G2 (*0* negative change respectively "too big" at T06; *2* no change; *4* positive change respectively "too small" at T06)

Figures 6 and 7 illustrate the added score of the system usability scale of group G9 respectively G2.

The rating sum has a range between 0 and 100. Values above roughly 75 indicate good usability (excellent usability if closer to 100), around 65 with medium usability, and below 60 poor usability. The average rating of group G9 is always above 60 and has values up to 82. However, even the first controller displays that look very much the same as current operational displays have ratings below 80. Therefore, a normalization with the baseline of the first display would be possible. Enhanced display revisions would then have quite comparable usability to today's

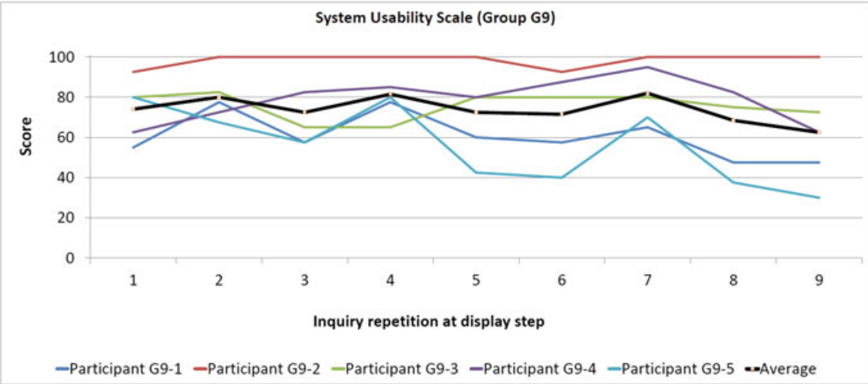


Fig. 6 System usability questionnaire results of group G9 (the higher the score, the better)

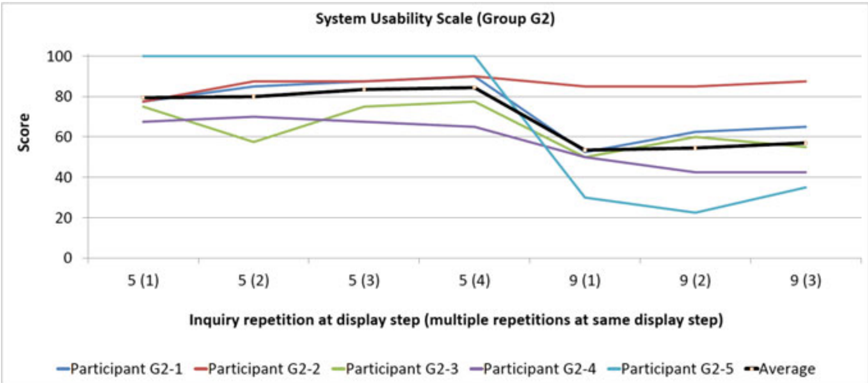


Fig. 7 System usability questionnaire results of group G2 (the higher the score, the better)

displays. Another interpretation would be an even better usability in some of the further display steps.

Group G2 rated system usability as good in the four iterative questions at step 5, but as poor in the three answers at step 9. This fact again shows the disadvantages of introducing new elements in single great steps when comparing results of G9 and G2.

6 Discussion of Study Results and Controller Comments

The comparison of system usability and transition questionnaire ratings revealed user acceptance differences between the two groups experiencing different numbers of intermediate steps. Results showed that participants' usability and learnability ratings were quite similar for display revision five comparing both groups, but dropped down for the group with only two transitions when using display revision nine. The adequate dimension of learning steps between all ten implemented display revisions was hit quite well due to controllers' feedback. The SUS scores were quite stable overall despite the big change from display revision 0 to 9.

Furthermore, controller tasks such as traffic supervision with conflict recognition were possible with the migrated display and new controller monitoring roles. No significant disadvantage of migration tolerant display steps could be stated compared to current controller displays. In contrary, some trends for reduced complexity and better conflict recognition were found. It was shown that introducing new ATC display functionalities in small steps is superior against a one-step approach.

7 Summary and Outlook

A migration tolerant concept for a set of controller human machine interfaces with small consecutive learning steps was implemented and evaluated. Results show good acceptance ratings of display characteristics such as learnability and task handling using the first display revisions. However, introducing further air traffic control display functionalities stepwise (group G9) results in improved user acceptance compared to broad integrations at a single point in time (group G2).

The need of migratory implementation steps with an evolutionary concept has also actually been stated by the SESAR ATM master plan [33]: "Provisions: These will be made for the training needs that enable effective and optimal change management. This will support a transition path that considers the influence of successive migratory implementation steps towards the agreed concept evolution and minimizes the extent to which the human system relies on phenomena such as mode switching".These high-level requirements for smooth transitions were instantiated with the display steps in this paper.

Following, the migration tolerant display steps are an eligible concept to face future requirements of the air traffic management domain and to improve user acceptance during stepwise ATC display functionality introduction.

References

1. Brooker, P.: Air traffic control separation minima: part 1—the current stasis. J. Navig. (Online) **64**, 449–465 (2011)
2. SESAR.: European ATM master plan: the roadmap for sustainable air traffic management, Brussels (2012)
3. SESAR.: Modernising the European Sky, SESAR, Brussels (2014)
4. Federal Aviation Administration: NextGen—Implementation Plan. NextGen Integration and Implementation Office, Washington (2014)

5. Brooker, P.: Air traffic control separation minima: part 2—transition to a trajectory-based system. J. Navig. (Online) **64**, 673–693 (2011)
6. Kesseler, E., Knapen, E.: Designing future advanced controller displays, NLR-TP-2004-307, NLR, Amsterdam (2004)
7. Metzger, U., Parasuraman, R.: Automation in future air traffic management: effects of decision aid reliability on controller performance and mental workload. Human Factors J. Human Factors Ergon. Soc. **47**(1), 35–49 (2005)
8. Neal, A., Flach, J., Mooij, M., Lehmann, S., Stankovic, S., Hasenbosch, S.: Envisaging the future air traffic management system. Int. J. Aviation Psychol. **21**(1), 16–34 (2011)
9. INDRA: DFS UAC Centre Karlsruhe Operational since Dec 2010 delivered by Indra, INDRA, DFS Deutsche Flugsicherung, Madrid (n.d.)
10. Kapp, V., Hripane, M.: Improving TMA sequencing process: innovative integration of AMAN constraints in controllers environment. In: 2008 IEEE/AIAA 27th Digital Avionics Systems Conference, IEEE, pp. 3.D.1-1–3.D.1-9 (2008)
11. König, C., Hofmann, T., Bruder, R.: Application of the user-centred design process according ISO 9241-210 in air traffic control. Work J. Prev. Assess. Rehabil. **41**(1), 167–174 (2012)
12. Ohneiser, O.: Migration tolerant human machine interfaces—are we ready for the future?, original German title: Migrationstolerante Mensch-Maschine-Schnittstelle für Lotsenarbeitsplätze – Sind wir auf die Zukunft vorbereitet?. In: Future Visualization Systems for Vehicle- and Process Guidance (DGLR-Report 2012-01), original German booktitle: Fortschrittliche Anzeigesysteme für die Fahrzeug- und Prozessführung, M. Grandt, S. Schmerwitz (eds.) German Society of Aerospace, Bonn, pp. 283–286 (2012)
13. Ohneiser, O., Heesen, M., Flemisch F., Rataj, J.: Migration tolerant human machine interface concepts in the domain of air traffic management and automotive. In: 63rd German Aerospace Congress, Augsburg, German Society of Aerospace (2014)
14. Air Traffic Management Requirements and Performance Panel (ATMRPP).: Proposal 15: for the development of TBO concept: Managing unpredictable evolution in trajectory based operations. ICAO, ATMRPP WG/25-WP/601, Toulouse (2014)
15. Kuenz, A., Korn, B.: Enabling green profiles for today's traffic mixture in high density. In: Proceedings of the 2009 Integrated Communications, Navigation and Surveillance Conference, IEEE, pp. 15–24 (2009)
16. Erzberger, H.: The automated airspace concept. In: NASA Ames Research Center, Santa Fe, Fourth USA/Europe Air Traffic Management Research and Development Seminar (2001)
17. Endsley, M., Sollenberger, R.L., Nakata, A., Stein, E.S.: Situation awareness in air traffic control: enhanced displays for advanced operations, Federal Aviation Administration, DOT/FAA/CT-TN00/01, Atlantic City (2000)
18. Duong, V., Gawinowski, G., Nicolaon, J.-P., Smith, D.: Sector-less air traffic management. In: EUROCONTROL, Santa Fe, Fourth USA/Europe Air Traffic Management Research and Development Seminar (2001)
19. Helmke, H.: Time-based arrival management: new controller support tools are necessary to ensure best-possible use of the available airspace, Air Traffic Technology International 2011 (2011)
20. Ohneiser, O.: Flight guidance support to integrate conventional equipped aircraft into a time based arrival flow, original German title: Führungsunterstützung zur Integration konventionell ausgerüsteter Luftfahrzeuge in einen zeitbasierten Anflugstrom. In: Future Visualization Systems for Vehicle- and Process Guidance (DGLR-Report 2012-01), original German booktitle: Fortschrittliche Anzeigesysteme für die Fahrzeug- und Prozessführung, M. Grandt, S. Schmerwitz (eds.) German Society of Aerospace, Bonn, pp. 175–192 (2012)
21. Ohneiser, O., Temme, M.-M., Rataj, J.: Trawl-net technology for timely precise air traffic controller turn-to-base commands, Lisbon. In: Eleventh USA/Europe Air Traffic Management Research and Development Seminar (2015)
22. Benhacène, R., Anthony, M., Rousselle, M.-P.: As rapid as paper strips? Evaluation of Vertidigi a new control tool for terminal sectors, Baltimore. In: Sixth USA/Europe Air Traffic Management Research and Development Seminar (2005)

23. de Leege, A.M.P., van Paassen, M.M., in't Veld, A., Mulder, M.: Time–space diagram as controller support tool for closed-path continuous-descent operations.J. Aircraft **50**(5), 1394–1408 (2013)
24. Axt, P., Axt-Gadermann, M.: From the fortune of small steps, original German title: Vom Glück der kleinen Schritte: Step by step - das Gewohnheitstier in uns überlisten - Das 52-Wochen-Gesundheitsprogramm, München, Herbig (2008)
25. Rauterberg, M.: Usability Engineering Methods and Tools: User Centered Design. Technical University Eindhoven, Eindhoven (1998)
26. Ohneiser, O., Beddig, S.: Air traffic controller support for conformance monitoring and active guidance of air Traffic, original German title: Fluglotsen-Unterstützung für die Konformitätsüberwachung und aktive Führung des Luftverkehrs, 62nd German Aerospace Congress, Stuttgart, German Society of Aerospace (2013)
27. Malakis, S., Kontogiannis, T., Kirwan, B.: Managing emergencies and abnormal situations in air traffic control (part I): taskwork strategies. Appl. Ergon. **41**(4), 620–627 (2010)
28. Malakis, S., Kontogiannis, T., Kirwan, B.: Managing emergencies and abnormal situations in air traffic control (part II): taskwork strategies. Appl. Ergon. **41**(4), 628–635 (2010)
29. Palmer, E.M., Clausner, T.C., Kellman, P.J.: Enhancing air traffic displays via perceptual cues. ACM Trans. Appl. Percept. **5**(1), 1–22 (2008)
30. Ohneiser, O.: Migration tolerant transitions of incremental display steps for air traffic controllers, original German title: Migrationstolerante Transitionen von inkrementellen Displaystufen für Fluglotsen, Dissertation Technische Universität Braunschweig, Braunschweig, (2016, to be published)
31. Brooke, J.: SUS - A quick and dirty usability scale. In: Jordan, P.W., Thomas, B., McClelland, I.L., Weerdmeester, B.A. (eds.) Usability Evaluation in Industry, pp. 189–194. Taylor and Francis, London (1996)
32. Likert, R.: A technique for the measurement of attitudes. Arch. Psychol. **22**(140), 5–55 (1932)
33. SESAR.: The roadmap for delivering high performing aviation for Europe. European ATM Master Plan, Brussels (2015)

EEG-Based Human Factors Evaluation of Conflict Resolution Aid and Tactile User Interface in Future Air Traffic Control Systems

Xiyuan Hou, Fitri Trapsilawati, Yisi Liu, Olga Sourina, Chun-Hsien Chen, Wolfgang Mueller-Wittig and Wei Tech Ang

Abstract Currently, Air Traffic Control (ATC) systems are reliable with automation supports, however, the increased traffic density and complex air traffic situations bring new challenges to ATC systems and air-traffic controllers (ATCOs). We conduct an experiment to evaluate the current ATC system and test conflict resolution automation and tactile user interface to be the inputs of the future ATC system. We propose an Electroencephalogram (EEG)-based system to monitor and analyze human factors measurements of ATCOs in ATC systems to apply it in our experiment. The EEG-based tools are used to monitor and record the brain states of ATCOs during the experiment. Real-time EEG-based human factors evaluation of an ATC system allows researchers to analyze the changes of ATCOs' brain states during the performance of various ATC tasks. Based on the analyses of the objective real time data together with the subjective feedback from ATCOs, we are able to reliably evaluate current ATC systems and refine new concepts of future ATC system.

X. Hou (✉) · Y. Liu · O. Sourina · W. Mueller-Wittig
Fraunhofer IDM@NTU, Nanyang Technological University, Singapore, Singapore
e-mail: houxy@ntu.edu.sg

Y. Liu
e-mail: liuys@ntu.edu.sg

O. Sourina
e-mail: eosourina@ntu.edu.sg

W. Mueller-Wittig
e-mail: askwmwittig@ntu.edu.sg

F. Trapsilawati · C.-H. Chen · W.T. Ang
School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore, Singapore
e-mail: fitritra001@e.ntu.edu.sg

C.-H. Chen
e-mail: mchchen@ntu.edu.sg

W.T. Ang
e-mail: askwmwittig@ntu.edu.sg

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_73

Keywords EEG · Human factors · Brain states monitoring · Air traffic control

1 Introduction

Although performance and reliability of current Air Traffic Control (ATC) systems has been improved with automation supports, the increased traffic density and complex air traffic situations bring additional requirements and new challenges to ATC systems and air-traffic controllers (ATCOs). To design a future ATC platform which can provide more effective and robust handling of heavy traffic situations, states of art technologies like touch and tactile human computer interface, interactive 3D situation displays and advanced CRA software will be integrated to current ATC platform. However, the standard evaluation method that uses questionnaires after each assessment can only give an overall rating of the performed task. It cannot tell designers how the workload and emotions are changing during the task performance in a complex traffic situation. This information may be estimated from some other performance factors, but it cannot be done in high time resolution using traditional methods. So there is a need for the tools to objectively estimate how novel interfaces affect ATCOs during operations. To solve this problem, we propose to use reliable brain computer interface (BCI) to measure performance of ATCOs in different ATC experiments. By using such bio-signal technology with standard evaluation methods, we can enhance and refine design and development of a new ATC platform. To our best knowledge, we are the first to develop real-time workload, emotion and stress recognition algorithms that use fewer electrodes and have good accuracy that could be used for ATM system evaluation. A real-time brain states monitoring system in which emotion, attention, workload, and stress recognition algorithms can be recognized in real-time is applied for evaluation of the future workplace of ATCOs.

We conduct an experiment to evaluate the costs and benefits of conflict resolution automation and tactile user interface in future ATC systems. In the user study, we evaluate the current ATC system with conflict resolution automation and tactile user interface used as inputs. ATCOs and students with ATC knowledge are instructed to complete ATC tasks in three conflict resolution aid scenarios including reliable, unreliable, and manual conditions with or without tactile user interface. During this user study, objective human factors measurements including mental workload, stress, and emotion of ATCOs while performing ATC tasks are obtained real time using an Electroencephalogram (EEG) device.

In this paper, we propose an EEG-based system to monitor and analyze human factors measurements of ATCOs in ATC systems. The EEG-based tools are used to monitor and record the brain states of ATCOs during the experiment. In subjective human factors studies, the data of mental workload, stress, emotion et al. are obtained through questionnaires that are administered upon completion of each task or/and after an experiment. However, this method only offers the overall evaluation of ATCOs performance. Real-time EEG-based human factors evaluation of an ATC system allows researchers to analyze the changes of ATCOs' brain states during the performance of various ATC tasks. The data can be analyzed during or at any time interval starting from 1/32 s. Machine learning techniques are applied to the EEG data to recognize levels of mental workload, stress and emotion during each ATC task.

2 Related Work

Air traffic controllers (ATCOs) have to handle a significant amount of information that has to be interpreted and analyzed in a time critical manner. The increase in air traffic density is becoming a major issue in air traffic control. As reported by Sheridan [1] and International Civil Aviation Organization [2], worldwide traffic density will be up to double in 2025 compared to 2006. Given the limited airspace available, the possibility of having more air traffic conflict is unavoidable [3]. Under this circumstance, ATCOs will inevitably need a better support to overcome their cognitive limitations in handling more aircrafts in airspace. Providing support through automation is considered to be an effective solution to minimizing the workload imposed on ATCOs [4]. Endsley and Rodgers [5] discovered that when the air traffic increased, the controllers' awareness of each aircraft declined rapidly and when the workload was excessive, operational errors appeared. Recent research showed that conflict resolution aid (CRA) software has the potential to support ATCOs in resolving air traffic conflict in an effective and efficient manner [6] regardless of its imperfection [7] by advising ATCOs all the possible maneuvers. The future work place should therefore be designed to reduce mental workload and stress of ATCOs for optimal performance.

In research and development of human-machine interfaces, the evaluation of workload is a key point. Workload is described as a noticeable relationship between the human cognitive capacity and the effort required to process a particular task [8]. There are mainly three classifications for measurement of workload: subjective, physiological, and performance-based measures [9, 10]. Subjective measurement of levels of workload is based on the use of question-answer type response to measure the amount of workload a person feels during a task.

Currently, there are many subjective measure procedures designed to evaluate the mental workload as NASA Task Load Index (TLX) [11], Subjective Assessment Technique (SWAT) [12], and Cooper–Harper Scale [13]. NASA-TLX uses mental workload, physical demand, temporal demand, performance, effort, and frustration as six dimensions scales to evaluate mental workload. SWAT uses different three dimensional scales time load, mental workload, and psychological stress load as three discrete levels. But, Hill et al. [14] have proved that NASA-TLX is superior to SWAT in terms of measurement sensitivity especially for measurement of low workload.

Physical workload is the measurable portion of physical response of body when performing a given task and is affected by a range of factors. These physical responses include brain activity, cardiac activity, respiratory activity, and eye activity. Performance-based measurement of workload relies on examining some key parameters during a specific task which can reflect the capacity of a subject. In physiological measurement, electroencephalogram (EEG) interface is more suitable for monitoring people's mental workload because that EEG signals are directly captured from brain activity. EEG-based analyses have been widely used in clinical diagnosis of mental diseases and in bioengineering research. A number of EEG-based methods and corresponding applications are designed and implemented in order to recognize the user's workload levels [15, 16]. In [17], mental workload is evaluated in online EEG monitoring during the security surveillance task. Comparing the mental workload index with the error rate for the subjects, the correlation coefficient is approximately 0.7, which indicates that when the workload increases people have a tendency to make more errors. The correlation between workload and EEG signals has been proved in [18, 19]. In [18], the driver's mental workload is significantly correlated with theta band power and alpha band power. In different driving tasks, the frontal theta activity shows significant increases when working memory load increases. In another experiment studying the workload and fatigue in aircraft pilots [19], increased EEG theta band power and decreased alpha band power are observed in high mental workload comparing with the low mental workload. Additionally, in [19] it is shown that when the pilots have high mental workload and mental fatigue, their EEG theta band power as well as the delta and alpha bands power increases.

In this research, the EEG-based workload recognition, subjective user studies, and task performance are used together for evaluation of ATCOs' workload in different scenarios.

3 EEG-Based Workload Recognition

3.1 Feature Extraction

In our previous work [20], the real-time EEG-based brain states monitoring system CogniMeter is proposed to recognized emotion, workload, and stress. So, in this paper, we implemented the same algorithm for ATCOs' mental workload recognition based on FD and statistical features.

FD measures the complexity and irregularity of time series [21]. It can be used as an index for characterizing the complexities of EEG signals. For a regular signal, the fractal dimension value is low. If the signal becomes irregular, the fractal dimension value increases accordingly. Wang et al. [22] proposed to use Higuchi fractal dimension to recognize different arithmetic mental tasks from EEG. It is also used in EEG-based serious games to identify attention level. In this paper, the Higuchi algorithm is used to calculate FD feature for real-time workload recognition.

Statistical features are widely used in EEG based brain states recognition including emotion recognition algorithms [23]. Six statistical features such as mean, standard deviation, mean of absolute values of the first differences, mean of absolute values of the first differences of normalized signals, mean of absolute values of the second difference, and mean of the second differences of the normalized signals are extracted from EEG for emotion recognition.

3.2 Mental Workload Recognition

The mental workload recognition algorithm has been proposed in [24], in which the algorithm has been tested on the EEG database with different feature combinations and classifiers. For different feature combinations, the average accuracy of SVM classifier is 9.56 % higher than k-NN classifier based on mental workload EEG data. By combining statistical and FD features and using SVM classifier, the best accuracy is 90.39 % for 2 levels mental workload recognition and 80.09 % for 4 levels mental workload recognition. Therefore, in this experiment, we use FD and statistical features calculated from 14 channels and SVM classifier for mental workload recognition.

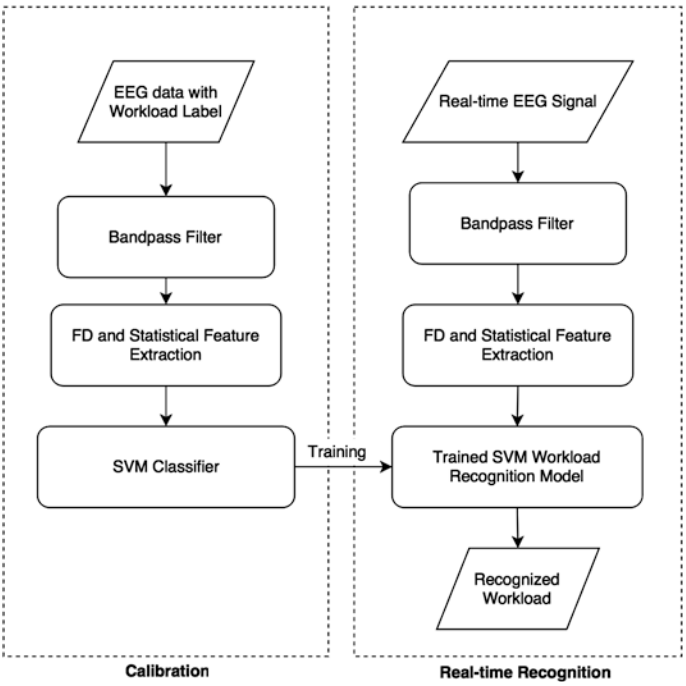


Fig. 1 The overall diagram of calibration and real-time brain states recognition algorithms for evaluation of mental workload [20]

The subject-dependent real-time workload recognition system consists of two parts: calibration and real-time mental workload recognition algorithm. The overall diagram of calibration and real-time workload recognition algorithms is shown in Fig. 2. In calibration, EEG data are labeled with different levels of mental workload for workload recognition correspondingly. Then, the EEG data are filtered, the corresponding features are extracted and the support vector machine (SVM) classifier is trained. After that, during real-time workload recognition, EEG signals are filtered and the FD and statistical features are extracted using a 4 s sliding window with 3 s overlapping. Next, new data features are input into the SVM classifier model trained in calibration. The classifier can recognize mental workload level based on each 4 s EEG signals input (Fig. 1).

4 Experiment

A preliminary experiment is designed and implemented to study the human factor in current ATC work place with some new features. 31 ATCOs and 5 students with ATC knowledge participated in the current user study and provided a signed-consent form that was approved by NTU IRB. All of them have received training of air traffic control and none of them has history of mental illness.

All participants were equally divided into three groups: Non-Display, Display, and Trajectory Prediction. Non-Display group was the baseline condition where participants were only equipped with the CRA. Participants in Display group were provided with the CRA and an additional display that depicted aircraft profile. In Trajectory Prediction group, participants were equipped with the CRA as well as an additional display that showed the prediction of aircraft trajectory including climb and descend rate information.

In every group, each participant performed ATC tasks in three CRA conditions: Manual, Reliable and Unreliable. In the reliable condition, the CRA was able to provide correct advisories to all the potential conflicts. In the unreliable condition, the maneuvering advisory provided an incorrect resolution advice that led to a conflict. In both reliable and unreliable CRA conditions, there was a conflict resolution advisory for each conflict and participants were free to either accept or reject the advisory by clicking a respective button. In the manual condition, participants were asked to resolve the potential conflicts by providing their own resolution maneuvering instructions.

In the experiment, there were three one-hour ATC scenarios corresponding to the three different CRA conditions. A balanced Latin square was adopted for the counterbalancing of CRA conditions to deal with any carry-over effects. In each scenario, participants were required to communicate with the pseudo-pilots to issue appropriate altitudes, to maintain separation between aircraft, to accept all aircraft that entered their sector, to hand-off aircraft that left their sector and to issue the correct radio frequency change.

4.1 Brain Computer Interface

In our experiment, Emotiv headset [25] is used to capture the users' EEG signals wirelessly with the USB receiver. It is a popular low-cost EEG device widely used for research including usability testing, neural marketing, serious games, etc. Emotiv EPOC has 14 channels located at AF3, F7, F3, FC5, T7, P7,O1, O2, P8, T8, FC6, F4, F8, and AF4 as shown in Fig. 2. During experiment, the EEG-based mental workload recognition system records EEG signals and recognizes ATCOs' workload in real time.

4.2 Workload Calibration and Recognition

As real-time EEG-based brain state recognition algorithms are subject dependent, calibration is required before real-time recognition. For calibration, four Stroop color-word test with different settings (congruent/incongruent ink colour or time limit) is used to induce different levels of workload. Each part of the test lasts for 1 min, and subject needs to fill a prompted questionnaire to evaluate his/her mental workload level on the scale from 1 to 9 and to describe his/her feelings in words as shown in Fig. 3. The calibration protocol using the Stroop color-word test is shown in Fig. 4. In the "Introduction" section, the subjects are briefly explained about the Stroop color-word test and get familiar with it; followed by the "Rest" section, which is used to record EEG data when the subjects are in the relaxed state. Then the subjects perform the Stroop test with three different levels and the self-assessment for each level is done at the end of each section as described above. To induce low workload, the word's meaning is the same with the word's font color (Congruent Section). To induce medium workload, the word's meaning is not the same with the word's font color (Incongruent Section 1). To increase workload to a higher level, the subject needs to react to the incongruent word within the limited time (Incongruent Section 2).

After calibration, the EEG data recorded during four tests were used for training of classifiers for each subject. Then, the EEG-based workload monitoring system

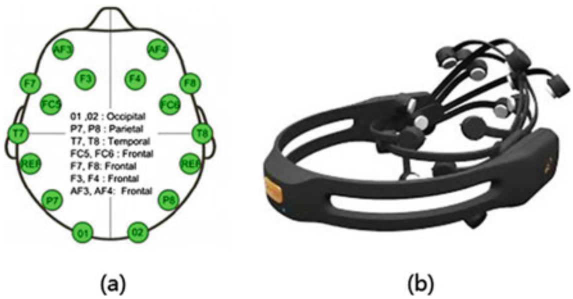


Fig. 2 The Emotiv brain computer interface. **a** The location map of 14 electrodes based on international 10–20 system. **b** Emotiv EPOC device records EEG signal at sampling rate 128 Hz with frequency response between 0.16 and 43 Hz

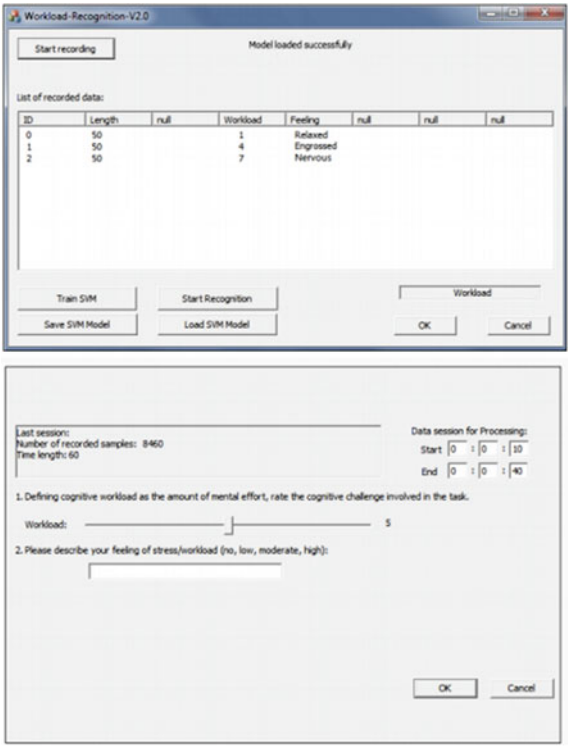


Fig. 3 Screenshots of the mental workload calibration interface and questionnaire interface

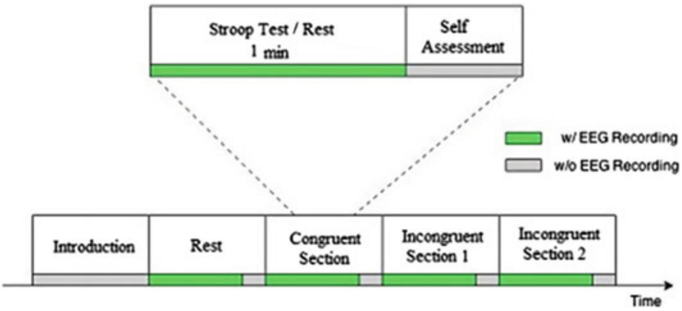


Fig. 4 The calibration protocol for EEG-based mental workload recognition

can recognize subject's mental workload each second. In Fig. 5, on the left, the subject's workload is visualized on the dynamic meter in real time. Besides color representation such as "red" color used for high workload and "green" color used for low workload, there is a word in the center of each meter to describe current workload level. After a completing of monitoring, a workload levels distribution

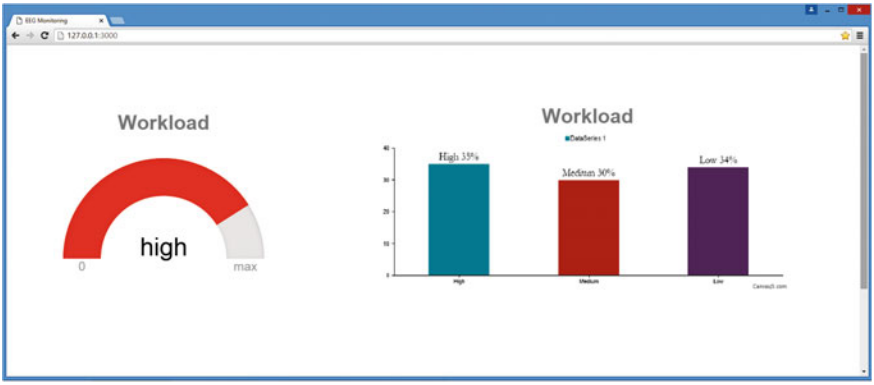


Fig. 5 Screenshot of EEG-based workload monitoring system. *Left meter* shows that current workload level is high. *Right diagram* shows overall distribution of different workload levels during the task performance

diagram can be generated to summarize the overall workload level during the task performance as it is shown in Fig. 5 on the right. This real-time workload monitoring system helps researcher to do more insightful analysis of subject's performance during ATC experiment.

4.3 ATC Simulation

In our current user study, we evaluate the current ATC work place and test interactive touch display that is the input of the development of the future ATC work place. The ATC simulator that used in our experiment is shown in Fig. 6. The middle monitor is used to display the primary radar informaiton. On the right side of the radar display, the monitor shows the Flight Progress Strips (FPS). FPS is an automation tool that provide aircraft updates including the latest altitude clearance, flight route as well as estimated outbound and inbound time for all departing and arriving aircraft, respectively. On the left side, a conflict resolution aid (CRA) display is intergrated to the current work place to support ATCOs in resolving conflict. The CRA is an automation aid that could advise ATCOs on the resolution of a potential conflict about 2 min in advance. In front of ATCOs, there is an interactive touch display to help ATCOs to understand the airspace situation. This display provides ATCOs with the information of aircraft speed profile, climb and descend rate along the time axes. During the experiment, the performance of percentage of resolved conflict and conflict resolution time are measured automatically through the data obtained from the simulator. Upon completion of each experiment scenario, the NASA-TLX questionnaire is used to measure mental workload of ATCOs. The EEG data are recorded throughout the experiment. The results of the

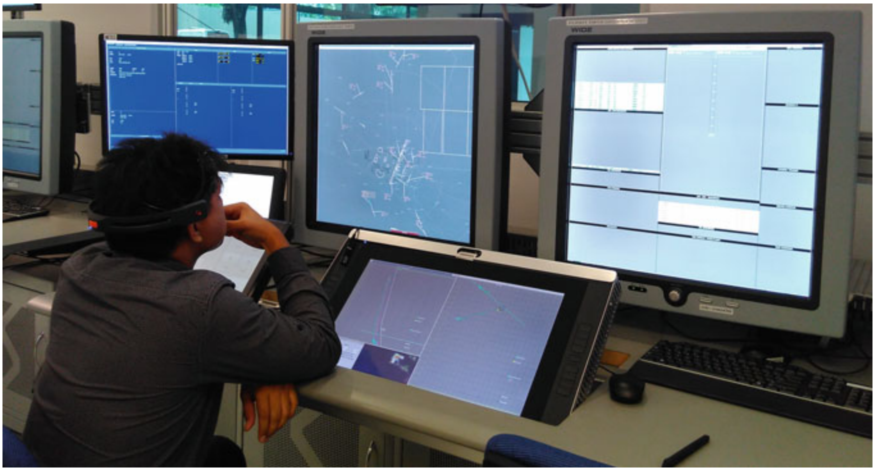


Fig. 6 The air traffice control work place integrated with interactive touch display and brain computer interface

study will drive the refinement and further development of both hardware configuration and software development.

5 Preliminary Results

Currently, we analyze the user study data of the ATC work place for three groups (Non-Display, Display, and Trajectory Prediction) in the three CRA conditions (Manual, Reliable and Unreliable). We studied the relation between the data received using traditional NASA-TLX method and the workload rating method used to label EEG data in the proposed EEG-based system for human factor study. Both methods were administered after each scenario. This analysis allowed for direct comparison between NASA-TLX and the proposed EEG-based evaluation system. Table 1 shows the correlation of workload rating received in 1–9 scale and NASA-TLX workload calculated after completion of each scenario of the experiment. Generally, the two evaluation methods were found to be highly correlated in most of the simulations. Only in unreliable CRA condition of non-display group and trajectory prediction group, the correlation between workload rating and NASA-TLX resulting data was not significant; however, the trend of positive correlation between the two methods' data in this condition could still be observed. These findings confirm that the method used for labeling of the EEG data with workload levels produces labels which are correlated with NASA-TLX workload evaluation data. Furthermore, with the reference to the labeling of the EEG data with workload levels, the EEG-based workload recognition algorithm can be used to calculate the workload levels in real time through all recorded EEG data with

Table 1 Correlation analysis between workload rating and NASA-TLX in the three groups (non-display, display, and trajectory prediction) and three CRA conditions (manual, reliable and unreliable)

	Manual	Reliable	Unreliable
Non-display	$r = 0.847$ $p = 0.001$*	$r = 0.748$ $p = 0.004$*	$r = 0.415$ $p = 0.180$
Display	$r = 0.661$ $p = 0.019$*	$r = 0.590$ $p = 0.043$*	$r = 0.716$ $p = 0.009$*
Trajectory Prediction	$r = 0.669$ $p = 0.017$*	$r = 0.529$ $p = 0.077$**	$r = 0.258$ $p = 0.419$

*Significant at $\alpha = 0.05$; **Significant at $\alpha = 0.1$

high time resolution. Thus, the EEG-based workload evaluation has been proven to validly assess workload and has a strong benefit since it could provide real-time workload data corresponding to the tasks performed throughout the experiment.

6 Conclusion

In this paper, we propose novel EEG-based tools for human factor study in Air Traffic Control (ATC) systems. The proposed system allows for recognition of mental workload, stress, emotions of subjects during the performance of ATC tasks to assess novel automation systems for future ATC. The EEG-based brain state recognition algorithms are implemented using machine learning techniques. We analyzed relation between mental workload calculated using traditional NASA-TLX method and the method used to label EEG data with different workload levels. It was found that the data are highly correlated in most of the simulations. Thus, the EEG-based system can be used to recognize workload during the task performance at any time. By utilizing the proposed EEG-based system, true understanding of ATCOs' working pattern can be obtained. Based on the analyses of the objective real time data together with the subjective feedback from ATCOs, we are able to reliably evaluate current ATC systems and refine new concepts of future ATC system.

Acknowledgments This research was supported by Civil Aviation Authority of Singapore (CAAS) and Air Traffic Management Research Institute (ATMRI) Project ATMRI: 2014-R5-CHEN.

References

1. Sheridan, T.B.: Next generation air transportation systems: human-automation interaction and organizational risks. In: Proceedings of the 2nd Symposium on Resilience Engineering, pp. 272–282 (2006)
2. Intensifying Asia-Pacific Collaboration to Address Efficiency and Safety.: International Civil Aviation Organization (2010)

3. Murphy, E.D., Albert, H.A., Chen, J.M., Anderson, G.G.: The role of mental computations in current and future en route air traffic control. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, pp. 110–114. Sage Publications, London (2012)
4. Endsley, M., Sollenberger, R., Stein, E.: The use of predictive displays for aiding controller situation awareness. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 43, pp. 51–55. SAGE Publications, London (1999)
5. Endsley, M.R., Rodgers, M.D.: Distribution of attention, situation awareness, and workload in a passive air traffic control task: implications for operational errors and automation. Federal Aviation Administration Office of Aerospace Medicine (1997)
6. Prevot, T., Homola, J.R., Martin, L.H., Mercer, J.S., Cabrall, C.D.: Toward automated air traffic control—investigating a fundamental paradigm shift in human/systems interaction. Int. J. Human-Comput. Inter. **28**, 77–98 (2012)
7. Trapsilawati, F., Qu, X., Wickens, C.D., Chen, C.-H.: Human factors assessment of conflict resolution aid reliability and time pressure in future air traffic control. Ergonomics 1–12 (2015)
8. Hart, S.G., Staveland, L.E.: Development of a multi-dimensional workload rating scale. Human Mental Workload 139–183 (1988)
9. De Waard, D., Studiecentrum, V.: The measurement of drivers' mental workload. Groningen University, Traffic Research Center Netherlands (1996)
10. Miller, S.: Workload measures. National Advanced Driving Simulator. Iowa City, United States (2001)
11. 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)
12. Reid, G.B., Nygren, T.E.: The subjective workload assessment technique: a scaling procedure for measuring mental workload. Adv. Psychol. **52**, 185–218 (1988)
13. Cooper, G.E., Harper Jr, R.P.: The use of pilot rating in the evaluation of aircraft handling qualities. DTIC Document (1969)
14. Hill, S.G., Iavecchia, H.P., Byers, J.C., Bittner, A.C., Zaklade, A.L., Christ, R.E.: Comparison of four subjective workload rating scales. Human Factors J Human Factors Ergon. Soc. **34**, 429–439 (1992)
15. Berka, C., Levendowski, D.J., Ramsey, C.K., Davis, G., Lumicao, M.N., Stanney, K., Reeves, L., Regli, S.H., Tremoulet, P.D., Stibler, K.: Evaluation of an EEG workload model in an Aegis simulation environment. Defense and security, pp. 90–99. International Society for Optics and Photonics (2005)
16. Kamzanova, A.T., Kustubayeva, A.M., Matthews, G.: Use of EEG workload indices for diagnostic monitoring of vigilance decrement. Human Factors J. Human Factors Ergon. Soc. **56**, 1136–1149 (2014)
17. Müller, K.-R., Tangermann, M., Dornhege, G., Krauledat, M., Curio, G., Blankertz, B.: Machine learning for real-time single-trial EEG-analysis: from brain–computer interfacing to mental state monitoring. J. Neurosci. Methods **167**, 82–90 (2008)
18. Lei, S., Roetting, M.: Influence of task combination on EEG spectrum modulation for driver workload estimation. Human Factors J. Human Factors Ergon. Soc. **53**, 168–179 (2011)
19. 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. **44**, 58–75 (2014)
20. Hou, X., Liu, Y., Sourina, O., Mueller-Wittig, W., Lim, W.L., Lan, Z., Wang, L.: CogniMeter: EEG-based brain states monitoring. In: LNCS Transactions on Computational Science (2016)
21. Higuchi, T.: Approach to an irregular time series on the basis of the fractal theory. Physica D **31**, 277–283 (1988)
22. Qiang, W., Sourina, O.: Real-time mental arithmetic task recognition from EEG signals. IEEE Trans. Neural Syst. Rehabil. Eng. **21**, 225–232 (2013)
23. Lan, Z., Sourina, O., Wang, L., Liu, Y.: Real-time EEG-based emotion monitoring using stable features. The Vis. Comput. 1–12 (2015)

24. Lim, W.L., Sourina, O., Wang, L., Liu, Y.: EEG-based mental workload recognition related to multitasking. In: 10th International Conference on Information, Communications and Signal Processing (ICICS) (2015)
25. The Emotiv EEG Device.: http://www.emotiv.com/epoc.php

Synergistic Allocation of Flight Expertise on the Flight Deck (SAFEdeck): A Design Concept to Combat Mode Confusion, Complacency, and Skill Loss in the Flight Deck

Paul Schutte, Kenneth Goodrich and Ralph Williams

Abstract This paper presents a new design and function allocation philosophy between pilots and automation that seeks to support the human in mitigating innate weaknesses (e.g., memory, vigilance) while enhancing their strengths (e.g., adaptability, resourcefulness). In this new allocation strategy, called Synergistic Allocation of Flight Expertise in the Flight Deck (SAFEdeck), the automation and the human provide complementary support and backup for each other. Automation is designed to be compliant with the practices of Crew Resource Management. The human takes a more active role in the normal operation of the aircraft without adversely increasing workload over the current automation paradigm. This designed involvement encourages the pilot to be engaged and ready to respond to unexpected situations. As such, the human may be less prone to error than the current automation paradigm.

Keywords Human factors · Human-systems integration · Function allocation · Flight deck · Aviation · Human error · Skill loss

---

P. Schutte (✉) · K. Goodrich
NASA Langley Research Center, Hampton, VA 23681, USA
e-mail: Paul.C.Schutte@nasa.gov

K. Goodrich
e-mail: Kenneth.H.Goodrich@nasa.gov

R. Williams
Analytical Mechanics Associates, Hampton, VA 23666, USA
e-mail: Ralph.A.Williams@nasa.gov

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_74

1 Introduction

The vision of a future flight deck for civilian and military aircraft is often one of no flight deck at all. Phrases such as ‘Increasing Automation’ and ‘Automation Autonomy’ dominate many research and development programs within NASA and the Department of Defense. The vision of uncrewed aircraft is one of increased efficiency, precision, and reliability and reduced costs and errors. Humans are often considered a liability to the system. The rationale for full automation is straightforward. Most accidents are found to be caused by human error. The human must be the weak link in the chain. Since automation has hardly ever been found at fault for causing an accident, more automation plus less human equals greater safety.

Unfortunately, this does not appear to be the case. The modern civil aircraft is highly automated and the human’s role has decreased substantially [1]. And yet, these aircraft are still involved in accidents. The fact that human errors are still named as the causes of these accidents gives rise to the question: Why hasn’t the reduction in human involvement in the operation of the aircraft resulted in a commensurate reduction in accidents caused by human error?

Similar questions have been raised regarding the issue of pilot workload in the flight deck. Since much of what the pilot used to perform has been allocated to the automation, why isn’t there a significant decrease in the pilot’s perceived workload? This question has been answered. Studies have shown that while the pilot’s physical workload decreased in highly automated aircraft, their mental workload increased. Automation did not reduce pilot workload; it simply changed the nature of that workload [2–4].

Perhaps the fact that human error and workload have not appreciably decreased in the modern flight deck is due, in part, to how the automation has been implemented as opposed to the amount of automation in the flight deck. There has been much research demonstrating how humans have fared poorly with increased automation; for example, automation complacency, overreliance on automation, loss of situational awareness and spatial orientation, and skill loss. These have contributed to human errors.

This paper presents a new flight deck design based on a function allocation approach called *Complemation* [5]. Complemation focuses on the role of the human in the flight deck; specifically, ‘Why *must* the human be in the flight deck?’ It uses automation and design to surround and support the human in performing that role. This is in contrast to substitution-based and machine-based forms of function allocation [6]. Substitution based function allocation considers all the tasks that have to be performed in the flight deck and determines whether the human or the automation can perform them better. The winner is given the task. Machine-based allocation operates under the assumption that machines are inherently better than humans and that the design should use automation to its fullest. In machine-based allocation schemes, the human is assigned the ‘leftovers,’ that is, the tasks that automation cannot handle. But these tasks are usually very difficult and can be disparate and non-cohesive from the human’s perspective. In Complemation, some tasks may be allocated to the human even though the automation may be able to perform some aspects of that task better than the human. Automation is used purposefully and deliberately as opposed to wherever it can be used.

The flight deck design produced by this approach is called the Synergistic Allocation of Flight Expertise in the Flight Deck (SAFEdeck). It has many commonalities with current flight deck designs and concepts but there are significant differences. One big difference is that the flight automation is controlled using the active feedback control inceptors (e.g., stick and throttle) rather than using an autopilot interface on the glare shield and a Flight Management System (FMS) Control and Display Unit (CDU). The inceptors are the only way in which the pilot can command the aircraft to move. The pilot is more actively involved in the progress of the flight. The automation supports the pilot by actively engaging and managing their attention so they do not forget to perform tasks and adhere to flight restrictions.

The first section of this paper sets the stage for describing SAFEdeck by providing an analogy using automobile driving. The next section constitutes the bulk of the paper where the SAFEdeck design is described. This section also concludes with a summary of results from an experiment that investigated this instantiation. The last section includes suggestions for future work and for implementation of the function allocation strategy.

2 An Automobile Analogy

Like aircraft operations, the SAFEdeck concept is extremely complex and detailed. As such, it is impossible to fully describe the concept in a paper such as this. In order to aid the reader in understanding SAFEdeck, an analogy may be beneficial. The analogy is based on automobile driving using advanced technology.

The Drive. A driver wants to travel from New York City to Orlando, Florida. The first step is to create a route using a route planner (e.g., Goggle Maps™, Garmin™, Tom Tom™). The planner uses published highway and roadway data to create the route. The driver can modify this route as needed. The route is loaded into the automobile's navigation system. The driver then begins the trip. Navigation information is provided on the driver's heads up display, as well as a top-down map display. Both displays depict the roadways and highways (even if they are not on the planned route).

The driver's first goal is to get to the interstate. There are several predetermined courses to get from the driver's house to the interstate (each course is a different packaged or chunked path out of the city). The driver turns the car onto one of these courses. The automation recognizes this as a predetermined route and offers the driver the option to drive this route. The driver accepts and engages the automation with a trigger switch on the steering wheel. From that point, the automation will drive the car. The driver can take hands off the wheel and gas pedal.

While still in the city, a van breaks down in the road in front of the car. The automation interprets this as a traffic backup and waits. The driver sees that it is not a traffic backup and disconnects the automation. The driver puts on the turn signal to inform other drivers and the automation that they want to drive into the oncoming traffic lane to go around the van. The automation's monitoring looks for cars approaching from either direction and gives the driver a green indicator when it is safe to do so. The driver drives the car around the van and returns to the road. The automation asks if the driver wishes to resume the departure from the city to the interstate, the driver says yes, pulls the trigger and the automation resumes driving the car to the interstate.

As the car approaches the interstate, it informs the driver that the entrance is ahead. It is the driver's responsibility to disengage the automation and manually drive the car onto the interstate. Once on the interstate, the driver can turn complete control back over to the automation. If the driver wishes to go faster, they accelerate to the new speed and couple to that speed. When the car approaches a change in interstates (e.g., leaving I-95 to get on I-495), the automation notifies the driver that the ramp is coming up. The driver then disconnects the automation and merges on to the new interstate and then reconnects the automation. If the driver fails to disconnect and take the exit. The automation will provide louder and more alarming alerts. The car will remain on the interstate. However the warnings will continue until the driver actively silences them.

If the driver wishes to stop for the night at a hotel, the driver can select the hotel on the map display. The automation will alert the driver when the car is approaching the exit ramp. Again, the driver manually takes control and transitions to the local road. If the hotel is much farther down the road, the driver can tell the automation to follow the road and provide reminders when approaching the point where they need to leave the road.

If the driver decides to simply drive around the countryside the next morning before returning to the interstate, the driver can drive on a road and then have the automation drive the car on that road. As the car approaches intersections, the automation alerts the driver of the intersection but nothing more. If the car comes to a T in the road where the driver must make a decision, the automation notifies the driver. If the car stops and the driver still has not intervened, the automation will sound the warning. If there is traffic behind the car, the automation will decide to turn one way or the other (to avoid obstructing traffic) but will pull over when able.

The Automation. Note that the car is not entirely self-driving. The driver does not enter a destination and then allow the car to independently drive all the way there. The car has limited automation, but that automation is extremely robust. The automation is responsible for all monitoring and for reminding the driver to make major transitions (e.g., home to city-exit course, city-exit course to interstate, interstate to interstate). But the driver must return to manual control to make those transitions. The automation will never willingly disconnect without the driver's approval. The automation can be overridden at any point.

The automation is aware of the road structure, the speed limits, facilities along the way, the weather conditions, and the plan. It is aware of its immediate surroundings (e.g., roads, other traffic, pedestrians and cyclists). It monitors road conditions. It is also aware of driver intervention and what the driver is doing (with regard to driving) even when the automation is not controlling the car. The logic of the automation has no high level reasoning skills and is entirely deterministic.

The Human. The human driver has the role of high level decision maker. If the driver is not situationally aware, they cannot perform this role and long durations of highly reliable automation can lead to complacency and distractibility. So the driver is called upon to periodically be part of the mission by making decisions at important junctures in the trip. It is unwise to expect that the driver will be paying attention otherwise.

The driver is also responsible for intervening in cases of automation failure or inability to appropriately perform. There may be cases where the automation doesn't know what to do or does not have authority. In these cases, the driver must return to manual control. If the driver has little or no regular experience with manual driving, there may be skill loss after a time. By requiring the driver to not only be involved in decision making at important junctures in the mission but also to manually drive, skill loss can be greatly reduced. Turns and decision points are more instructive in car handling skills than manual driving on a long stretch of highway.

3 Synergistic Allocation of Flight Expertise in the Flight Deck (SAFEdeck)

The SAFEdeck approach expands on the automobile system described above and applies it to aviation. The street/highway map is replaced with High and Low Altitude charts, Arrival, Departure, Approach and other terminal area charts and procedures. Nearly all of that information is contained in modern FMSs. The steering wheel is replaced by the active inceptors for the control surfaces and the gas pedal is replaced by an active throttle. The pilot has the role of high-level decision maker, risk manager and backup for the automation. Unlike today's aircraft, the entire mission may be planned, but execution of that mission requires human intervention at critical junctures. Unlike today's aircraft where there are three ways to control the aircraft (stick and throttle, autopilot/mode control panel (MCP), and FMS), there is only one way to control the aircraft and the automation—the active stick and throttle.

SAFEdeck is based on several design concepts, H-mode [7], the Naturalistic Flight Deck [8], and the Haptic Flight Control System [9]. The philosophy behind this approach is that one of the human's primary roles is to step in and deal with emergencies, non-normals, and highly complex or unanticipated situations. In some cases they must act as a backup for the automation or other resources. One of the goals of the design is to keep the pilot in the proper condition to perform these duties. To do this, SAFEdeck seeks to actively engage the pilot in the mission to maintain their situation awareness. This engagement will take the form of manually flying the aircraft at certain times in order to maintain skill level. In addition, the automation will conform to standard Crew Resource Management [10] principles as if it were another crew member. Finally, it is expected that the human will have deficiencies that lead to errors and the design must accommodate these deficiencies.

The SAFEdeck concept will be described by first defining the hardware components required, then the flight management functions and other functions critical to the concept. Finally, the results of an evaluation experiment will be briefly described.

3.1 SAFEdeck Hardware Components

There are many ways in which the SAFEdeck design can be implemented and so some of these descriptions will be deliberately vague. But the basic components either exist or are easily implemented in flight decks today. In many ways, the SAFEdeck flight deck will look very similar to modern flight decks. There are no dramatically new technology or display requirements. The main difference is how the automation and instrumentation is implemented, rather than the automation and instrumentation itself.

Active Inceptors. The two primary inceptors—the stick (or wheel and column) and the throttle—are active force-feedback inceptors. They are capable of transmitting haptic cues (such as pulses and vibrations) and are able to produce artificial force shaping to allow for resistance to envelope departures, detents, and other feedback signals. One force shaping feature is to allow virtual slots/tracks (similar to the slots/tracks in a standard transmission gear shifter pattern). Each inceptor has at least three switches: A trigger for engaging the automation, a button for disengaging the automation, and a selection device (e.g., thumb wheel, hat-switch). The inceptor position always corresponds to the actual commands given to the control surfaces and engines.

Navigation Display. A large, easily accessible, high definition map display that can not only show the path that the aircraft is flying and waypoints, but also existing route structure and available options. The map should have a top-down perspective and a vertical perspective. Touch or cursor control will likely be a requirement. The Navigation Display is a primary instrument for normal flight and will likely be consulted as often as the Primary Flight Display.

Primary Flight Display. This display contains the usual symbology found on modern primary flight displays. In addition a perspective view is presented behind the symbology. On this perspective view, not only is the current path portrayed in something like a highway in the sky, but also existing route structure. Waypoints in the form of ‘waypoles’ are also presented. Waypoles are vertical representations of waypoints. These waypoles may be flat earth representations if they are significantly far away.

Target Control Panel. Similar in many respects to modern autoflight interfaces (e.g., Mode Control Panels), the Target Control Panel allows the pilot to dial in specific headings, altitudes, airspeeds, and ascent/descent profiles. The major difference is that manipulating these parameters will not affect the aircraft's flight path. They merely create targets for the pilot to aim at/fly to. There may be additional parameter controls such as time of arrival or latitude/longitude.

Flight Planner. The Flight Planner is a separate device that is used to create plans, create what-if and alternate scenarios, and perhaps simulate the mission in fast time. The Flight Planner may be a portable device that communicates with the automation so that the pilot can make plans prior to flight or can use it as a tablet in flight, but this is not required. Ideally, the Navigation Display should not be used for the Flight Planner. The Navigation Display should always present real-time tactical information and it should not have additional clutter involved with the Flight Planner interface needs.

3.2 SAFEdeck Flight Management Functions

There are five basic Flight Management Functions in the SAFEdeck concept: Envelope Protection, Collision/Danger Avoidance, Self Preservation, Precision Assistance, and Active Flight Control. The first three are always on, however they can be overridden by the pilot. The last two assist the pilot in short-term, tactical maneuvers and control the aircraft over longer periods of time to manage workload and improve efficiency. They are used at the pilot's discretion.

Envelope Protection. This function impedes the ability to stall, overspeed, underspeed or barrel-roll the aircraft. When the aircraft is approaching one of these states, the pilot is alerted prior to this protection engaging. Before the aircraft actually enters one of these states, the inceptor will exert an artificial force that counteracts the condition. If the pilot does not intervene, this force will automatically return the aircraft to a safe orientation. The pilot can overpower this force and use the full capabilities of the aircraft. If the pilot releases the inceptor, the automation will seek to stabilize the aircraft.

Collision/Danger Avoidance. This function serves to automatically avoid dangers such as other aircraft, severe weather, terrain, or restricted airspace. The pilot will be alerted as soon as possible so that they can avoid the danger themselves as they see fit. If the aircraft continues to advance towards the danger, the automation will increase the level of alert and it will provide artificial counter pressure on the inceptor to move away from the danger. Again, the pilot can override this feature.

Self Preservation. This function is used when the pilot is not responding due to incapacitation or impairment. If the pilot has not responded to an alert or failed to take control of the aircraft when the automation requests it, the automation will enter into self-preservation mode. In this mode, the automation emits an emergency transponder signal, and air traffic control (ATC) and all aircraft in the area are given notice that the automation is taking control of the aircraft on its own. The automation will then plan a route to the nearest acceptable airport and proceed to perform an automated emergency landing. ATC and the other aircraft are responsible for clearing the way for this aircraft as they would in any emergency.

As this is one of the few times the automation will make a mode change on its own and because it takes control away from the pilot, every effort will be made to ensure that the pilot can override this mode. It may be that the automation's first task is to descend to a breathable altitude in case the pilot is hypoxic. Additional safeguards may be necessary for this mode such as concurrence by ATC.

Precision Assistance. There are three aspects of this function. They may be thought of as a 'snap-to' feature, a 'restrict axis' feature, and a 'reset' feature.

Snap-To. If the aircraft is approaching a target (such as a heading that has been selected on the target control panel, or a published flight path or waypoint), and the pilot performs an action that appears to be trying to lock on to that target, the automation will home in on it so that the pilot does not have to struggle to make the precise corrections. The automation will stabilize on that target. It is important to note that the automation will not hold that target under precision assistance. If the pilot wants to hold to that target, they must use *active flight control*.

Restrict Axis. Often the pilot may wish to make a turn without changing altitude or may want to make an altitude change while staying on the lateral route (e.g., a jetway). The stick inceptor has a slight artificial force shaping in the form of two virtual slots forming a cross with the center of the cross located at the current position. If the pilot pulls directly back on the stick, they can feel the vertical slot created by the artificial force shaping. This will restrict movements to the vertical dimension only and will hold the lateral path constant. Likewise, if the pilot moves the stick to the left into horizontal slot, only lateral changes are made and the aircraft remains at the same altitude. It is important to note (yet difficult to describe) that the force shaping conforms to the actual pattern that would be required to maintain either axis. For example, when turning to the left, one might have to raise the nose slightly (pulling back on the stick) to maintain altitude. The slot will then bend slightly back to add this correction. This is due to a SAFEdeck constraint that the stick position always reflects what is happening in the aircraft. By shaping the alleyway to reflect real flight control corrections, the pilot always feels what should be done to maintain that axis. That way, if the precision assistance automation fails, the pilot will still be making the same stick movements that would be made by the automation. The pilot will never have to move the stick in one manner while the automation is active and another manner while the automation is disengaged.

Reset. This function is used to return the aircraft to a stabilized straight and level configuration. If the pilot finds themselves losing control of the aircraft, the pilot can call upon this function to have the automation right the aircraft. The pilot would use this feature if they become spatially disorientated. One possible implementation for this feature would be for the pilot to press and continue to hold the trigger on the stick. This appears congruent with the expected human physical response to disorientation—to grip the inceptor tightly. When in the reset mode, the stick inputs of the pilot are ignored.[1]

Active Flight Control. In the automobile analogy offered in Sect. 2, the automation would couple to a road or highway and follow it without driver intervention. The SAFEdeck equivalent of this would be to couple to a jetway. The pilot would fly to the jetway, the automation would recognize it as something to follow, and the pilot would tell the automation to follow it. The airspace system is significantly more complicated than a country's road and highway system. SAFEdeck uses a category of objects called behaviors to handle the diversity of air travel. A jetway is a behavior. A holding pattern is a behavior. An approach is a behavior. A performance climb is a behavior. A takeoff is a behavior. A go-around is a behavior. A heading hold is a behavior. Behaviors are actions or sets of actions that the automation can perform autonomously. Behaviors can be published (e.g., jetways, approaches) or they can be created (e.g., heading hold, holding pattern, performance climb). In the automobile analogy above, the predetermined route out of the city would be a single behavior and is equivalent to a standard instrument departure from an airport). Behaviors are generally geographically—based and have a start and an end—however they do not have to be (e.g., holding patterns can occur anywhere above a certain altitude and continue until the pilot decides to leave the pattern). The pilot can join a behavior at any point along its three dimensional path.

To couple to any behavior, the pilot performs the following:

- Fly the aircraft to the behavior and align it to the behavior.
- Select the behavior (there may be more than one available at that location).
- Pull the trigger and engage the automation.

Align, Select, Trigger is all the pilot has to remember to couple to a behavior. Precision assistance aids the pilot in aligning the aircraft to the behavior. In some cases it may be possible to create a behavior when you select it. For example, when the pilot points the aircraft at a waypoint, the automation gives the pilot the option to create a 'go-to' behavior to that waypoint. Pulling the trigger creates that behavior and couples the automation to it.

Disengaging the Automation. A dedicated button on the inceptors is used to disengage the automation. This is the preferred method. When the button is used to disengage an audible notification is given that a normal disengagement has taken place. Pressing the disengage button is the equivalent of telling the automation, "I have control of the airplane." The audible notification is the automation's way of saying that it concurs. Another way to disengage the automation is by force. If the pilot grabs the stick and provides a reasonable amount of pressure, the automation will disconnect and a caution alert will sound. This indicates that the automation

[1]This is one of the few exceptions where the stick position may not agree with the actual control surface commands.

was disengaged in a non-normal manner. Unless there is a failure, the automation will not automatically disengage. Following good Crew Resource Management [10] principles, it will not relinquish control until there is someone to receive it, thus ensuring that someone always has control of the aircraft.

Modifying a behavior. It is common for an aircraft to change altitude on a route due to weather or traffic. In such cases, the Restrict Axis precision assistance is used. Moving the stick either directly backwards or directly forwards so that it 'slides' into the artificial force alleyway will cause the automation to stay on the current lateral track of the behavior. Dialing an altitude into the Target Control Panel will create an altitude target. When the aircraft is approaching this altitude, the pilot receives a notification that it will be time to level off. When the pilot levels off near that altitude, the Snap-To precision assistance will home in on that altitude.[2]

At the Behavior's end. If the aircraft comes to the end of the behavior and the pilot has not transitioned to a new behavior or disconnected the automation, the automation will go into a *safe state*. A safe state varies depending on the type of behavior and if there is a plan in the system. At the end of an airway behavior, if there is a planned transition to another airway and the pilot has done nothing, the automation will make that transition on its own. However, this will constitute a warning that requires significant pilot input to silence. The significant input is to keep the pilot from becoming reliant on this feature (e.g., not bothering to disconnect, align, select, trigger because the automation will do it for them). If the pilot does not respond for the next behavior transition listed in the plan, the self-preservation mechanisms described above will engage. If there is no planned transition but there are published behaviors connected to the end of the current behavior, the automation will make an educated guess and pick one. If there are no planned or published transitions, the automation will transition to an altitude/heading hold behavior.

3.3 SAFEdeck Notification and Alerting Functions

SAFEdeck requires that the pilot must have a more interactive role in flying the aircraft and this includes making time—and position-critical inputs such as leveling off at the proper altitude. However vigilance and prospective memory (i.e., remembering to do something) are weak traits in human behavior [11]. It is vitally important that the SAFEdeck design includes a robust notification system that will ensure that the pilot remembers to intervene. Fortunately, vigilance and prospective memory are automation's strong suits. The SAFEdeck automation can provide notifications for:

[2]There is more detail to correctly achieving this procedure.

- A behavior transition or parameter target is coming up (allowing the pilot to stop what they are doing and get back into the loop),
- It is time to make the behavior transition (e.g., disconnect, align, select, trigger),
- The pilot has failed to make the behavior transition, or
- The behavior has ended and the automation has gone into a safe state.

These notifications increase in urgency, saliency, and alert level (Advisory, Caution, Warning). The goal is that the pilot will respond to the first two notifications in order to avoid the last two alerts.

3.4 SAFEdeck Filtering and Decluttering Functions

As mentioned previously, all published behaviors are presented on the navigation display. This is important because it allows the pilot to easily transition to an unplanned change/behavior by essentially flying to it and coupling the automation to the behavior instead of having to program the changes in a flight management system. However, there are far too many published airways, waypoints, arrivals, departures, and other procedures to present all of them on the navigation display. Another critical element of the SAFEdeck design is robust and efficient contextual behavior filtering and display decluttering. These functions would use context such as aircraft equipage, current plan, current altitude, phase of flight, direction, range, current airport information, and perhaps probability to filter out a significant amount of choices. Of course, the pilot should be able to select the filtering/decluttering methods so as to see more or fewer choices.

3.5 Experiment Results

The SAFEdeck concept increases the amount of physical workload on the pilot by bringing them back into the loop at junctures in the mission. The primary controls are now all routed through the 'stick and throttle' inceptors. These two factors have led to speculation that this is a step backwards to the flight decks of old rather than a step forward into the future. A part-task simulator study was performed to assess impact of the SAFEdeck concept on workload when compared to manual flying and fully automated flying. Additionally, the impacts on situational awareness, primary and secondary task performance, and subject preferences were assessed [12, 13].

Twenty-four high-time, non-instrument-rated pilots planned and flew four different flights in a fictitious airspace using a moderate-fidelity, part-task simulation. Each of the four runs was approximately one hour long. Three different flight control paradigms were tested: Manual Control (MC), Full Automation (FA)—a path-coupled automatic control typical of modern commercial aircraft, and a simplified version of the SAFEdeck concept. Subjects were required to make both tactical and strategic flight changes as well as perform two secondary tasks (target recognition and numeric calculation). An automation failure was introduced in the FA and the SAFEdeck conditions and the time to detect the failure was measured. Workload was measured using the NASA-TLX [14]. Situational awareness was measured using the SAGAT [15] protocol and subjective responses.

To summarize the statistically significant findings: (1) the SAFEdeck condition reduced Mental Demand and Effort when compared to the MC condition; (2) subjects detected a failure of the automation in the SAFEdeck condition sooner than they detected it in the FA condition; and (3) subjects preferred the SAFEdeck condition over both the FA and the MC conditions when considering just flying the aircraft and when considering flying the aircraft with secondary tasks.

The statistically significant results themselves are encouraging and they reinforce the claims of increased situation awareness, reduced workload, and high subject preference when using the SAFEdeck concept. While many of the results were not statistically significant, they all favored the SAFEdeck concept over the other two.

4 Summary

SAFEdeck has six major features that make it unique from current automation strategies in the flight deck. The first combats mode confusion and skill loss and supports graceful degradation. SAFEdeck uses the manual control inceptors to manage and direct the automation rather than having the pilot use three uniquely different interfaces the control inceptors, the autopilot (mode control panel) and the flight management system interface. The second feature addresses mode confusion and complacency while improving situational awareness. The pilot is involved in all major trajectory changes such as major heading and altitude changes. The third feature combats typical human errors that stem from forgetfulness. SAFEdeck takes advantage of automation's memory capacity and retrieval (retrospective, prospective, declarative, and procedural) to backup the pilot. A fourth feature addresses mode confusion as well as allowing more fluid tactical trajectory management support. It is the use of enhanced graphics on both the primary flight display and the map display that show all flight path options that are available and appropriate to the pilot. The fifth feature is the use of the automation as backup for human error, and the pilot as backup for the automation. This feature addresses problems of complacency and other types of human error. Finally, complacency and fatigue are addressed by imbuing the automation with 'self preservation' features that make it 'resistant but not insubordinate' to making blunders such as flight into terrain or continuing an unstable approach below a safe altitude.

SAFEdeck does not sacrifice efficiency or capability. The human automation team is as efficient and precise as today's flight management system/autopilot combination. All functions performed in the current automation scheme can be performed using the new paradigm. But it is much more natural and simplified to perform those functions using the SAFEdeck design approach. The SAFEdeck design can be implemented using today's technology and does not rely on advances in artificial intelligence, access to big data, or changes in the airspace system.

The next steps in research are to fully implement the SAFEdeck design and then perform usability studies on the design using pilots and non-pilots.

References

1. Aviation Safety Network (2012): http://aviation-safety.net/statistics/period/stats.php?cat=A1
2. Hart, S.G., Sheridan, T.B.: Pilot workload, performance, and aircraft control automation. In: Human Factors Considerations in High Performance Aircraft (AGARD-CP-371), pp. 18/1–18/12. Neuilly Sur Seine, France: NATO-AGARD (1984)
3. Wiener, E.L.: Cockpit automation. In: Wiener, E.L., Nagel, D.C. (eds.) Human Factors in Aviation, pp. 433–461. Academic Press, San Diego (1988)
4. Wiener, E.L.: Life in the second decade of the glass cockpit. In: Jensen, R.S., Neumeister, D. (eds.) Proceedings of the Seventh International Symposium on Aviation Psychology, pp. 1–11. Department of Aviation, The Ohio State University, Columbus, OH (1993)
5. Schutte, P.: Complemation: an alternative to automation. J. Inf. Technol. Impact **1**(3), 113–118 (1999)
6. Cook, C.A., et al.: Function allocation for future systems: developing an improved methodology. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 40(19). SAGE Publications (1996)
7. 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. Technical Memorandum No. NASA/TM—2003-212672. NASA Langley Research Center, Hampton, VA (2003)
8. Schutte, P.C., Goodrich, K.H., et al.: The Naturalistic Flight Deck System: An Integrated System Concept for Improved Single-Pilot Operations, NASA/TM-2007-21509. National Aeronautics and Space Administration, Washington, DC (2007)
9. Goodrich, K.H., Schutte, P.C., Williams, R.A.: Haptic-multimodal flight control system update. In: 11th AIAA Aviation Technology, Integration, and Operations Conference (ATIO), Virginia Beach, VA (2011)
10. Kanki, B., Helmreich, R., Anca, J. (eds.): Crew Resource Management. Academic Press, San Diego, CA (2010)
11. Dismukes, R.K., Nowinski, J.L.: Prospective memory, concurrent task management, and pilot error. In: Attention: From Theory to Practice, pp. 225–236 (2006)
12. Schutte, P.C., Goodrich, K.H., Williams, R.A.: Towards an improved pilot-vehicle interface for highly automated aircraft: evaluation of the haptic flight control system. In: Proceedings of 4th International Conference on Human Factors and Ergonomics, July 2012
13. Schutte, P.C., Goodrich, K.H., Williams, R.A.: Evaluation of the Haptic Flight Control System: A Common Interface for Manual and Automatic Control Modes
14. Hart, S., Staveland, L.E.: Human Mental Workload, Development of NASA TLX: Results of Empirical and Theoretical Research. North Holland Press, Amsterdam (1988)
15. Endsley, M.R.: Situation awareness global assessment technique (SAGAT). Proceedings of the National Aerospace and Electronics Conference (NAECON), pp. 789–795. IEEE, New York (1988)

Developing a Mental Model in ATC: 1—Learning Situational Assessment

Richard W. Rohde

Abstract This is the first in a series of papers on the mental process of the En Route 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

The concept of a 'Situational Awareness' in Air Traffic Control has been around since at least the 90s in academia [1], and much longer than that in the field, where it is better known as 'the picture' [2] or 'the flick'.

Some early papers used 'Mental Model' for what now is generally considered 'Situational Awareness' [3]. Over time, however, the two terms were recognized as being different [4, 5], though the precise definitions for both terms are still somewhat nebulous.

This paper will first attempt to provide a working definition and in-depth explanation of both the 'Mental Model' and 'Situational Awareness' (SA) as they apply to En Route Air Traffic Control from a controller's point of view. It will then explore how a newly hired radar trainee learns SA. A second paper will explore how Situational Awareness is attained in an active air traffic environment.

A TRACON or an approach controller's Mental Model and Situational Awareness will be similar to an En Route controller. Tower controllers, because they are in an environment where they must focus their attention several different

R.W. Rohde (✉)
Fort Hill Group, 660 Pennsylvania Ave SE, Suite 204, Washington, DC 20003, USA
e-mail: Rory.Rohde@FortHillGroup.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_75

places to do their job and because they work as a group, have different systems of Mental Models and Situational Awareness [6].

More narrowly defining these terms is important because they directly or indirectly factor into the concept of ‘workload’, another important and ill-defined concept that is at the crux of both ATC studies and NextGen goals.

2 Basic Definitions

Situational Awareness 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” [7].

For the Controller, Situational Awareness is the 3-dimensional picture of the sector and traffic that he has constructed in his brain. He uses Situational Awareness to assess the traffic situation, project into the future, and make control decisions accordingly. Situational Awareness is ‘working memory’ combined with the Mental Model and resides in short-term memory. Note that Situational Assessment is not the same as Situational Awareness but instead is the process of attaining Situational Awareness.

The Mental Model is the controller’s overall knowledge of everything that might affect his or her air traffic situation. This includes applicable knowledge from the mental libraries of long-term memory and relatively current information such as weather, Traffic Management Initiatives (TMIs), the controller’s assessment of their own mental state, etc.

> In the case of ATC, there appear to be two components of the controller’s mental model. The first is a Domain Model that encompasses airspace, aircraft characteristics, and ATC procedures. The second factor is a Device Model, which is an understanding of the electronic systems (including the computer-human interface) designed to support ATC. Both kinds of knowledge are essential if the air traffic controller is to accomplish the task of separating and guiding aircraft. This is analogous to the need to know some geography as well as automobile operation to arrive successfully at a destination. [8]

A simplified way of differentiating the two terms is to say that the Mental Model is based on Long-Term memory, the mental library, while SA is based on short-term or working memory.

It is easy to see why these terms can be confusing. The term ‘mental model’ could easily (and arguably more accurately) have been used to fit the definition that has been given to SA. It is almost impossible to describe Situational Awareness without using the term ‘mental’. EATCHIP, for example, refers to SA as Mental Picture (MP).

3 Learning Situational Awareness: Non-radar

Prospective En Route controllers go through several weeks of non-Radar training at the FAA Academy in Oklahoma City. Among other things, this pass/fail course requires students to memorize a map of 'Aerocenter' (Fig. 1), a generic low altitude sector.

The Aerocenter map, which the student must be able to draw in its entirety from memory, includes VORs, airports, airways, mileages, radials, intersections, adjacent sectors and centers, approach control airspace, frequencies, minimum altitudes, and more. The student will also familiarize themselves thoroughly with Aerocenter Letters of Agreement (LOAs) and Standard Operating Procedures (SOPs). This is the foundation of the mental domain model that the trainee will use to move traffic safely and efficiently in non-Radar simulation scenarios.

During several weeks of classroom training, students will add to their domain models through classroom instruction on "the rules of Air Traffic Control" from FAA order 7110.65 (the 'ATP'). They will learn proper phraseology, non-radar separation rules, and other applicable information. After successful completion of this phase, they move on to the Manual simulation lab.

For a 'Manual' (Non-Radar) problem, the trainee will sit down at the sector where they will be presented with a number of strips in two main 'active' bays

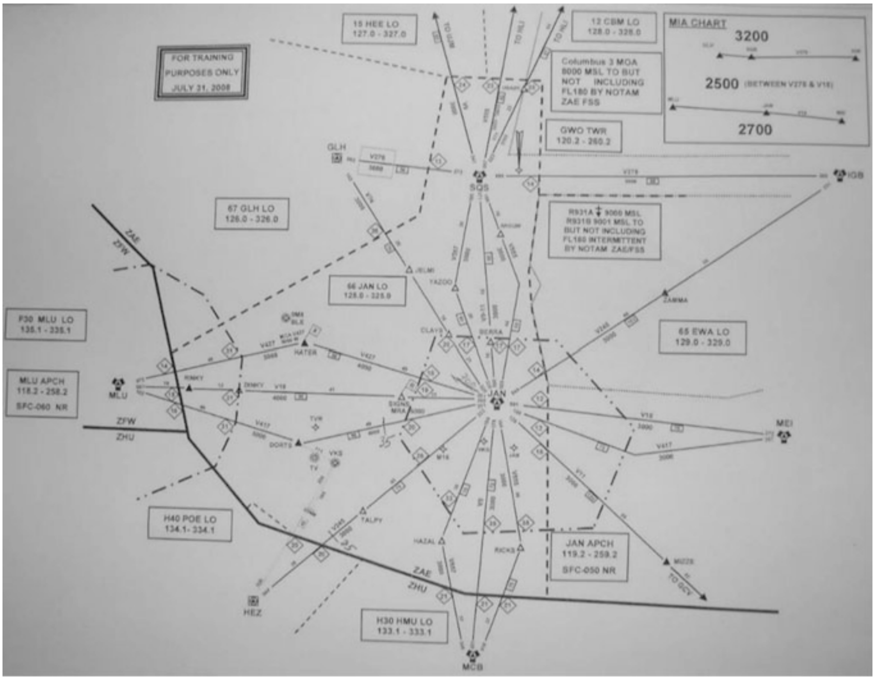


Fig. 1 Aerocenter

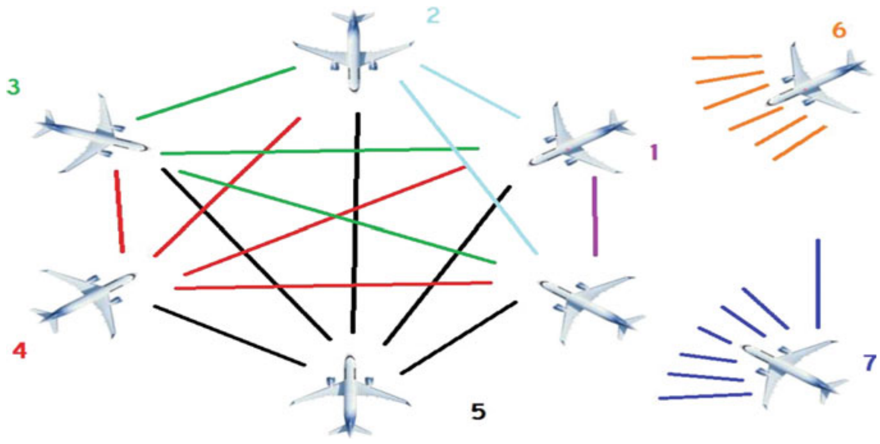


Fig. 2 Potential conflicts

(based on geographical location) and a few more in the 'suspense' or proposal bay. Aircraft in the suspense bay will depart from airports within the sector during the course of the scenario.

As the trainee progresses through the various scenarios, they will augment their Mental Model. The 'numbers' memorized on the Aerocenter map will develop meaning as the trainee learns to apply them to accomplish the tasks presented. Frequencies will become second nature. They will also learn the 'personality of the sector', including confliction points, shortcuts, tricks, traps, and other information that can be filed away for future use.

With each scenario, the total number of aircraft will gradually increase. Because each added aircraft must be checked with all proceeding aircraft, complexity grows quickly in a 'triangle number' progression. For example, a scenario with six aircraft will have 15 potential conflicts, while adding two more aircraft almost doubles the number of potential conflicts (Fig. 2).

In Aerocenter, the two main active bays are JAN (Jackson, MS) and SQS (Sidon, MS), the two VORs[1] in the sector (Fig. 3). These are also the two main crossing/confliction points (circled in red in Fig. 3). Aircraft traveling East/West over one of these VORs will have one strip in the respective bays, while aircraft traveling North/South over both will have two strips, one in each bay.[2]

The trainee is given several minutes to 'pre-plan'. During this time the student is may mark the strips using a red pen. For example, they may add direction arrows as a quick reference to which way the aircraft is going (over time the student will

[1]VORs are the basic navigational aids that define airways.

[2]There is a third "active" bay, VQS, which is used for traffic in and out of Vicksburg (VQS) and Byerly (0M8), but this bay is generally not a factor for separation as described in this paper.

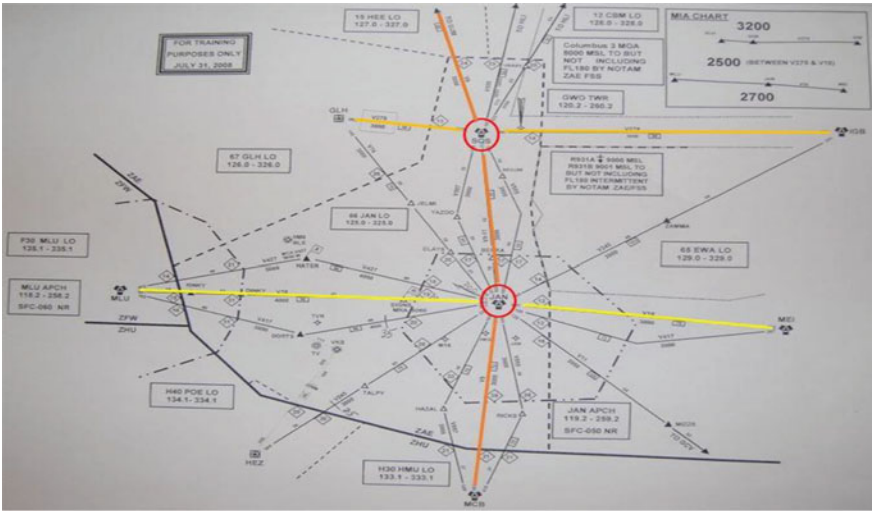


Fig. 3 Confliction points

become adept at reading the strips and will most likely only use direction arrows for unusual routes). It is during this time that the trainee builds their 'picture', combining their mental model with the information on the strips to achieve SA.

4 Separation

The controller will first look for separation issues. Of the three forms of separation-lateral, vertical, and longitudinal-lateral (geographical) is the surest. However, the strip bays already account for this. Aircraft in different strip bays are geographically separated.

Vertical (altitude) separation is the second best form of separation. The trainee will scan the altitudes on the strips in each bay and look for 'pairs'. It cannot be stressed enough that altitude is the primary way to classify aircraft within a strip bay. For separation purposes, aircraft at different altitudes are on different planets.

When a 'pair' is found, the student will further examine the strips to see if there is a potential conflict. Two aircraft with a common fix-posting will either be on the same route following each other (in-trail), or crossing each other's routes.

In-trail aircraft will need to have longitudinal separation ensured while aircraft on crossing routes will need to have lateral separation. Aircraft should not be head-on because that would mean one of them would be IAFDOF—Inappropriate Altitude for Direction of Flight, something else the student will scan for during the pre-planning period.

Because this is a non-radar environment, the exact position is not readily available to the controller. All longitudinal and lateral separation will initially be accomplished using time and position reports obtained from the simulation pilots. Because of the uncertainty involved in non-radar, 10 min is the minimum standard separation.

The student will examine potential lateral and longitudinal conflicts for time separation. If the aircraft are following each other (in-trail), the student will also check the airspeeds to make sure the trailing aircraft is not overtaking the lead aircraft.

For example, let us assume the controller finds two strips in the SQS bay showing 160 (Sixteen thousand feet) in the altitude box (Fig. 4). Upon further examination, the student sees both are routed IGB.V278.GLH and the faster one is in front. These westbound aircraft are not in conflict (he may assume the imaginary "previous controller" will ensure traffic coming into the Aerocenter controller's sector are separated).

Now let us assume the controller then comes across a third aircraft in the SQS bay at 160. This is DAL7231, routed UJM.V9.MCB (Fig. 5). DAL2731 is southbound and will cross the paths of both of the FDX aircraft at 160. The student will next look at the times for each aircraft at SQS. If there is not the required 10' difference for non-radar separation, he will put a red 'W' in on both strips and offset them as a 'prospective memory' aid—a reminder that action will need to be taken to avoid a separation error.

DAL7231 is separated from FDX278 by the required 10' minimum, but is in direct conflict with FDX524. The trainee may 'preplan' an altitude change for one of the aircraft during the pre-planning phase, writing the planned altitude in red on the strip. The conflict occurs at 1612z, so it must be resolved 10' prior, or by 1602z in this case. Since this is only 2 min after the start time of the scenario, this will be one of the student's initial actions.

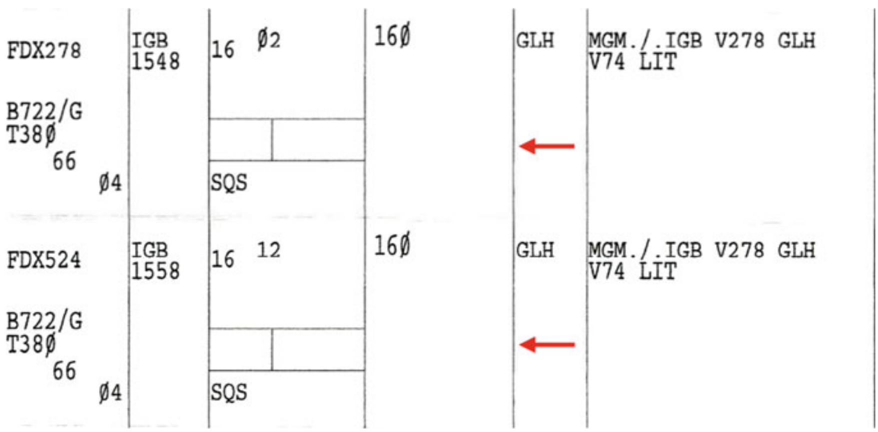


Fig. 4 Flight progress strips

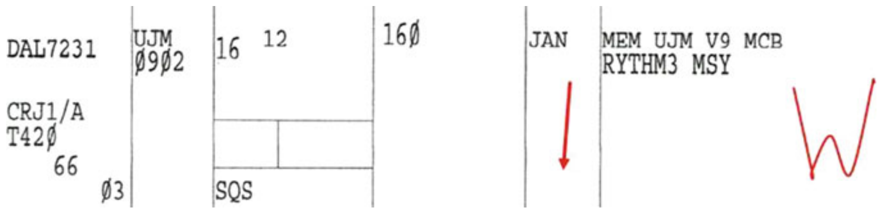


Fig. 5 Flight progress strip marking

After scanning for basic conflicts, the student will next look for active aircraft where other action needs to be taken. Most of the time this will be arrivals that need to be descended. For any aircraft that will need an altitude change to meet the SOPs or LOAs for Aerocenter, the controller will need to rescan the strips for all altitudes between the active aircraft's current altitude and the altitude it must ultimately be at the aircraft will be travelling through 'multiple worlds.'

If conflicts exist, clearances will be formulated that will ensure separation AND that incorporate other changes (such as the new altitude for either DAL2731 or FDX254) into his existing SA. He will then update the picture and continue with the active aircraft.

Once the trainee has completed this process, he will examine the proposal aircraft in the 'suspense' bay and work out clearances to safely get them to their requested altitude.

When the scenario begins, the trainee will be able to solicit pilot reports on aircraft positions and/or altitudes as necessary to use other methods of separation.

The successful student will have everything envisioned and all moves planned before the instructor 'starts the clock'. Then it is just a matter of exercising prospective memory and executing the plan.

This is all a simplified version. The student has several other non-radar rules at his or her disposal (the 44-knot rule, 'paper stops', etc.), and there are other factors and requirements to consider, but hopefully it conveys the general idea of not only how the non-radar SA works, but how En Route controllers learn to form mental models and use them to 'get the flick'.

5 Gaining SA in a Radar Simulation

> Situation awareness (SA) is considered the product of the process situation assessment that takes place at three levels: perception (SA1), interpretation (SA2) and anticipation (SA3). Attention management strategies are crucial to keep this ever changing 'picture' up-to-date. [4]

After four weeks of non-radar training, the students will move on to basic radar training. They will spend a few weeks in the classroom learning how the En Route

Automation (ERAM) works and then have five weeks of simulation training strictly as a D-side (Radar assistant controller).

Those that pass will be sent to an ARTCC where they will receive more D-side training, this time on simulations of one or two of the sectors from their assigned area of specialization. Following successful completion of this, they will get On-The-Job-Training (OJT) with live traffic. Only after “checking out” on the D-sides in their area of specialization and several months of “seasoning time” will they begin radar simulation training.

Using a radar scope provides much more information to the trainee. It also allows them to use significantly tighter separation standards, 5 miles instead of 10 min (which can be over 80 miles for a jet aircraft). However, it does not change the basic way he will work traffic. Until a little over a decade ago, the controller would usually look at the flight strips to begin to gain situational awareness and then ‘fine tune’ his picture on the radar. With the advent of URET, which has a much smaller footprint that a strip bay with a small display showing less than half of the information that was available on Flight Progress Strips, the student controller will go right to the radar to gain Situational Awareness.[3]

The student will apply the principles learned in non-radar training, scanning the traffic for aircraft at the same altitude, noting almost simultaneously if the routes will cross. When such pairs are found, the controller will examine them in detail to see if there is a potential conflict. If a potential conflict is detected, the controller will pre-plan how he wants to alleviate it. This process is what researchers commonly refer to as Trajectory Prediction (TP).

The trainee will then move on to scanning for aircraft that need ‘to have something done to them’ such as receiving a departure clearance and climb to their requested altitude or changing the altitude of other aircraft to meet LOA and SOP criteria. He will most likely look over the strip bay/URET to scan for potential conflicts and any unusual traffic such as IAFDOF, block altitudes, Non-RVSM, etc.

Once the scenario starts, SA is constantly ‘refreshed’ through scanning of the control environment including the radar scope, paper or electronic flight strips, and audio information.

6 Summary

Learning to attain Situational Awareness in Air Traffic Control is similar to many other environments. It takes a lot of practice to master. Mogford’s car analogy [8] is useful here. One first learns about driving a car while sitting in the passenger’s seat

[3]This assumes there is active traffic on the scope. Up until 2008, all ARTCC training was done on the DYSIM system, which did not allow for traffic to be “active” at the beginning of the scenario. EERTS and later ERAM simulations in addition to be being much more realistic than DYSIM allow for this capability. But since many training scenarios were just converted over from DYSIM it is still much more common to have trainees taking over a sector empty of aircraft.

and watching others drive. Once you reach a certain comfort level, the person teaching you how to drive will most likely take you to an empty parking lot or some other open space with little or no traffic so you can get used to actually driving, i.e. how to use the breaks, the accelerator, the steering wheel, and maybe even how to operate a manual transmission.

This paper has attempted to define the terms related to Situational Awareness and explain how developmental controllers learn the process of Situational Assessment. The purpose is to provide academics and researchers with insights into the inner mental working of Air Traffic Controllers in the hopes of further improving the quality of future research.

Returning to the car analogy once more, driving around in an open parking lot is very similar to controlling traffic under simulation conditions. Only after you have mastered these will you be ready to train with live traffic…which is completely different from the sterile environment of a simulation. That will be the focus of the next paper in this series.

Acknowledgments A special thank you to Patrick Long, ATCS at Seattle Center, for all his assistance, especially the accuracy of the description of Aerocenter.

References

1. Stein, E., Garland, D.: Air Traffic Controller Working Memory: Considerations in Air Traffic Control Tactical Operations. DOT-FAA Technical Center (1993, Sept)
2. Shorrock, S.: Errors of memory in air traffic control. Saf. Sci. **43**, 571–588 (2005)
3. 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. (1992, Jan)
4. 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. ATM2009
5. Wickens, C.: Situation awareness: review of Mica Endsley's 1995 articles on situation awareness theory and measurement. Hum. Factors 397–403 (2008, June)
6. Broach, D.: Selection of the Next Generation of Air Traffic Control Specialists: Aptitude Requirements for the Air Traffic Control Tower Cab in 2018. CAMI (2013, Mar)
7. Endsley, M.: Toward a theory of situation awareness in dynamic systems. Hum. Factors J. Hum. Factors Ergonomics Soc. 32–64 (1995, Mar)
8. Mogford, R., Guttman, J., Morrow, S., Kopardekar, P.: The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature. DOT/FAA (1995, July)

Sleep Environment Recommendations for Future Spaceflight Vehicles

Zachary A. Caddick, Kevin Gregory and Erin E. Flynn-Evans

Abstract Evidence from spaceflight and ground-based missions demonstrate that sleep loss and circadian desynchronization occur among astronauts, leading to reduced performance and, increased risk of injuries and accidents. We conducted a comprehensive literature review to determine the optimal sleep environment for lighting, temperature, airflow, humidity, comfort, noise, privacy and security in the sleep environment. We reviewed the design and use of sleep environments in a wide range of cohorts including among aquanauts, expeditioners, pilots, military personnel, and ship operators. We also reviewed sleep quality from every NASA spaceflight mission. We found that the optimal sleep environment is cool, dark, quiet, and is perceived as safe and private. There are wide individual differences in the preferred sleep environment; therefore modifiable sleeping compartments are necessary to ensure all crewmembers are able to select personalized configurations for optimal sleep. We provide recommendations to aid in the design of deep space sleep chambers.

Keywords Extreme environments · Habitability · Human factors · Sleep

1 Introduction

Sleep quality—including the ability to fall asleep and remain asleep—and sleep duration are dependent upon circadian phase, length of prior wake duration, and time within the sleep episode [1–3]. Proper alignment of scheduled sleep episodes to the circadian pacemaker is important for sleep consolidation and sleep structure [4, 5]. High sleep efficiency is best maintained for eight hours when sleep is initiated approximately six hours before the endogenous circadian minimum of core

Z.A. Caddick · K. Gregory
San Jose State University Research Foundation, San Jose, CA, USA

E.E. Flynn-Evans (✉)
NASA Ames Research Center, Moffett Field, CA, USA
e-mail: erin.e.flynn-evans@nasa.gov

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_76

body temperature [4, 5]. This phase relationship between the rest-activity cycle and the endogenous circadian timing system implies that even small circadian phase delays of the sleep propensity rhythm with respect to the rest-activity schedule can result in sleep onset insomnia or substantial wake after sleep onset.

In order to quantify the impact of a sub-optimal sleep environment on sleep quality and duration, it is important to measure sleep outcomes when sleep is appropriately timed relative to the circadian and homeostatic drives for sleep. It is possible for an individual to experience sleep disruption in an optimal sleep environment due to the imposed sleep schedule. Similarly, it is possible for an individual to experience high sleep efficiency in a sub-optimal sleep environment when accumulated sleep debt is present, which dampens the arousal threshold. Our aim was to compile the evidence associated with sleep disruption due to controllable, environmental stimuli in order to aid NASA engineers and operational personnel in the optimal design of crew sleep accommodations for deep spaceflight.

2 Methods

We conducted a comprehensive literature review summarizing optimal sleep hygiene parameters for lighting, temperature, airflow, humidity, comfort, intermittent and erratic sounds, privacy and security in the sleep environment. We reviewed the design and use of sleep environments in a wide range of cohorts including among aquanauts, expeditioners, pilots, military personnel and ship operators. We also reviewed the specifications and sleep quality data arising from every NASA spaceflight mission, beginning with Gemini.

3 Recommendations

The sleep environment required for long duration missions will differ from the sleep accommodations that NASA has developed in the past. Our review revealed several modifications that will be important to make in order to ensure that deep space crews have sleep environments that will provide them with quality sleep.

3.1 Sleep Chamber Location

The location of the sleep station within the vehicle is key to reducing noise and light pollution. Noise emanating from common areas has been shown to be disruptive to sleep [6, 7]. Given that there are individual differences in sleep timing preference, it is likely that some crew will chose to be awake, while others are asleep [8, 9]. In order to ensure that morning-types and evening-types are both afforded adequate rest, it is desirable to position crew quarters away from the galley area and exercise machinery. We also found that individuals living in a variety isolated and confined environments reported experiencing sleep disruption due to other crewmembers using the waste management system during sleep episodes [9–11]. Therefore, the waste management system should be located far enough away from sleeping quarters that noise is buffered, but close enough that crewmembers are able to quickly access the facility and return to sleep without having to travel too far. It may be appropriate to locate waste management facilities in a module adjacent to the sleep stations.

It is likely that watch schedules will be necessary during deep space missions. We found that in the early history of human spaceflight, watch schedules were very disruptive to sleeping crewmembers due to the close proximity of the sleeping crewmember to the "on watch" crewmember [12]. According to studies of military personnel and pilots, locating the sleep chambers for off-duty crewmembers away from the command and communication area is desirable [11, 13–15]. However, the sleep chambers should be positioned near enough to the vehicle command center that crewmembers may quickly respond in an emergency situation [11].

3.2 Privacy

It is imperative that each crewmember is provided with a private sleep chamber for the duration of the mission. We found that shared sleep spaces and common bunkrooms are associated with frequent sleep disruption due to other crewmembers [13]. The practice of "hot bunking" has been virtually eliminated from all occupations that we evaluated due to hygiene concerns and the impact that hot bunking has on psychological mood and health [13, 16]. We found that individuals view their sleep location not just as a place for sleep, but also as a space for privacy [7, 14, 16–25]. Access to a private space is viewed as critical to the psychological well-being of individuals living in isolated and confined environments [26]. Similarly, provision for storage of personal items within the sleep chamber was viewed as highly desirable [9, 27]. The sleep chambers for deep space vehicles should also allow crewmembers to customize the space with personal items and reconfiguration of stowage compartments [9].

There have been situations where crewmembers have been displaced from private quarters during spaceflight missions [28]. In these situations it is very difficult for the displaced individuals to obtain adequate sleep [29–31]. Given that the loss of a sleep chamber would likely also be associated with a breach of the spaceflight vehicle, the resulting anxiety may further reduce crewmember sleep quality and quantity. As a result, it is possible that the loss of a sleep chamber could greatly impact the physical and psychological health of crewmembers at a time when successful performance of duties is essential. Given the importance of sleep in conferring fitness for duty, future crew vehicles should include back up, deployable sleep chambers in order to ensure that individuals have access to a private sleep environment throughout the mission.

3.3 Habitable Volume

The crew quarters that are presently on ISS appear to provide enough habitable volume for crewmembers to move as desired during sleep [32, 33]. We found one case where a crewmember was too large to fit in the assigned sleep chamber during spaceflight [34]. Although it may be necessary to design all sleep chambers and sleeping bags to the same standard, it is important to consider that larger crewmembers will have less habitable volume relative to smaller crewmembers. As such, it is important to ensure that the crewmembers selected for a deep space mission are able to evaluate the size of the sleep stations in advance of the mission. It may also be desirable to design two sizes for the sleep stations to accommodate larger and smaller crewmembers.

The optimal sleep environment for a planetary excursion will be necessarily different from the optimal sleep environment for spaceflight. During a long duration planetary excursion, larger crew quarters are necessary due to the comparatively reduced habitable space available in a partial gravity environment. We found that individuals living in isolated and confined environments on Earth use their sleep rooms as a place to be alone and to work in addition to sleep [7, 27, 35]. As a result, the crew rooms on a planetary excursion should include space for a bed (placed horizontally on the floor), a desk and storage of personal belongings. The use of bunkrooms or shared sleep spaces is only appropriate for a short-duration planetary excursion. In these cases, bunks or cots may be used to accommodate crewmembers [7]; however, even during such short excursions private crew quarters would be preferable [27].

3.4 Light

Sleep chambers in spaceflight and on the ground must include features that protect individuals from being awoken by external forces such as light, noise, inadequate temperature and poor air quality. Light is the primary resetting cue for the human circadian pacemaker [36]. Exposure to light at inappropriate times leads to circadian misalignment, which causes sleep disruption [37]. Similarly, exposure to light is alerting and suppresses the drive to sleep [38]. The intensity, spectra, duration, and timing of light determine the magnitude and direction of phase shifting and potency of acute alerting [39]. All wavelengths of light have a negative impact on sleep, but blue light elicits the strongest effect due to the stimulation of intrinsically photosensitive retinal ganglion cells [38]. Exposure to green light is capable of enhancing alertness and suppressing sleep [38], while exposure to red light has the weakest effect on alertness and circadian phase shifting [40]. Evidence from the laboratory, field and subject matter experts support the notion that exposure to light during sleep episodes is disruptive to sleep quality and quantity [12–14, 29, 41–49]. Based on this evidence, all light should be eliminated from the sleep environment.

If indicator lights are necessary for identifying egress points, then they should be dim and red [40].

There is strong evidence to suggest that individuals living in isolated and confined environments away from typical solar light dark cues are prone to circadian desynchrony due to self-selecting inappropriate patterns of light exposure [8, 50–54]. This circadian misalignment leads to individuals experiencing a drive to sleep during scheduled wake and an inability to sleep during scheduled sleep opportunities. In order to preserve a stable 24-h pattern of work and sleep among the crewmembers, it may be desirable to provide a strong cycling of light and darkness in common spaces to mimic the solar light dark cycle and help crewmembers maintain a regular sleep-wake schedule and circadian entrainment [55, 56]. However, if such a strategy is utilized, it is important that crewmembers maintain some autonomy in controlling dimmer, personal lighting as would be the case at home on Earth. Similarly, crewmembers scheduled to be on night watch may benefit from supplemental lighting in the vehicle command center in order to enhance alertness and performance [57].

3.5 Noise

Noise is ever-present on space vehicles. We found that noise has been a major cause of sleep disruption throughout the history of spaceflight [12, 19, 29, 58]. The current guidelines allow for exposure to continuous noise above the WHO recommended guidelines [33, 59]. In addition, the current NASA guidelines do not provide mitigations against impulsive or intermittent noise [33]. We found that exposure to intermittent noise is at least as disruptive to sleep as continuous noise exposure [11, 12, 15, 19, 29, 58, 60, 61]. Given this evidence, exposure to noise be limited to below 35 dB, because exposure to noise above this level is associated with a reduction in sleep quality and quantity, even when individuals do not wake fully [59]. In addition, intermittent noise should be minimized, so that it does not vary beyond 5 dB from background noise levels. There is some evidence to suggest that exposure to continuous white noise less than 25 dB is sufficient to mask intermittent noises [62], therefore it is desirable to allow crewmembers access to white noise in their sleep chamber if desired. Earplugs and/or noise canceling headphones should also be made available for crewmembers [63]. Due to crewmember concerns about missing alarms while wearing earplugs, it may be desirable to develop multi-sensory alarms that include auditory and visual stimulation [64–66].

3.6 Temperature and Humidity

The ambient temperature on early space vehicles varied widely. For optimal sleep, an individual needs to reach his or her thermoneutral equilibrium and should have sufficient bedding available to create a microclimate of between 25–35 °C (77–95 °F) [67, 68]. Given that there are wide individual differences in the optimal temperature for sleep, the sleep environment on future space vehicles should be cool, but there should be sufficient insulation available for crewmembers to modify their environment to suit individual preferences [69–71]. This may mean providing crewmembers with sleeping bags of different thicknesses, or a mechanism for layering sleeping bags together. It is also desirable for sleeping bags to include vents to release heat, because the human core body temperature falls and rises during a typical sleep episode [72]. Warming of proximal and distal skin temperature has been associated with faster sleep onset [73–75] and crewmembers have reported having difficulty sleeping due to cold feet and hands [19, 34], therefore providing a way for crewmembers to warm their extremities prior to sleep may be desirable.

The level of humidity in the environment can also influence sleep quality and quantity. The optimal humidity range for human health is between 40 and 60 % [19]. The presence of humidity in the environment changes the perceived temperature. Humidity above 60 %, combined with high temperatures is disruptive to sleep [76]. Therefore, lower humidity of 50–60 % is optimal for sleep, particularly when ambient temperature is increased.

3.7 Air Quality

The optimal ambient gas mixture for sleep is equivalent to the air experienced at sea level on Earth (78 % nitrogen, 21 % oxygen, 1 % other gases) [16, 21, 77–86]. Similarly, the optimal air pressure during sleep is equivalent to the pressure on the Earth at sea level [87, 88]. Air mixtures that deviate from these conditions, such as what mountaineers experience during expeditions, results in disrupted sleep and periodic breathing [80, 82, 84, 88–90]. In depressurized environments, such as at elevation on Earth, supplemental oxygen can reduce headaches, periodic breathing, and can improve sleep outcomes [91, 92]. Airflow is also associated with positive sleep outcomes and aids in the reduction of O_2 [85, 93] and intrusive odors, such as body odor, food, and mechanical smells [12, 34, 85]. Although there is little information on the impact of air pollution and particulates on sleep quality and quantity, reports from lunar expeditions suggest that dust from planetary extra vehicular activities may build up in the habitable environment [29, 34]. As a result, the vents providing airflow to crew sleep chambers should include air filters to protect against crewmembers breathing particulate matter and dust during sleep.

3.8 Involuntary Movement

Involuntary movement due to turbulence is associated with sleep disruption [94]. Therefore, vehicle movement and vibration should be minimized as much as possible. Similarly, the microgravity environment results in the potential for crewmembers to free-float during sleep episodes. Although some crewmembers have reported that they enjoyed that experience, other crewmembers have reported that they prefer to be restrained while sleeping [95]. Given that some individuals may not use harnesses and other attachments, such attachments should be designed, so that they can be removed or secured out of place when not in use. Similarly, separate attachments should be available to secure the sleeping bag to the wall of the sleep chamber if desired.

3.9 Summary

In summary, sleep is critical to crewmember health and performance. In order for crewmembers to achieve optimal sleep, they must be provided with a sleep environment that allows them to achieve quality sleep, free of external disruption. We found that the optimal sleep environment is cool, dark, quiet, and is perceived as safe and private. There are wide individual differences in the preferred sleep environment; therefore modifiable sleeping compartments are necessary to ensure all crewmembers are able to select personalized configurations for optimal sleep. A sub-optimal sleep environment is tolerable for only a limited time, therefore individual sleeping quarters should be designed for long duration missions. In a confined space, the sleep environment serves a dual purpose as a place to sleep, but also as a place for storing personal items and as a place for privacy during non-sleep times. This need for privacy during sleep and wake appears to be critically important to the psychological well being of crewmembers on long duration missions. Designing sleep chambers for optimal sleep health should produce benefits beyond simply improving sleep quality and quantity on long duration missions.

References

1. Klerman, E.B., et al.: Comparisons of the variability of three markers of the human circadian pacemaker. J. Biol. Rhythms **17**(2), 181–193 (2002)
2. Akerstedt, T., Gillberg, M.: Effects of sleep deprivation on memory and sleep latencies in connection with repeated awakenings from sleep. Psychophysiology **16**(1), 49–52 (1979)
3. Wilkinson, R.: Some factors influencing the effect of environmental stressors upon performance. Psychol. Bull. **72**(4), 260–272 (1969)
4. Dijk, D.J., Czeisler, C.A.: Paradoxical timing of the circadian rhythm of sleep propensity serves to consolidate sleep and wakefulness in humans. Neurosci. Lett. **166**(1), 63–68 (1994)
5. Dijk, D.J., Czeisler, C.A.: Contribution of the circadian pacemaker and the sleep homeostat to sleep propensity, sleep structure, electroencephalographic slow waves, and sleep spindle activity in humans. J. Neurosci. **15**(5 Pt 1), 3526–3538 (1995)
6. Compton, D.W., Benson, C.D.: Living and Working in Space: A History of Skylab (1983)
7. Hoffman, S.J.: Antarctic Exploration Parallels for Future Human Planetary Exploration: A Workshop Report. National Aeronautics and Space Administration (2002)

8. Basner, M., et al.: Mars 520-d mission simulation reveals protracted crew hypokinesis and alterations of sleep duration and timing. Proc. Natl. Acad. Sci. U.S.A. **110**(7), 2635–2640 (2013)
9. Yan, X.W., England, M.E.: Design evaluation of an arctic research station from a user perspective. Environ. Behav. **33**(3), 449–470 (2001)
10. MSFC Skylab Structures: MSFC Skylab structures and mechanical systems mission evaluation. In: NASA TMX-64824, June. National Aeronautics and Space Administration (1974)
11. International Federation of Air Line Pilots' Associations: In-flight flight crew rest facilities (2013)
12. Hacker, B.C., Grimwood, J.M.: On the Shoulders of Titans: A History of Project Gemini. NASA SP-4203. NASA Special Publication, 4203 (1977)
13. Caldwell, J.A., et al.: Is Fatigue a Problem in Army Aviation: The Results of a Survey of Aviators and Aircrews. US Army Aeromedical Research Laboratory (2000)
14. Green, P.S.: Special Report: Sleep Deprivation During Deployment and PTSD. https://vanwinkles.com/for-our-returning-troops-post-traumatic-sleep-disorders-are-the-new-ptsd (2015)
15. Watt, C.G.: Aircraft Fatigue Management, in Air War College. Air University (2009)
16. Stuster, J.W.: Space station habitability: recommendations based on a systematic comparative analysis of analogous conditions. National Aeronautics and Space Administration (1986)
17. Carrere, S., Evans, G.W.: Life in an isolated and confined environment a qualitative study of the role of the designed environment. Environ. Behav. **26**(6), 707–741 (1994)
18. Hendrickx, B.: The Kamanin Diaries 1969–1971. J. Br. Interplanetary Soc. **55**(9/10), 312–360 (2002)
19. Johnston, R.S.: Skylab Medical Experiments Altitude Test (SMEAT) (1973)
20. Koros, A., Wheelwright, C., Adam, S.: An evaluation of noise and its effects on shuttle crewmembers during STS-50/USML-1 (1993)
21. MacCallum, T.K., Poynter, J.: Factors affecting human performance in the isolated confined environment of biosphere 2. In: Third Annual Mid-Atlantic Human Factors Conference, pp. 82–87 (1995)
22. Rosekind, M.R., et al.: Crew Factors in Flight Operations XIII: A Survey of Fatigue Factors in Corporate/Executive Aviation Operations. Moffett Field, CA, NASA (2000)
23. Vander Ark, S.T., Holland, A.W., Wood, J.: La Chalupa-30: Lessons Subsea Learned Mission from a 30-day Analogue (1994)
24. Weiss, K., Feliot-Rippeault, M., Gaud, R.: Uses of places and setting preferences in a French Antarctic station. Environ. Behav. **39**(2), 147–164 (2007)
25. Willshire, K.F.: Human response to vibroacoustic environments of space vehicles. National Aeronautics and Space Administration, Langley Research Center (1984)
26. Harrison, A.A., et al.: Implications of Privacy Needs and Interpersonal Distancing Mechanisms for Space Station Design (1988)
27. Yan, X.W., et al.: A critical review of design and use of field tent shelters in polar regions. Polar Rec. **34**(189), 113–122 (1998)
28. Legler, R.D., Bennett, F.V.: Space Shuttle Missions Summary. National Aeronautics and Space Administration: Mission Operations. Johnson Space Center (2011)
29. Aeronautics, National, Administration, Space: Apollo 11 Mission Report. Houston, TX (1971)
30. National Aeronautics and Space Administration: Apollo 13 Mission Report. Houston, TX (1970)
31. Shepard, Jr., A.: Apollo 14 Mission Report (1972)
32. Allen, C., Denham, S.: International Space Station Acoustics—a status report. In: 41st International Conference on Environmental Systems (2011)
33. Broyan, Jr., J.L., Borrego, M.A., Bahr, J.F.: International Space Station USOS Crew Quarters Development. SAE International (2008)
34. Bluth, B., Helppie, M.: Soviet Space Stations as Analogs (1986)
35. European Space Agency: Mars 500: Isolation Study (2010)

36. Czeisler, C.A., Gooley, J.J.: Sleep and circadian rhythms in humans. Cold Spring Harb. Symp. Quant. Biol. **72**, 579–597 (2007)
37. Flynn-Evans, E.E., et al.: Circadian misalignment affects sleep and medication use before and during spaceflight. npj Microgravity. **2**, 15019 (2016)
38. Lockley, S.W., et al.: Short-wavelength sensitivity for the direct effects of light on alertness, vigilance, and the waking electroencephalogram in humans. Sleep **29**(2), 161–168 (2006)
39. Lockley, S.W.: Timed melatonin treatment for delayed sleep phase syndrome: the importance of knowing circadian phase. Sleep **28**(10), 1214–1216 (2005)
40. Mien, I.H., et al.: Effects of exposure to intermittent versus continuous red light on human circadian rhythms, melatonin suppression, and pupillary constriction. PLoS ONE **9**(5), e96532 (2014)
41. Bierman, A., Figueiro, M.G., Rea, M.S.: Measuring and predicting eyelid spectral transmittance. J. Biomed. Opt. **16**(6), 067011 (2011)
42. Cho, J.R., et al.: Let there be no light: the effect of bedside light on sleep quality and background electroencephalographic rhythms. Sleep Med. **14**(12), 1422–1425 (2013)
43. Dijk, D.-J., et al.: Sleep, circadian rhythms, and performance during space shuttle missions. In: The Neurolab Spacelab Mission: Neuroscience Research in Space, pp. 211–221 (2003)
44. Figueiro, M.G., Rea, M.S.: Short-wavelength light enhances cortisol awakening response in sleep-restricted adolescents. Int. J. Endocrinol. (2012)
45. Figueiro, M.G., Plitnick, B., Rea, M.S.: Pulsing blue light through closed eyelids: effects on acute melatonin suppression and phase shifting of dim light melatonin onset. Nat. Sci. Sleep **6**, 149 (2014)
46. Grandner, M.A., et al.: Short wavelength light administered just prior to waking: a pilot study. Biol. Rhythm Res. **44**(1), 13–32 (2013)
47. Potter, J.J., et al.: Polar field tent shelters and well-being of users. Environ. Behav. **30**(3), 398–420 (1998)
48. Thompson, A., et al.: Effects of dawn simulation on markers of sleep inertia and post-waking performance in humans. Eur. J. Appl. Physiol. **114**(5), 1049–1056 (2014)
49. Zeitzer, J.M., et al.: Millisecond flashes of light phase delay the human circadian clock during sleep. J. Biol. Rhythms **29**(5), 370–376 (2014)
50. Arendt, J.: Biological rhythms during residence in polar regions. Chronobiol. Int. **29**(4), 379–394 (2012)
51. Halberg, F., et al.: Human biological rhythms during and after several months of isolation underground in natural caves. Bull. Natl. Speleol. Soc. **32**, 89–115 (1970)
52. Kleitman, N.: Sleep and wakefulness. Am. J. Med. Sci. **249**(2), 140 (1965)
53. Miller, N.L., Nguyen, J.: Working the nightshift on the USS John C. Stennis: implications for enhancing warfighter effectiveness. In: Proceedings of the Human-Systems Integration Symposium (2003)
54. Siffre, M.: Biological rhythms, sleep, and wakefulness in prolonged confinement. In: ESA, Proceedings of the Colloquium on Space and Sea, pp. 53–68 (1988)
55. Duplessis, C.A., et al.: Submarine watch schedules: underway evaluation of rotating (contemporary) and compressed (alternative) schedules. Undersea Hyperb. Med. **34**, 21–33 (2007)
56. Young, C.R., et al.: At-sea trial of 24-h-based submarine watchstanding schedules with high and low correlated color temperature light sources. J. Biol. Rhythms **30**, 144–154 (2015)
57. Barger, L.K., et al.: Sleep and cognitive function of crewmembers and mission controllers working 24-h shifts during a simulated 105-day spaceflight mission. Acta Astronaut. **93**, 230–242 (2014)
58. Flynn-Evans, E.E., et al.: Sleep duration, disruption and hypnotic use among 76 astronauts on short duration missions. In: NASA Human Research Program Meeting. Houston, TX (2012)
59. World Health Organization: Night Noise Guidelines for Europe (2009)
60. Fyhri, A., Aasvang, G.M.: Noise, sleep and poor health: modeling the relationship between road traffic noise and cardiovascular problems. Sci. Total Environ. **408**, 4935–4942 (2010)

61. Schmidt, F.P., et al.: Effect of nighttime aircraft noise exposure on endothelial function and stress hormone release in healthy adults. Eur. Heart J. eht269 (2013)
62. Stanchina, M.L., et al.: The influence of white noise on sleep in subjects exposed to ICU noise. Sleep Med. **6**(5), 423–428 (2005)
63. Department of the Army, Manual 6-22.5: Combat and Operational Stress Control Manual for Leaders and Soldiers, p. 111. US Department of the Army, Washington, DC (2009)
64. Kawada, T., Suzuki, S.: Change in rapid eye movement (REM) sleep in response to exposure to all-night noise and transient noise. Arch. Environ. Health **54**, 336–340 (1999)
65. Muzet, A.: Environmental noise, sleep and health. Sleep Med. Rev. **11**, 135–142 (2007)
66. Rechtschaffen, A., Hauri, P., Zeitlin, M.: Auditory awakening thresholds in REM and NREM sleep stages. Percept. Mot. Skills **22**(3), 927–942 (1966)
67. Okamoto, K., et al.: A survey of bedroom and bed climate of the elderly in a nursing home. Appl. Human Sci. **17**(3), 115–120 (1998)
68. Okamoto-Mizuno, K., Tsuzuki, K., Mizuno, K.: Effects of head cooling on human sleep stages and body temperature. Int. J. Biometeorol. **48**, 98–102 (2003)
69. Haskell, E.H., et al.: The effects of high and low ambient temperatures on human sleep stages. Electroencephalogr. Clin. Neurophysiol. **51**, 494–501 (1981)
70. Lin, Z., Deng, S.: A study on the thermal comfort in sleeping environments in the subtropics—measuring the total insulation values for the bedding systems commonly used in the subtropics. Build. Environ. **43**, 905–916 (2008)
71. Häuplik-Meusburger, S.: Architecture for Astronauts: An Activity-Based Approach. Springer Science & Business Media (2011)
72. Van Someren, E.J.W., et al.: Circadian and age-related modulation of thermoreception and temperature regulation: Mechanisms and functional implications. Ageing Res. Rev. **1**, 721–778 (2002)
73. Kräuchi, K., et al.: Warm feet promote the rapid onset of sleep. Nature **401**, 36–37 (1999)
74. Raymann, R.J.E.M., Swaab, D.F., Van Someren, E.J.W.: Cutaneous warming promotes sleep onset. Am. J. Physiol. Regul. Integr. Comp. Physiol. **288**, R1589–R1597 (2005)
75. Raymann, R.J.E.M., Swaab, D.F., Van Someren, E.J.W.: Skin deep: enhanced sleep depth by cutaneous temperature manipulation. Brain **131**(2), 500–513 (2008)
76. Okamoto-Mizuno, K., et al.: Effects of partial humid heat exposure during different segments of sleep on human sleep stages and body temperature. Physiol. Behav. **83**, 759–765 (2005)
77. Daues, K.R.: A history of spacecraft environmental control and life support systems (2006)
78. Gothe, B., et al.: Effect of quiet sleep on resting and CO_2-stimulated breathing in humans. J. Appl. Physiol. **50**(4), 724–730 (1981)
79. Lo, Y.-L., et al.: Genioglossal muscle response to CO_2 stimulation during NREM sleep. Sleep **29**(4), 470 (2006)
80. Reite, M., et al.: Sleep physiology at high altitude. Electroencephalogr. Clin. Neurophysiol. **38** (5), 463–471 (1975)
81. Robin, E.D., et al.: Alveolar gas tensions, pulmonary ventilation and blood pH during physiologic sleep in normal subjects. J. Clin. Invest. **37**(7), 981 (1958)
82. Salvaggio, A., et al.: Effects of high-altitude periodic breathing on sleep and arterial oxyhaemoglobin saturation. Eur. Respir. J. **12**(2), 408–413 (1998)
83. Schiffman, P., et al.: Sleep deprivation decreases ventilatory response to CO_2 but not load compensation. CHEST J. **84**(6), 695–698 (1983)
84. Selvamurthy, W., et al.: Sleep patterns at an altitude of 3500 metres. Int. J. Biometeorol. **30**, 123–135 (1986)
85. Strøm-Tejsen, P., et al.: The effects of bedroom air quality on sleep and next-day performance. In: Indoor Air (2015)
86. Szymusiak, R., Satinoff, E.: Maximal REM sleep time defines a narrower thermoneutral zone than does minimal metabolic rate. Physiol. Behav. **26**, 687–690 (1981)
87. Miller, J.C., Horvath, S.: Sleep at altitude. Aviat. Space Environ. Med. **48**(7), 615–620 (1977)
88. Mizuno, K., Asano, K., Okudaira, N.: Sleep and respiration under acute hypobaric hypoxia. Japan. J. Physiol. **43**, 161–175 (1993)

89. Dietz, T.E.: An Altitude Tutorial. http://www.ismmed.org/np_altitude_tutorial.htm (2006)
90. Muza, S.R., Fulco, C.S., Cymerman, A.: Altitude Acclimatization Guide. Army Research Institute of Environmental Medicine Natick MA Thermal and Mountain Medicine Division (2004)
91. Barash, I.A., et al.: Nocturnal oxygen enrichment of room air at 3800 meter altitude improves sleep architecture. High Altitude Med. Biol. **2**(4), 525–533 (2001)
92. Luks, A.M., et al.: Room oxygen enrichment improves sleep and subsequent day-time performance at high altitude. Respir. Physiol. **113**(3), 247–258 (1998)
93. Federal Aviation Administration: Flightcrew Member Duty and Rest Requirements (2011)
94. Matsangas, P., Shattuck, N.L., McCauley, M.E.: Sleep duration in rough sea conditions. Aerosp. Med. Hum. Perform. **86**, 901–906 (2015)
95. Petty, J.I.: Space Shuttle Life. http://spaceflight.nasa.gov/history/shuttle-mir/history/h-b-sslife.htm (2013)

Part XI
Road and Rail—Eco-driving and Electric Vehicles

Well Worth a Detour?—Users' Preferences Regarding the Attributes of Fast-Charging Infrastructure for Electromobility

Ralf Philipsen, Teresa Schmidt and Martina Ziefle

Abstract The spread of electric vehicles can be a partial solution for reducing greenhouse gas emissions, which are a major challenge of modern industrial nations. However, the limited ranges and the fragmentary charging infrastructure are currently impediments to adoption. To develop a need-based fast-charging network users' requirements on preferred charging locations have to be factored in. Therefore, the present study aimed at quantifying users' preferences regarding the fast-charging infrastructure and identifying possible trade-offs by using conjoint-analysis. Both current and potential battery electric vehicle users were addressed through an online questionnaire (N = 283). It was revealed that the waiting time for an available charging station, the necessary detour and the charging costs are the most important attributes for the selection of charging locations, whereas possible on-site activities to bridge the charging time were less important. While the attributes' importance was largely independent from trip length, participants' BEV experience contributed significantly to found variance.

Keywords Fast charging · Electromobility · Battery electric vehicles · Infrastructure planning · Evaluation criteria · User requirements · Conjoint analysis

R. Philipsen (✉) · T. Schmidt · M. Ziefle
Human-Computer Interaction Center (HCIC), RWTH Aachen University,
Campus-Boulevard 57, 52074 Aachen, Germany
e-mail: philipsen@comm.rwth-aachen.de

T. Schmidt
e-mail: schmidt@comm.rwth-aachen.de

M. Ziefle
e-mail: ziefle@comm.rwth-aachen.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_77

1 Introduction

Global warming as part of the climate change is one of the major challenges mankind is facing today. Amongst others, the man-made emission of greenhouse gases, in particular CO_2, caused by the burning of fossil fuels is one of the most important factors contributing to the warming of the atmosphere [1]. Therefore, a major aim of modern societies and especially of industrial nations should be to reduce those emissions.

Besides industrial production, the transport sector, especially the motorized private transport, is one of the domains that is heavily responsible for the emission of greenhouse gases [2]. New technologies in this area might initiate a change leading to a more efficient and sustainable mobility. Specifically, a switch of drive technology, i.e., from internal combustion engines to electric drives, should help to reduce emissions and reduce the dependence on fossil, and thereby limited, resources.

Consequently, the German Federal Government formulated the objective to increase the number of vehicles with electric drives in Germany to 1 million by 2020 in its “National Electromobility Development Plan” in 2009 [3]. However, at the present time, only a little more than 25,000 battery electric vehicles (BEVs) are registered in Germany [4]. Although the public perception of electric vehicles is predominantly positive, particularly with regard to the perceived environmental benefits [5, 6], there are still major impediments to adoption that prevent a widespread dissemination of this technology. In addition to the high costs of BEVs and their integrated batteries [7], especially the limited range and the long charging times of electric vehicles as well as the lack of a comprehensive charging network are perceived as major barriers [8, 9].

Therefore, fast-charging technology that enables recharging of a vehicle’s battery up to 80 % of its capacity in 20–30 min could alleviate these problems and enable new motion patterns with BEVs, e.g., long-haul travelling. However, the fast-charging infrastructure in Germany is still in the very early planning and build-up phase and predominantly dependent on public funding programs. A major and yet not fully answered question is, where to place the fast-charging stations to establish a both comprehensive and needs-oriented charging network.

2 Related Work and Question Addressed

Implementing a more user-focused design is an important consideration to achieve the aforementioned goals [10]. Therefore, it is necessary to understand both the users’ requirements for charging locations and their individual charging behavior. In response to this last point, there has been already a lot of fundamental research done focusing on driving in critical range situations and the psychological concept of range anxiety [11, 12], perceived range comfort zones [13], and the driver’s charging behavior in general [14, 15]. The results of these research approaches might be incorporated into recommendations for charging grid densities and the required absolute numbers of charging stations, but the selection of concrete locations for charging stations needs further research.

Indeed, there are already some approaches to factor the user in during the planning and development of charging infrastructure. However, current methods are mostly limited to activity-based [16, 17] or discrete-choice [18] models that are based on users' activity and movement profiles and try to cope with the increasing number of electric vehicles by simulating the regional technology spread [19]. The resulting information about optimal placement is normally limited to local district level; exact locations, for example, on the street level, cannot be proposed or evaluated. Therefore, a further integration of user factors and information about how users assess and select charging locations into the planning models is needed.

In previous focus group discussions, several requirements on charging locations and eight corresponding evaluation criteria have been identified: *dual use of time and route*, *habit compatibility*, *accessibility*, *visibility*, *reliability* in terms of an adequate number of functional and free charging points, *safety*, *connection to the public transportation network*, and *necessity* [20]. It was revealed that users' willingness to spend extra time for detours, waiting for an available charging point and the charging process itself is rather limited. Consequently, users requested to use the charging time meaningfully by parallel on-site activities.

However, the weighting of these evaluation criteria during the selection of charging stations is yet unknown. In real-world scenarios, charging locations rarely can fulfill all the users' requirements in an ideal way. Therefore, the selection process is always a trade-off between possibly conflicting objectives that presumably depend on type and length of the trip. Additionally, the prior experience with BEVs could influence the trade-off decisions, as it affects the handling of range limits [11, 12] and the general evaluation of electromobility [21].

The present work aimed at identifying possible trade-offs by quantifying the preferences of both early adopters and potential future users regarding the attributes of fast-charging infrastructure in different trip scenarios.

3 Method

A questionnaire utilizing conjoint analysis was used to address the research question. In the following, the development of the questionnaire and the design of the conjoint analysis will be presented in detail. Subsequently, a short summary of data acquisition, preparation, and statistical analysis will we reported. Last, the gathered sample will be introduced.

3.1 Questionnaire

Based on the aforementioned findings from focus group discussions [20], a questionnaire was developed. It aimed at quantifying the trade-offs between different attributes of charging stations that are considered during the driver's decision making process. Therefore, the questionnaire consisted of two parts. The first part dealt with personal information of the participants, especially demographic data like age, gender, occupation, and income. Additionally, (e-)mobility-related behavior in terms of general attitude toward electromobility, car use, annual mileage, charging possibilities at home or at work, and, if appropriate, range and connector types of the used BEV were queried. To gain an impression of the participants' technical affinity, Beier's inventory to gauge technical self-efficacy [22] was used. The second part of the questionnaire used a conjoint measurement to research the participants' selection of charging locations.

3.2 Conjoint Study

To explore participants' preferences regarding the selection of charging locations, a choice-based conjoint (CBC) design was used. This method addresses the problem that trade-offs between user requirements can hardly be measured by conventional question designs in terms of an isolated rating of charging station attributes.

To take different trip lengths into consideration, a scenario-based approach was used. One scenario dealt with long-haul trips, mainly on highways, and the other broached the issue of short-haul trips in terms of everyday urban driving. The trip scenario was designed as within-group factor; thus, every participant had to answer two separated sets of conjoint-related questions. To address possible knowledge gaps and to create a more uniform basis of decision-making, participants were told that they would need 25 min to recharge their BEV and their battery's capacity was sufficient for a 100 km drive which complies with current BEV's typical performance characteristics.

The selection of attributes and their levels (see Table 1) was based on focus group findings. The first two attributes were derived from criteria of *dual use of time and route*. Consequently, possible on-site activities to spend the charging time constitute the first attribute. Its levels distinguish between a productive use of time, activities just to keep busy, and no parallel activities at all. The second attribute dealt with the detours that are necessary to reach the charging location. Concrete distances between 0 and 15 km were used to specify no, small, medium, and large detours. The distances varied between the short- and the long-haul scenario to be more adequate. The third attribute was deduced from the *reliability* criteria. Users in focus groups requested an immediate charging possibility after arriving; therefore, the waiting time for an available charging station was the next used attribute. The levels varied between no waiting time at all and 30 min. The last attribute was not

Table 1 Attributes and associated levels used in conjoint analysis

Attributes	Possible parallel activities to spend charging time …	Charging costs	Necessary detour	Waiting time for an available charging station
Levels	… productively	100 % of fuel costs	15 km [a]	30 min
	… keeping busy	75 % of fuel costs	10 km	15 min
	None	50 % of fuel costs	5 km	5 min
		Free charging	2 km [b]	None
			None	

[a]used only in long-haul scenario
[b]used only in short-haul scenario

based on focus group findings but should help to research whether financial incentives influence the charging location selection. Due to the current confusing and very heterogeneous market situation, it was expected that concrete prices are not assessable for non-users. Therefore, more abstract levels were used, matching charging costs to conventional fuel costs needed to achieve the same range. All levels were introduced to the participants before the actual questions and illustrated with additional graphics. Especially the levels of the parallel activities attribute were described in detail, e.g., activities to spend the charging time keeping busy meant activities the participants might enjoy but would not do without charging, like getting a coffee to keep boredom at bay.

For both scenarios, participants had to carry out ten tasks each. A task consisted of a selection of the preferred charging location out of four randomly composed alternatives based on the predefined attributes and levels. There was no possibility to skip tasks or use a "none of these"—option.

3.3 Data Acquisition, Data Preparation and Analysis

The survey was conducted via online-channels. It was spread in the social environment of the authors and at the university as well as by using social networks and expert forums dealing with (e-)mobility. The aim was to acquire both early adopters, in terms of current BEV users, and potential users, i.e., participants who are strongly interested in electromobility but have no purchase intention yet due to the already introduced barriers. Although a total of 326 participants completed the survey, 43 who stated that they are absolutely not interested in electromobility were removed to focus on the aforementioned user groups.

Hierarchical Bayes (HB) estimation was used to analyze the CBC-data by computing the attributes' relative importance scores as well as the corresponding part-worth utility values for each participant. In addition, non-parametric methods were used for scenario and user group comparisons. The level of significance was set to $\alpha = 0.05$ and two-tailed tests were used for the statistical analysis.

3.4 Sample

A total of 283 (N) participants have been included in the analysis. 214 (78.1 %) were male, 60 (21.9 %) female. The age ranged from 19 to 82 years with a mean of 42.40 years (SD = 14.80). The educational level was rather high; in particular, 183 participants, and thus nearly two-thirds of the sample (64.6 %), stated that they had achieved a university degree. 16.2 % (n = 46) of them were even graduated. The second most frequent educational attainment was graduation from high school (n = 59, 20.8 %), followed by completed vocational trainings (n = 27, 9.5 %) and secondary school (n = 14, 5.0 %). Most participants were employed (n = 169, 59.7 %), whereas 62 (22.0 %) were still in training or education and 25 (8.8 %) were retired. The technical self-efficacy in the sample was rather high with M = 3.71 (SD = 1.01, scale min = 0, scale max = 5).

Mobility Behavior and Charging Possibilities. 63.0 % of the participants (n = 178) reported that they drive more than 10,000 km per year. 199 participants (70.8 %) stated they would be able to charge a BEV at home with no or little effort, whereas 74 (27.1 %) could charge at work. Concerning prior experience, 55 participants (19.4 %) indicated that they already use a BEV as private or company car. 92.7 % of these (n = 51) used it on a daily base and 29.1 % (n = 16) used BEVs with fast-charging compatible plugs like CCS or CHAdeMO. The most frequent ranges achievable with one battery charge were 100–150 km (n = 27, 49.1 %) and 50–100 km (n = 12, 21.8 %).

4 Results

Hereinafter, the results of the survey will be presented in detail. First, the relative importance of the charging stations' attributes as well as the part-worth utilities will be introduced for the complete sample. Second, the effects of BEV experience on the attributes' relative importance will be presented.

4.1 Model Fit and Relative Importance Scores

The HB model's goodness of fit was determined by the root likelihood (RLH) that could vary between 1 (best) and 0.25 (worst) under the present CBC design with

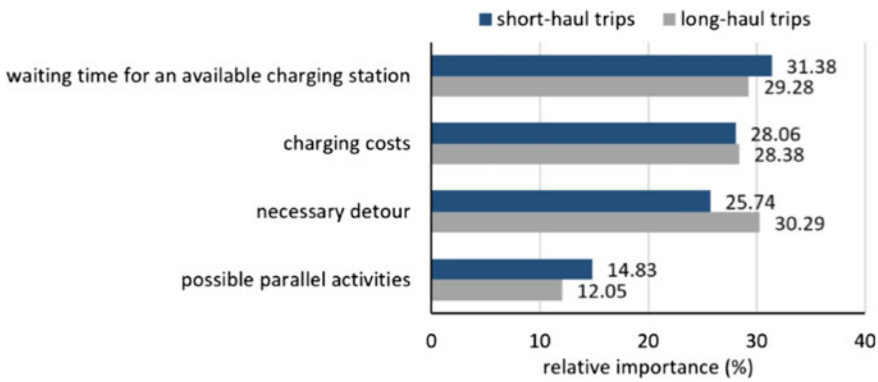


Fig. 1 Average importance of charging stations' attributes for short-haul and long-haul trips

four choices. The calculated RLH was 0.61 for the short-haul and 0.64 for the long-haul scenario.

Average importance scores for all attributes were calculated to quantify how they contribute to the participants' decision making. A full overview about the scores for both short-haul and long-haul trips can be found in Fig. 1.

Concerning short-haul trips, the necessary waiting time for an available charging station was most important, followed by charging costs and the detour necessary to reach the charging location. These three most relevant attributes had a relative importance of roughly 30 %. In contrast, the lowest importance value deviated clearly and was the possibility to bridge the charging time with parallel activities.

In comparison to short-haul trips, the order of importance changed in the long-haul scenario. In particular, necessary detours got the highest average importance value. The most important attribute for short-haul trips, i.e. the needed waiting time for an available charging station, was only the second most important factor regarding the selection of charging stations on long-haul trips. However, the importance of the latter two as well as of the third most important attribute, i.e. the costs for charging, again were approximately the same with values approximately 30 %. Likewise, the on-site possibility for parallel activities had the lowest relative importance value by a clear margin.

The comparison of short-haul and long-haul trips revealed several differences. Both the waiting time for an available charging station and possible parallel activities are significantly more important for short-haul than long-haul trips ($Z_{waiting\_time} = -3.809$, $p < 0.001$ and $Z_{activities} = -4.982$, $p < 0.001$). In contrast, the necessary detour was more important for long-haul trips with $Z_{detour} = -6.424$, $p < 0.001$. The relevance of charging costs for the selection of charging locations did not differ depending on trip length.

4.2 Part-Worth Utilities

Looking at the part-worth utility values of the attribute levels (see Fig. 2), it can be shown that the order of preferences mirrored the designed order of precedence. Accordingly, participants' average preferences decreased from no waiting time for available charging stations at all to the longest waiting time of 30 min. However, the part-worth utility values for no waiting time and a period of 5 min lay close together, whereas the distances between the other attribute levels were considerably higher.

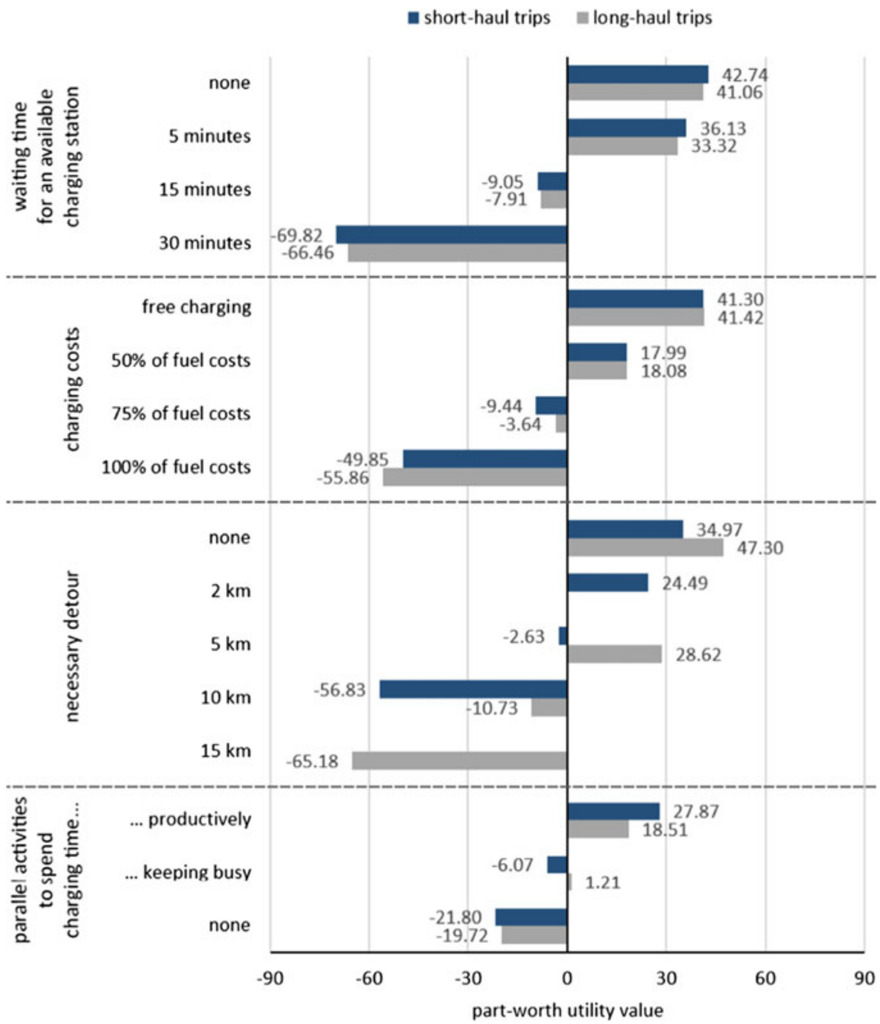


Fig. 2 Part-worth utility values (zero-centered diffs) for charging station attributes distinguished by trip length

Concerning charging costs, it becomes obvious that there is a strict best to worst order. As expected, free charging is preferred most, followed by charging costs about half as expensive as conventional fuel. The other two cost levels, i.e., charging costs of 75 and 100 % of the amount of fuel costs, were considerably less preferred.

A comparable picture has also been obtained for the detour attribute; that is, preferences decreased from no detour at all to the presented maximum detour. Concerning the possible activities to pursue during charging, activities to spend time productively had clear priority over activities to spend time keeping busy. Absence of on-site possibilities for activities parallel to charging got the lowest utility values and was thereby the least preferred level.

Although most part-worth utility values differed between the short- and long-haul scenario to a small extent, there were no differences with regard to the relative order of preferences for all attribute levels. Relating to the detour attribute, the absolute differences of part-worth utilities between the scenarios were bigger in comparison to the other attributes, but due to the different parametrization of this attribute, they were only partially comparable.

4.3 Effects of Prior Experience with BEVs

Taking prior experience with BEVs into account, several differences between the user groups were revealed in both trip scenarios. First, the effects of BEV use in the short-haul scenario will be described (see Fig. 3).

The importance of two charging station attributes differed significantly between BEV users and interested non-users with regard to short-haul trips. The waiting time for an available charging station was more important to BEV users than non-users ($U = 4247$, $p < 0.001$, $r = -0.22$), whereas the necessary detour was

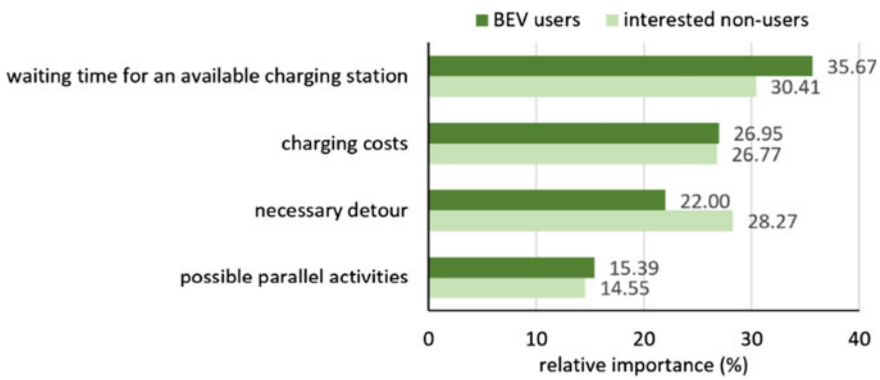


Fig. 3 Average importance of charging stations' attributes for **short-haul trips** distinguished by user groups

more important to participants without BEV experience (U = 4536, p = 0.002, r = −0.19). In contrast, the importance values of both charging costs and possible parallel activities did not differ significantly depending on the user groups in this scenario.

The mentioned attributes' differences result in different orders of importance, too. Both BEV users and non-users matched concerning the most and least important attribute, whilst detours were more important to BEV users than charging costs, this order reversed for non-users.

Concerning long-haul trips, a comparable picture emerged (see Fig. 4). The attributes' order of importance for BEV users matched the short-haul scenario and the waiting time for an available charging station is more important to users than non-users (U = 3907, p < 0.001, r = −0.26), whereas the necessary detours were more important to non-users (U = 4725, p = 0.005, r = −0.17). Once more, there was no significant difference between user groups regarding the charging costs. However, the importance of possible parallel activities was influenced by BEV use in the long-haul context. These activities were more important to non-users than users (U = 4951, p = 0.015, r = 0.14).

Therefore, the attributes' order of importance concerning non-users changed in comparison to short-haul trips. The necessary detours were the most important attribute for non-users during the selection of charging locations, whereas charging costs and the waiting time for an available charging station followed jointly in second place (n.s. differences between attributes). Accordingly, possible parallel activities at charging locations played by far the least important role.

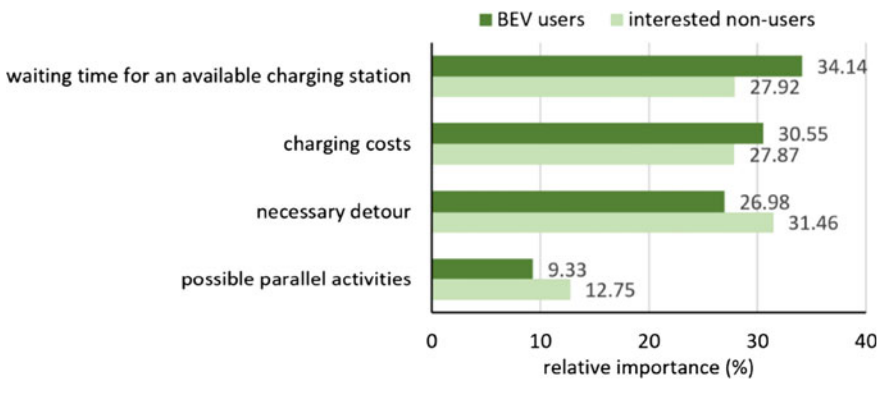


Fig. 4 Average importance of charging stations' attributes for **long-haul trips** distinguished by user groups

5 Discussion

The present study aimed at quantifying users' preferred attributes of fast-charging stations for electromobility to factor them in during planning and development processes. To know users' preferences is necessary for establishing a comprehensive charging network that is need-based and no longer perceived as barrier to adoption.

It was revealed that the waiting time for an available charging station, the charging costs, and the necessary detour to reach a charging location are the most important attributes considered during the user's selection process. All three attributes shared nearly the same importance level, i.e. accounting for about 30 % of the user's decision. Following the 'time is money' principle, monetary incentives in terms of low charging costs theoretically can almost equally compensate high waiting times or long detours. However, the nearly equal importance of waiting times and detours was not expected, because while both attributes imply a loss of time, detours additionally cause a range reduction and extra charging costs. This finding might be caused by different possible uses of time; in particular, activities during detours are limited to driving, whereas more varied actions are possible in or around parked vehicles during the wait for an available charging station. Furthermore, detours are predictable for familiar charging locations, whereas the occupancy rate at a station can be a factor of uncertainty, if not retrievable in advance, for example, by web-based services. Therefore, further studies should address the effects of underlying factors, e.g., contrasting loss of time and loss of range as well as uncertainty and calculability. Surprisingly, in contrast to expectations after the focus group discussions [20], the fourth researched attribute in terms of possible activities at a charging site was identified as least important for a user's selection process. Although locations that enable activities to bridge the charging time productively were preferred to locations that only enable to kill time or offer no activities at all, this characteristic might be disregarded, if the other location attributes offer enough perceived benefits. As a consequence, charging stations at shopping or recreation facilities as requested in [20] can be a good choice of location, but stations will only gain market share there, if waiting times, charging costs, and necessary detours will not contradict this locational advantage.

Further, the meaningfulness of revealed part-worth utilities is somehow limited, because, due to the measurement scale that is limited to relative preferences, it was not possible to identify absolute deal breakers that would prevent the selection of a certain charging station completely. It is still unclear what the turning points of attribute levels are, for example, which amount of detour, waiting times, or charging costs would lead to a complete dismissal of a charging location. Currently, it can be only stated that, for example, waiting times of 0 or 5 min seem to be rated nearly equal, whereas delays beyond 15 min are considerably less preferred. To identify crucial criteria for location selection, further studies with a more adaptive conjoint analysis design flanked by isolated ratings will be necessary.

Concerning trip scenarios, the current study could not reveal major differences. Although, detours were more important on long-haul trips, which might also be influenced by deviating parameterization of distance levels, the trip length had no decisive influence on the attributes' order of importance. However, the current study used only two abstract scenarios to utilize short- and long-haul trips. The next step will be to use more precise distance information to contrast trip length and maximum BEV range, especially with regard to the number of necessary charging stops. Furthermore, the types of trips, e.g., business trips, trips for shopping or leisure activities, or holiday journeys, should be taken into consideration.

Finally, the effects of BEV experience were analyzed. It was shown that users and non-users currently differ regarding the importance of waiting times and detours, whereas there basically is consensus regarding charging costs and parallel activities. Although the emerged picture was not completely uniform, the revealed commonalities should be helpful for establishing a charging infrastructure that can cope with the requirements of both the current early adopter group and potential future users who, as of yet, perceive the status of the charging network expansion as impediment to purchase.

However, the effects of prior BEV experience were only small and therefore can explain only a small amount of the identified variance. Consequently, further user factors should be taken into consideration in future studies. To start with, emphasis should be placed on demographic factors, e.g., gender, age, or residential information, that can usually be acquired from public authorities at street or at least district level. Although a deeper understanding of the charging location selection, that might also include further human factors, should be aspired, the consideration in simulation tools for infrastructure planning is limited to the available data base.

In conclusion, the present work is one of the first steps to measure user requirements on concrete charging station placement. The obtained findings could be implemented in concepts and simulations for charging location evaluation that aim to predict how specific site characteristics contribute to the acceptance and use of fast-charging infrastructure to be built. Based on the relative weights of individual importance levels, information about the planned location, and with respect to underlying business models, it can be assessed, for example, whether the site offers the opportunity for parallel activities and, thereby, allows a positive utilization forecast, or whether a location is well worth a detour.

Acknowledgments The authors thank all participants for their patience and openness to share opinions on a novel technology. Furthermore, thanks go to Chantal Lidynia, Julia van Heek, Luca Liehner, and Hans-Ulrich Voß for their research assistance.

References

1. The Royal Society, Joint Science Academies' Statement.: Global response to climate change. https://royalsociety.org/~/media/Royal_Society_Content/policy/publications/2005/9649.pdf. Last visited on 11 Mar 2016 (2005)

2. IPCC.: Climate change 2014—synthesis report. In: Pachauri, R.K., Meyer, L.A. (eds.) Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Switzerland (2015)
3. Bundesregierung, Nationaler Entwicklungsplan Elektromobilität [National Electromobility Development Plan]. https://www.bmbf.de/files/nationaler_entwicklungsplan_elektromobilitaet.pdf. Last visited on 11 Mar 2016 (2009)
4. Kraftfahrt-Bundesamt, Personenkraftwagen am 1. Januar 2016 nach ausgewählten Merkmalen. http://www.kba.de/DE/Statistik/Fahrzeuge/Bestand/Ueberblick/2016_b_barometer.html. Last visited on 11 Mar 2016
5. Hoffmann, C., Hinkeldein, D., Graff, A., Kramer, S.: What do potential users think about electric mobility? In: Evolutionary Paths Towards the Mobility Patterns of the Future, pp. 85–99. Springer (2014)
6. 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—User Experience Design for Everyday Life, pp. 628–639. Springer (2014)
7. 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)
8. Egbue, O., Long, S.: Barriers to widespread adoption of electric vehicles: an analysis of consumer attitudes and perceptions. Energy Policy **48**, 717–729 (2012)
9. 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)
10. International Energy Agency.: Global EV outlook—understanding the electric vehicle landscape to 2020. http://www.iea.org/publications/freepublications/publication/GlobalEVOutlook_2013.pdf. Last visited on 11 Mar 2016 (2013)
11. Rauh, N., Franke, T., Krems, J.F.: Understanding the impact of electric vehicle driving experience on range anxiety. Hum. Factors J. Hum. Factors Ergonomics Soc. (2014)
12. Rauh, N., Franke, T., Krems, J.F.: User experience with electric vehicles while driving in a critical range situation—a qualitative approach. IET Intel. Transport Syst. **9**(7), 734–739 (2015)
13. 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, Austria, 5–6 June 2014
14. Franke, T., Krems, J.F.: Understanding charging behaviour of electric vehicle users. Transp. Res. Part F Traffc Psychol. Behav. **21**, 75–89 (2013)
15. Bühler, F., Franke, T., Schleinitz, K., Cocron, P., Neumann, I., Ischebeck, M., Krems, J.F.: Driving an ev with no opportunity to charge at home—is this acceptable? In: Proceedings of the Human Factors and Ergonomics Society Europe (2013)
16. 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)
17. 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 Transport Environ. **41**, 165–176 (2015)
18. Bernardo, V., Borrell, J.R., Perdiguero, J.: Fast charging stations: network planning versus free entry (2013)
19. Namdeo, A., Tiwary, A., Dziurla, R.: Spatial planning of public charging points using multi-dimensional analysis of early adopters of electric vehicles for a city region. Technol. Forecast. Soc. Chang. **89**, 188–200 (2014)
20. Philipsen, R., Schmidt, T., Ziefle, M.: A charging place to be—users' evaluation criteria for the positioning of fast-charging infrastructure for electro mobility. Procedia Manufact. **3**, 2792–2799 (2015)

21. Zaunbrecher, B.S., Beul-Leusmann, S., Ziefle, M.: Laypeople's perspectives on electromobility—a focus group study. In: Internet of Things. IoT Infrastructures, pp. 144–149. Springer International Publishing, Berlin (2014)
22. Beier, G.: Kontrollüberzeugung im Umgang mit Technik [Locus of control when interacting with technology]. Rep. Psychol. **24**(9), 684–693 (1999)

Introducing Wireless Charging for Drivers of Electrical Vehicles in Sweden—Effects on Charging Behaviour and Attitudes

Jonas Andersson, Maria Nilsson and Stefan Pettersson

Abstract This paper reports on a Swedish large-scale research and demonstration study of wireless charging of electric vehicles. The study is the first of its kind outside North America. The purpose of the 18-month study was to test the technology during real life working conditions using 20 electric vehicles located at eight municipality and company sites in Sweden. The study indicates that the charging behaviour will most likely be different with inductive charging. There are clear benefits of inductive charging that have the potential to increase the attractiveness of electric vehicles, and there are no substantial evidence that perceived safety should hinder a wider adoption of inductive charging. Further, we conclude that the usability of the technology can have a high impact on perceived attractiveness, and should therefore be of focus in future developments of the technology.

Keywords Inductive charging · Electric vehicles · Demonstration · User study · Attitudes

1 Introduction

Electrical vehicles have the potential to decrease our fossil dependency. However, to fulfil that vision electrical vehicles have to become wide spread. Factors such as range anxiety [1, 2], grid overload (cf. [3]), (cumbersome) charging process [4, 5], among others, have the potential to become barriers that may hinder this development. Wireless (inductive) charging is a new technology that may overcome

J. Andersson · M. Nilsson (✉) · S. Pettersson
Viktoria Swedish ICT, Lindholmspiren 3A,
SE-417 56 Göteborg, Sweden
e-mail: maria.nilsson@viktoria.se

J. Andersson
e-mail: jonas.andersson@viktoria.se

S. Pettersson
e-mail: stefan.pettersson@viktoria.se

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_78

many of these identified issues by eliminating the cord [6]. However, large-scale field trials are yet to be seen. Previous studies have typically been performed in laboratory settings.

Hence, this paper reports on a Swedish large-scale research and demonstration trial of inductive wireless charging project called WiCh - Wireless Charging of electric vehicles (www.wich.se), see [7] for further information. The study is financed by the Swedish Energy Agency together with the project partners. This is the first study of its kind outside North America [8]. The overall goal of the study was to demonstrate and test the technology during real life setting in Swedish conditions (e.g., cold and snow), to explore the pros and cons in practical use of inductive charging. Another aim was to study the drivers' attitudes towards using inductive charging for operating the electric vehicle in their daily working context. In total, inductive wireless charging technology, provided by Plugless Power (www.pluglesspower.com), was installed in 20 vehicles located at eight sites in Sweden, to be used during an 18-month period. Employees in the public sector were to use the vehicles in their daily work, except one (private).

Of interest in the study was to examine whether or not the practicalities of not dealing with a cord would make electrical vehicles more attractive. Franke and Krems' [4] show that 87 % of the drivers agreed that charging was easy. Although they report: "However, several users (57 %) reported that handling the charging cable was cumbersome". That is, even though the authors' conclude that charging was not a major barrier for users in this study, the cumbersome cord is noted. In addition, it has been shown that the cord can become stiff, especially in cold weather [9], which may influence the attractiveness of electrical vehicles. It was of further interest to examine the potential change in charging behaviour when introducing inductive wireless technology. Studies have highlighted that wireless charging has the potential to solve some grid issues [3]; e.g., the vehicles has the potential to be charged more often but for less time. However, what the drivers, who are to implement these effects of inductive charging, think of the technology is yet to be revealed. Previous studies have shown that electric vehicle drivers adopt their charging behaviour to the environmental and structural setting they exist in [1], e.g., most drivers are likely to charge their vehicles over night (as they do not use the vehicles then) and does not see a problem with e.g., the length of charging.

Hence, the purpose of the research is to explore whether inductive charging actually has the advantages that can further promote electric vehicles as an alternative to fossil fuelled vehicle. In this paper, the first results from the study are presented.

2 Method

The study was carried out in two stages, as a comparative before (Stage1) and after study (Stage 2), to explore change in attitudes towards practical use, attractiveness and perceived safety when introducing wireless charging. The data collection consisted of interviews and digital surveys as well as vehicle loggers. Stage 1 of the study involved capturing current attitudes and experiences of using electrical vehicles charged with cord. Stage 2 involved capturing the experience of using electrical vehicles with the wireless charging technology installed. So far, the collected data consists of sixteen interviews (before introduction of inductive charging, and after six months of use), approximately 200 survey responses, and logged driving and charging data from the 20 vehicles. In this paper, results from the first three digital surveys are presented. Further statistical analyses are yet to be presented.

2.1 Research Questions

The main research objective of the WiCh-project is to answer the question if inductive charging changes the charging behaviour of the users. That is, the following questions are addressed: (1) Does inductive charging change charging behaviour? (2) Does inductive charging make electric vehicles more attractive compared to charging with cable? (3) Does perceived safety hinder use of inductive charging? In this paper, initial results on all three questions are presented. Further statistical analyses are yet to be presented.

2.2 Study Setting

Within the WiCh-project 20 vehicles has been equipped with inductive wireless charging technology supplied by Plugless Power (www.plugglesspower.com) [7].

More specifically, the wireless inductive technique uses strongly coupled magnetic resonance that enables high-energy transfer over large distances [10]. The charging equipment is made up by two different units, a sender placed in the ground at a parking space and a collector mounted underneath the vehicle. A coil in the sender generates a magnetic field that induces an electric current in a coil in the collector. The induced current is then used to charge the battery in the car. To get high efficiency the frequencies of the sender and the collector must be coupled. This, so called resonant coupling, occurs when the natural frequencies of the two units are approximately the same. When this happens the two units can exchange energy through their oscillating magnetic fields, without losing too much energy to the surroundings [10].

The vehicles equipped with the technology are distributed over 8 different sites in the south of Sweden, in the cities of Stockholm, Uppsala and Göteborg. The vehicles, are used in daily operations for transportation purposes by municipality representatives (except one vehicle, which is used by a private user). The representatives uses the vehicles in their ordinary (official) travel typically performed during office hours. The sites (except the private user) have all a history of using electric vehicles, typically more than 12 months experience. The majority of the vehicles are parked in a garage but other parking facilities include carport and enclosed car park with no weather shelter. The electric vehicles that are being used in the study are Nissan Leaf and Chevrolet Volt. The sites have access to one or more vehicles and are typically set up as a carpool accessible via a booking system. Two sites have dedicated drivers for the equipped electric vehicles. The vehicles are typically used several times a week. At some sites, the vehicles are used one or more times every (working) day. The typical length of a journey is about 20 km.

2.3 Survey Design

The surveys contained twenty questions in total regarding user background, vehicle use patterns, charging pattern, and attitudes towards charging of electric vehicles. After completing the user-, vehicle use-, and charging profile questions, the respondents performed ratings on a six grade Likert-scale, combined with free text comments to capture opinions and attitudes towards charging electric vehicles. Finally, the respondents could leave free text opinions about the survey. The survey design was adapted not to exceed a response time of about ten to fifteen minutes. The surveys were designed using the Survey Monkey© online tool and were distributed via e-mail to the participants. Before receiving the e-mail, prospective participants were informed of the study by the researchers and gave their consent to participate. Alternatively, the participants were contacted by a representative from their own organisation who distributed the survey link.

The first survey was distributed to the end users in May 2015 to capture their attitudes to charging (with cord) and electrical vehicles in general (Stage 1). The second survey was distributed approximately two months after installation to capture initial experience of using the new technology, and the third survey about four months after installation of inductive charging. A fourth questionnaire is planned to be sent out at the end of the test period. For number of respondents of each survey, see Table 1.

Table 1 Number of respondents, gender distribution and experience of EVs

	Q1: Cable	Q2: Induction 1	Q3: Induction 2
Number of resp.	65	56	44
Male/female (%)	65/35	59/41	63/37
Experience of EVs			
0–6 months (%)	20	18	13
6–12 months (%)	18	8	13
>12 months (%)	62	74	74

2.4 Participants

All participants came from the project partner organisations within WiCh. The age distribution was 26 to 76+ years for all three surveys and the main part of respondents were between 36 and 55 years of age. The main part of the users can be considered as experienced electrical vehicle (EV) drivers (Table 1), and they use the vehicles regularly in their daily work tasks.

The number of responses differed between the participating organisations. The distribution of responses from each site was roughly 5–15 % of total responses with one deviating organisation with much higher (35–55 % of total) response rate for the three surveys.

3 Results and Analysis

The results are analysed and presented graphically together with qualitative assessments of text answers. For presentation purposes, and first assessment, the responses from the six grade Likert-scales (e.g. “entirely positive” to “entirely negative”) used in the survey have been divided into e.g. positive/negative side as presented in the diagrams in the results section. The selected survey results presented in this paper address the emerged charging pattern of inductive charging, the attractiveness of electrical vehicles, and perceived safety of charging electrical vehicles, before and after the introduction of inductive charging.

3.1 Emerging Charging Behaviour

The usability of the charging equipment is perceived as slightly poorer compared to charging with cable, and was further reduced at the time of the third survey (Fig. 1). The negative ratings increased in correspondence to the drop on the positive side. With cable charging, 90 % of the respondents were positive towards the charging procedure (Fig. 2). After introduction of inductive charging, the number of positive respondents was reduced to 72 %, followed by a drop to 64 % after four months of inductive charging use. The users without opinion in the question were about 4 % in all three surveys. A qualitative analysis of the free text answers indicates that the drop in perceived usability is related to difficulties of parking correctly over the charging pad. The experience of having to make several attempts to park the car correctly seems to cause annoyance affecting the answers.

This is further highlighted by the change in opinions related to the time it takes to handle the charging equipment (Fig. 3), showing a drop in positive responses with inductive charging. It is noteworthy that for cable charging about 80 % of the respondents report that it takes 0–1 min to start the charging. For inductive

Fig. 1 Perceived usability

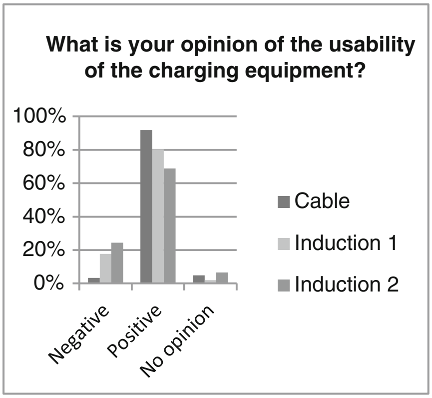


Fig. 2 Charging procedure

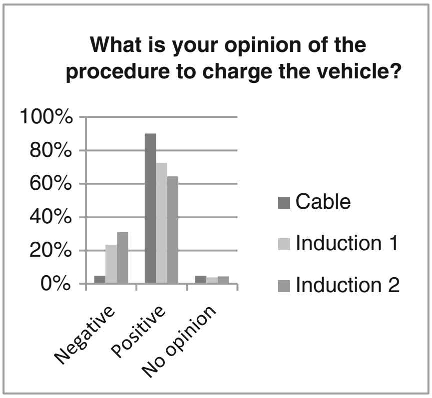
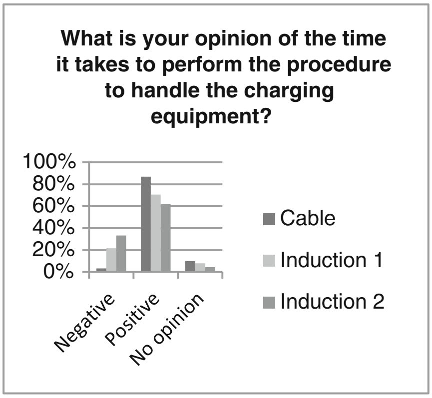


Fig. 3 Time to perform charging procedure



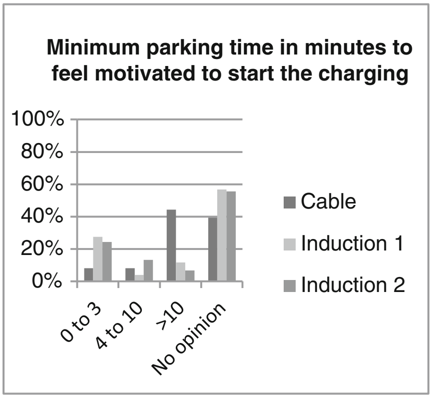


Fig. 4 Minimum parking time to feel motivated to start charging

charging the same figure are approximately 50 % of respondents, while the remaining part says it takes 2–10 min to get the inductive charging started.

The minimum parking time to feel motivated to start the charging (Fig. 4) was however judged to be lower when using inductive technology. This despite the decreasing rate of positive answers regarding the time it takes to perform the charging procedure.

Also, on the multiple-choice question "*At the following occasions I have not started the charging despite available charging*", the alternative "*The car is parked too short period of time*" dropped from 36 % (Cable) to 3 and 8 % for inductive charging indicating that the parking time becomes less of an issue with induction when deciding to start charging or not. The inductive charging thus seem to have a potential of leading to more charging occasions, if used in a working context with many short stops (and available inductive charging at stop locations).

3.2 *Perceived Attractiveness*

In addition, we were interested in if electric vehicles become more attractive when using inductive charging. On the question if the electric vehicle is an attractive mode of transport the positive answers increased from 75 to 88 % after introduction of inductive charging (Fig. 5). This figure dropped to 68 % after having used the technology for four months. Thus, the first impression of inductive charging seems very positive but due to technical issues it did unfortunately not meet the expectations of the users. In the written comments the respondents have positive things to say about the technology as such, although, they are not satisfied with the implementation at their specific workplace. Briefly said, the technology is good when it works but there are room for improvements. On the question of preferred mode of charging, the preference for cable charging is almost unchanged between surveys two and three (Fig. 6). The preference for inductive charging dropped from 49 to 34 % between surveys, while the indifference alternative increased slightly.

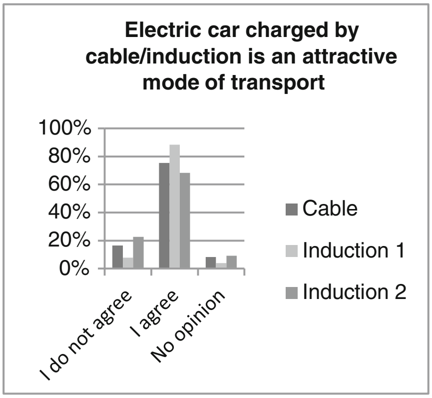


Fig. 5 Attractive mode of transport

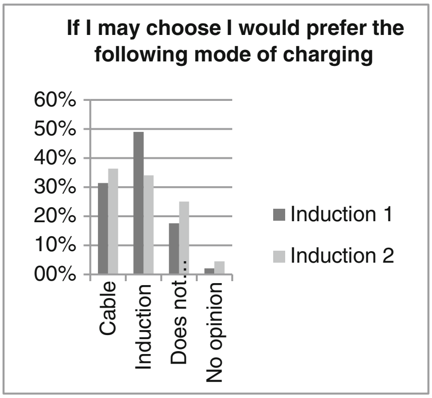


Fig. 6 Preferred mode of charging

An interesting finding is that despite the perceived problems with the charging equipment, users still maintain the positive attitude toward electric vehicles in general. The question on experience of using electric vehicles showed a slide towards increasingly positive responses after having used the inductive charging for four months (Fig. 7).

At this point it is however difficult to say what has caused the shift. It could be societal factors lying outside the scope of the project to study. When looking at the free text responses, one reoccurring comment is however that inductive charging is more convenient since you don't need to handle the (at times dirty) cable. Examples of quotes are: "*It is more convenient than charging with cable and there is no risk of forgetting to connect the cable…*"; "*Good to get rid of the socket and cable*"; "*Nice to avoid handling a cable that easily becomes dirty…*". Getting rid of the cable seems to contribute to the positive experiences. The quote "*Inductive charging, as such, is good…*" also indicates that it could be the case that the users separate the concept and idea of inductive charging from their own implementation.

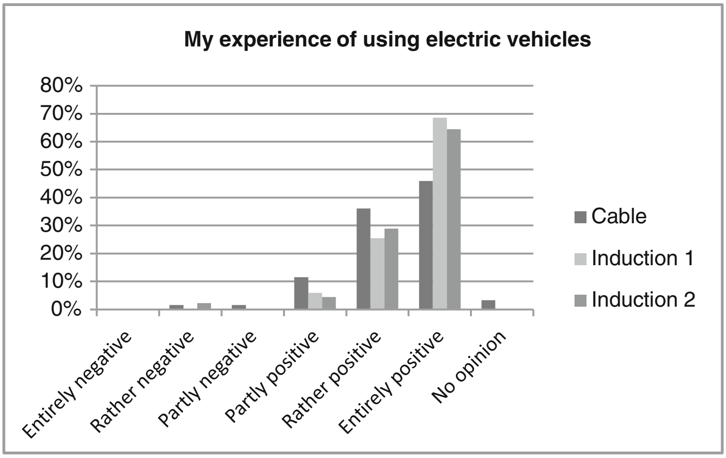


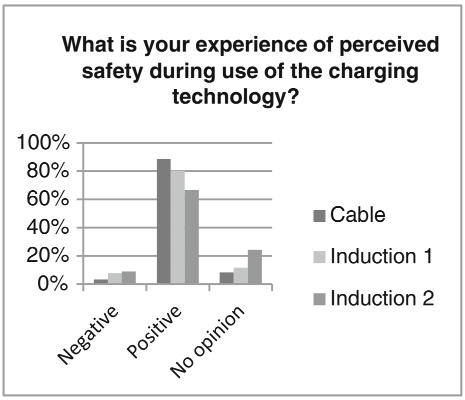
Fig. 7 Experience of using electric vehicles

Some of the participants seem to like the new technology despite it hasn't worked optimally at their specific location.

3.3 *Perceived Safety*

One concern with inductive charging is the perceived safety. The survey results show that the positive opinion drops over time, similarly to the dimension of practical use (Fig. 8). However, the negative side does not increase accordingly. Instead the opinions seem to have shifted to "No opinion". An explanation for this could be that the safety aspects regarding electro-magnetic radiation from the inductive chargers are difficult to assess as an individual user. The radiation is

Fig. 8 Perceived safety of the charging technology



invisible and less tangible and it is impossible to know if you are being exposed while starting or stopping the charging, which might be a source of anxiety.

4 Discussion

The discussion attempts to answer our three main research questions. The main findings from the study are discussed.

4.1 Does Inductive Charging Lead to a Change in Charging Behaviour?

The results from surveys indicate that new charging patterns are likely to appear when introducing inductive charging. The drivers say that the parking time to motivate a start of charging is shorter compared to cable. Interestingly, when using inductive charging, the task of connecting the cable is replaced by the task of parking correctly over the charging pad (at least in the evaluated charging solution). Since the parking task will still be made, there should be a reduced effort required from the driver. However, the evaluated technology needs refinement for an optimal user experience helping the driver to hit the precise charging spot. (With future autonomous vehicles or integrated automatic parking aid this issue might disappear.) Yet, our results show that the users think it is convenient to get rid of the cable, which promotes use of inductive charging. Combining these results tells us that there is a potential for an increased number of charging opportunities given that the charging infrastructure is present. This means that the charging behaviour will be different with inductive charging. Exactly how the charging behaviour will change is yet to be studied.

4.2 Does Inductive Charging Make Electric Vehicles More Attractive?

According to our results, the first impression of inductive charging seems to be perceived very positively. The technical problems in the project however caused the technology not to live up to the expectations of many of the drivers. Also, the difficulties of parking the car correctly over the inductive charger pad in the first attempt, with several attempts needed to find the exact spot for the charging to start, the perceived benefit of not having to handle the charging cable was very much reduced. The positive comments in the free text answers of the technology *as such* still points to a positive attitude although being discontent with the implementation.

Other free text answers also points to perceived convenience of inductive charging, not having to handle the cable. The results as a whole points to the importance of a well designed driver-charger interaction where the evaluated solution has room for improvements. When the technology works, the drivers also see clear benefits of inductive charging which reasonably also makes the electric vehicle more attractive. As our study shows, poor interaction design can have the opposite effect.

4.3 Does Inductive Charging Become Hindered by Perceived Safety Issues?

Despite the drop from the positive side of the perceived safety ratings, very few users have reverted to feeling more negative regarding their perceived safety. In general there is a predominance of positive ratings. Malfunctioning equipment could also add to the insecurity of if the equipment is safe or not, in other aspects than electromagnetic radiation. In the study, we have not seen any substantial evidence that perceived safety should be a hindrance for wider adoption of inductive charging. However, the case could be different for home users where children and pets will be closer to the charging equipment.

5 Conclusions

To summarise, we found the following initial answers to our three research questions: (1) the charging behaviour will most likely be different with inductive charging, (2) there are clear benefits of inductive charging that have the potential to increase the attractiveness of electric vehicles, and (3) we have not seen any substantial evidence that perceived safety should be a hindrance for wider adoption of inductive charging. Further, poor interaction design can have a high impact on perceived attractiveness. Indeed, this far in the demonstration project, the evaluations shows that the tested inductive charging technology has potential to provide an easier way of charging, and hence, never having to "re-fuel" the vehicle. According to the surveys, the most difficult thing, compared to charging with cord, is to park correctly over the charging pad on the first attempt. Due to the placement of the charging pad at the rear of the vehicle, exact positioning can be difficult. However, the tested equipment still seems somewhat immature to provide an optimal user experience.

It has been evident that the technical problems with the surrounding equipment (e.g. malfunctioning ground circuit breakers, parking difficulties, charging of 12 V batteries) affected the attitudes towards the inductive charging in a negative way. In addition, it should be noted that the conclusions made here are drawn from a study with professional drivers, not private drivers using the vehicle in their home context. The findings are probably not valid for home users since it can be expected that charging behaviours and attitudes will change with the context of use.

Further statistical analysis are yet to be performed that include the forth questionnaire.

Acknowledgments Participating organisations in the WiCh project are (in alphabetic order); City of Stockholm Environment and Health Administration, Göteborgs Gatu AB (Gatubolaget), Viktoria Swedish ICT, Swedish Radiation Safety Authority, Test Site Sweden and Vattenfall. The project granted by the Swedish Energy Agency.

References

1. Nilsson, M.: Electrical vehicles: the phenomenon of range anxiety. Public deliverable task (1300) ELVIRE Project. Available at: http://www.elvire.eu/IMG/pdf/ (2011)
2. The_phenomenon_of_range_anxiety_ELVIRE.pdf
3. Nilsson, M.: Electric vehicles: an interview study investigating the phenomenon of range anxiety. Public deliverable task (5000) ELVIRE project. Available at: http://www.elvire.eu/IMG/pdf/An_interview_studyinvestigating_the_phenomenon_of_range_anxiety_ELVIRE-2.pdf (2011)
4. US department of energy, Evaluating electric vehicle charging impacts and consumer charging behaviours-experience from six smart grid investment grant projects. Available at: http://energy.gov/sites/prod/files/2014/12/f19/SGIG-EvaluatingEVcharging-Dec2014.pdf (2014)
5. Franke, T., Krems, J.F.: Understanding charging behaviour of electric vehicle users. Transp. Res. Part F Traffic Psychol. Behav. **21**, 75–89 (2013)
6. Nilsson, M., Habibovic, A.: Identifying EV drivers´ needs for information communication technology to ease the EV charging process. In: Proceedings of Automotive UI 2013, Eindhoven, Netherlands. Available at: http://www.auto-ui.org/13/ (2013)
7. Lukic, S., Pantic, Z.: Cutting the cord: static and dynamic inductive wireless charging of electric vehicles. IEEE Electrification Mag. **1**(1), 57–64 (2013)
8. Pettersson, S., Wedlin, J., Fransson, T., Andersson, J., Klingegård (f. Nilsson), M.: Large scale testing of wireless charging in Sweden. In: EVS29 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium (2016)
9. Evatran and Swedish Agencies to Install PLUGLESS Systems across Sweden. Available at: http://sustainablerace.com/evatran-and-swedish-agencies-to-install-plugless-systems-across-sweden/
10. Nilsson, M.: Experience electrical vehicle the case of fast charging. Public Deliverable North Sea Region Electric Mobility Network. Available at: http://archive.northsearegion.eu/files/repository/20130716113709_FastCharge_Pilot_experiences.pdf (2012)

Exploring the Value of Information Delivered to Drivers

Jaemin Chun, SeungJun Kim and Anind K. Dey

Abstract New IT functions have greatly increased the amount of in-car information delivered to drivers. Although valuable, that information can distract drivers when delivered during vehicle operation. By inferring driver state from sensor data, prior research has shown that it can accurately identify opportune moments to deliver information. Now that we know when to best deliver information, it raises the question: what information should we deliver at those interruptible moments? To answer this question, we conducted a series of surveys and interviews and compiled a list of representative in-car information items and context factors that affect the importance of these items. By combining and exploring those context factors, we identified driving situations when each of the in-car information items is highly valuable, and verified these situations through a large online survey of drivers. Lastly, we examined what technology is available for detecting these driving situations, and which situations require further advanced technologies for detection. Results from our study offer important insights for understanding the diversity of drivers' experiences about the value of in-car information and the ability to determine situations in which this information is valuable to drivers. With these results, researchers can then build information delivery systems that can deliver information to drivers both when they are interruptible and when they find the information valuable.

Keywords In-car information · Information value estimation · Driver experience

J. Chun (✉) · S. Kim · A.K. Dey
Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, USA
e-mail: jaemchun@andrew.cmu.edu

S. Kim
e-mail: sjunikim@cs.cmu.edu

A.K. Dey
e-mail: anind@cs.cmu.edu

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_79

1 Introduction

Connected IT devices have greatly increased the amount of information that is being provided to drivers. Currently, more than 60 types of information are shown in-car and this number increases with each new function added to the car (e.g, collision warnings, infotainment systems, telematics). Various functions to improve driver comfort, safety and entertainment are constantly being developed and deployed by automakers. Additionally, as drivers focus on what is happening outside of the car, they are presented with a variety of attention-demanding information sources.

However, human attention is finite [1]. Delivering information while driving can distract drivers and thus detract from driving performance. According to the National Highway Traffic Safety Administration (NHTSA), such distractions can be classified into four categories: visual, auditory, cognitive and biomechanical [2]. Incoming information may demand drivers' visual and auditory attention, while planning and reacting to the delivered information may cause additional cognitive and biomechanical distractions.

Once information is delivered, drivers are forced to choose whether to pay attention to the information or ignore it by quickly calculating trade-offs between its perceived value and the distraction cost it is likely to incur [3]. Although drivers may have their own strategies of allocating attention to in-car information sources, distracted driving is still one of the major causes of car accidents: 10 % of fatal crashes and 17 % of injury crashes in 2011 were caused by driver distraction [4].

To address the issue of distracted driving, researchers have developed intelligent systems that monitor driver states to detect opportunities for interruption. Vision-based technologies (e.g, capturing eye gaze, eye blinking face pose, head orientation) have been the most popular method for driver monitoring and have been used mainly to detect driver fatigue and drowsiness [5]. Vehicle monitoring systems, such as OBD (on board diagnosis devices), have been used to collect data about car movements (e.g, speed, acceleration, angle of the steering wheel) [6, 7]. Driver behaviors also have been tracked by advanced wearable sensing technologies [8, 9]. By applying those technologies, systems have been developed that can detect interruptible moments (e.g., pushing information while the driver is conducting driving-related tasks). As an example, in our own recent research, we built a system that achieved 94 % accuracy in detecting driver interruptibility using a combination of wearable sensing technologies and a vehicle monitoring device [3]. Current interruptibility detection technologies have nearly reached the point where they are both practical and deployable.

In addition to determining *when* to interrupt drivers, researchers must address the question of *what type of information* to provide at those moments. Information pushed to drivers during vehicle operation may divide attention. As such, information should be delivered to drivers whenever its *value* can offset the *cost* of receiving it. Further, as the value of in-car information may vary according to a driver's situation, we need to identify the situations when the information value is high. Our paper contributes the following: *identifying valuable in-car information items and the driving situations in which their value is particularly high.*

2 Research Procedure

In the early stage of our study, we identified several unanticipated issues that needed to be addressed. First, we found that the type and the number of in-car information items vary among car models and car manufacturers. Some drivers do not sufficiently understand the meaning of in-car information (shown as icons or symbols), leading to challenges in estimating their value. This matter becomes more serious for novice drivers or when the information is of a type that is rarely shown to drivers (e.g., anti-lock braking system warning lamp, supplemental—restraint system warning).

Second, drivers have difficulty in describing driving situations. Often, participants feel it is difficult to select the context factors that are relevant to the situation they picture. Sometimes they think their scenario (the described driving situation) lacks detail or contains irrelevant details by selecting irrelevant context factors.

Third, the value of the in-car information cannot be accurately measured by observing drivers' behavior (such as eye gaze) since valuation is an internal, mental process. In fact, information-seeking behaviors are the result of a driver's negotiated decision: from time to time, drivers may choose an information source that they can actually attend to rather than the one they need most at the moment. For this reason, evaluating the value of information based on driver behavior could lead different interpretations than by asking the driver directly.

After carefully examining the issues listed above, we took the following steps to achieve our research goal:

- Step1: Identify representative in-car information.
- Step2: Identify context factors that affect the value of the in-car information.
- Step3: Develop driving scenarios when the value of selected in-car information item is expected to be high.

3 Step 1: Selection of Representative in-Car Information

Due to space limitations, we cannot cover all the information we showed to driver. Therefore, we selected representative information by assessing its importance through surveys.

3.1 List of in-Car Information Items

We first collected the list of in-car information from drivers' manuals. Seven models with high sales volume from five car manufacturers (Benz, BMW, Ford, Honda and Hyundai) and 2 GPS devices (Garmin and TomTom products) were selected to extract 61 common information items. Among them, 31 items were the type of information that drivers can check at any time (e.g., speedometer, gear position). The remaining 30 items are shown to drivers only when a relevant event happens (e.g., brake warning lamp, battery check lamp). We used different evaluation criteria for the two types of information. Anytime information was evaluated according to two criteria: importance and frequency (i.e., how frequently drivers check the in-car information). The relevant-event information was evaluated based only on importance. Examples of in-car information items are shown in Fig. 1.

3.2 Representative in-Car Information Items

Because we observed that drivers' understanding of in-car information depends on their driving experience, we considered two groups of drivers: novice and experienced. Because there is not a definitive way to separate those groups, we asked the participants to review their driving experience (e.g., for how long have they driven, how often they drive, how far they drive daily) and asked them to self-select whether they considered themselves to be novice or experienced drivers. Ten participants in their 20 s and 30 s were recruited for the each group. The novice group (age average = 25.5, age range = 21–33) reported an average driving frequency of 1–2 day/week over an average of 1.60 years of driving. The experienced

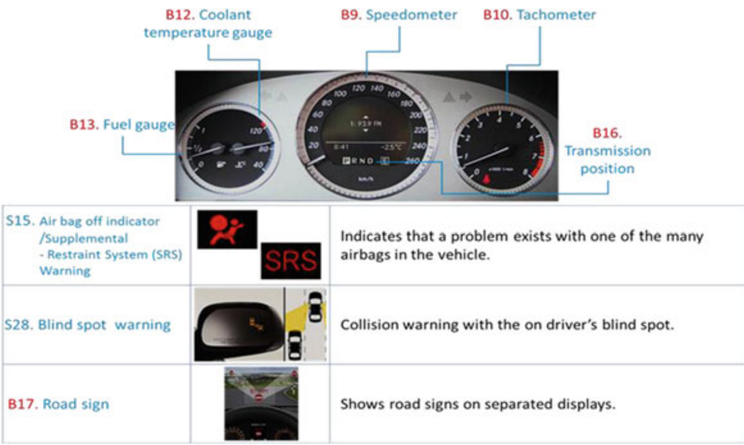


Fig. 1 Examples of explanations for information shown in car

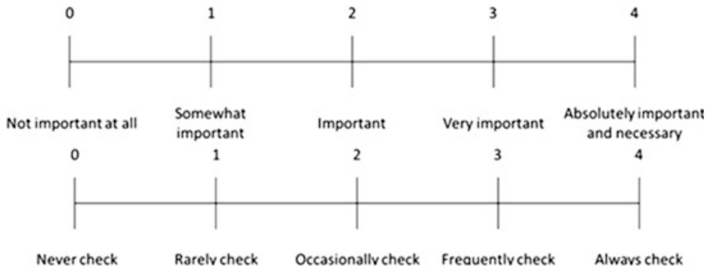


Fig. 2 Scales used for estimating importance (*above*) and frequency (how often drivers need the information, *below*) of in-car information

group (age average = 29.6, age range = 24–35) reported an average driving frequency of 4–5 days/week over an average of 7.30 years.

The participants were asked to rate the importance of the in-car information using a 5-point scale (Fig. 2). We also asked the participants to determine weights for the criteria (i.e., importance of the information and frequency of need for the information) to determine their relative importance. The average weights for the criteria (the importance and frequency) were 0.479 and 0.521 for novice group and 0.560 and 0.440 for the experienced group.

Table 1 summarizes the in-car information items identified as very important (average score >3) by the novice and experienced groups. As shown, the novice group tends to regard in-car information that tells the driver about external environments (e.g., collision warnings and road sign) as highly important. On the other hand, the experienced group put more importance on things happening inside the car.

Often, novice drivers are not very confident with their driving skill and thus want to be assisted by the car. In fact, some of these drivers reported that they often fail to identify the movement of other cars or miss road signs because they are busy checking road conditions. In contrast, experienced drivers reported that they do not have trouble in checking outside situations unless they are driving in a very new

Table 1 In-car information items chosen to be important by the novice and experienced drivers

Type of information	Always shown	Selectively shown
Novice drivers	– Speedometer – Fuel gauge – **Estimated time of arrival**	– Fuel reserve indicator – Engine overheating light – **Road sign** – **FCW/BSW/RCW**
Experienced drivers	– Speedometer – Fuel gauge – **Clock**	– Fuel reserve indicator – Engine overheating light – **Brake warning lamp** – **Check engine light**

In-car information items in bold are distinctively selected by the groups
**FCW* Forward Collision Warning, *BSW* Blind Spot Warning, *RCW* Rear-end Collision Warning

(e.g., driving in other country) or a severe driving condition (e.g, night driving in a heavy rain). Rather, they focused on the significance of the information and wanted to receive the information as reminders.

Although the brake warning lamp and the check engine light indicate vehicle malfunctions, the novice drivers scored those items lower than did the experienced drivers. The novice drivers knew that these items were important (because they are shown in red) and related to vehicle malfunction, but they ignore them because the meanings of those warnings are unclear and they believe that their car would continue to operate anyway. This was particularly true with icons that connote multiple meanings: the brake warning lamp could indicate a brake system problem or that the hand/emergency brake is engaged. The check engine light indicates that there is some problem in the engine but the reason underlying the problem is unspecified.

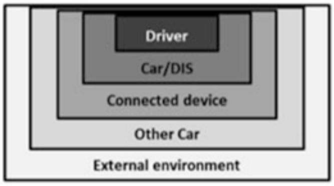
Both novice and experienced drivers selected as important some basic and essential information related to vehicle state, including speed, fuel, and engine overheating warnings. However, those groups selected those items for slightly different reasons. For example, most of the novice drivers checked the fuel gauge because they worried about gas depletion, whereas experienced drivers are concerned with finding a gas station with cheap gas in a reachable distance. From these results, we realized that the value of in-car information is truly affected by a driver's experience and thus affected by a different set of context factors.

4 Step 2: Identification of the Context Factors that Affect the Value of in-Car Information

To help participants specify driving situations, we provided them with a list of driving-related context factors. For each selected in-car information item, we selected and highlighted context factors that are closely related to them.

4.1 Context Factors as Building Block for the Driving Situation

For participants to describe a driving situation, they need to know which context factors can be used as building blocks for the driving situation. To obtain potential context factors, we modified the works of [10, 11]. Then, we verified the factors by asking 5 novice and experienced drivers to examine the list. Based on their responses, we derived 37 context factors in five categories—driver's internal state, driver's behavior/task, car/Driving Information System (DIS), other cars, and road environment (Fig. 3).



Category	Context factor	Example of context factor
Driver	Profile	Age, gender...
	Driving experience	Novice, Experienced, Expert
	Physical state	Sweating, cold, warm, back pain...
	Emotional state	Anger, neutral, dull, fear, panic...
	...	...
External environment	Weather	Cloudy, foggy, snowing, raining, sunny...
	Traffic	Light, heavy, jammed...
	Traffic signal	Stop, go...
	Type of road	School zone, highway, exit, ramp...
	Road condition	Dry, wet, paved, unpaved, cracked, frozen...
	Road sign	Stop, speed limit, no right turn...
	...	...

Fig. 3 Example of context factors for drivers

Contrary to our expectations, simply providing participants with the list of context factors did not work well. Influences of each context factor are valued differently among in-car information items. For example, the speed limit may affect the value of information displayed on the speedometer, yet have no effect on the value of information shown on the fuel gauge. In fact, not all the derived context factors are equally useful for the in-car information items we selected. Also, considering all 37 context factors at once is a complex task for participants, who often felt overwhelmed by the long list.

4.2 *Context Factors that Affect the Value of in-Car Information*

To address this problem, we highlighted highly related context factors for each selected in-car information items. This allowed participants to begin describing driving situations more easily. We then collected drivers' opinions concerning which context factors to highlight for each in-car information item.

Ten novice drivers (age average = 25.8, age range = 21–34) and 10 experienced drivers (age average = 28.4, age range = 24–33) drivers were recruited. The novice group reported an average driving frequency of 1–2 day/week over an average of 1.47 years of driving. The experienced group reported an average driving frequency of 4–5 days/week over an average of 6.78 years. The participants were provided with the context factor list (Fig. 3) and list of selected in-car information items (Table 1). Participants rated the effect of context factors on each selected in-car information item and debriefing questions were followed to solidify their decision. Participants rated the effect of each context factor using a 5-point scale (Fig. 4). During the debriefing session, participants were asked to briefly describe a driving situation related to their selected context factors and to double check the other context factors to verify that they had not overlooked any important factors.

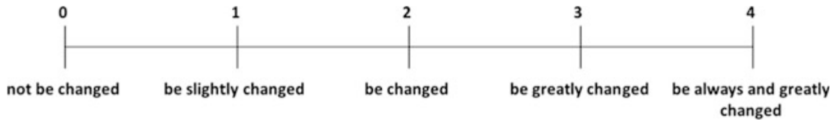


Fig. 4 Example of the scales used for estimating effect of context factors on in-car information items

Category	Context factor	Speedo meter	Fuel gauge	Brake warning lamp	…
	…	…	…	…	…
Driver	Primary task	o			…
	Secondary task				…
	Hours of driving		o		…
…	…	…	…	…	…

Fig. 5 Example of the context factors which has a strong effect on the value of in-car information (denoted by 'o')

For each in-car information item, we highlighted the context factors with an average score greater than 2.5 (Fig. 5). Our results showed that highlighted context factors are different among the in-car information items, which reflects that the value of each in-car information item can vary depending on the driving situation (that can be defined by a set of context factors).

5 Step 3: Deriving Representative Driving Situations at Which the Selected in-Car Information Is Valuable

In this step, we collected situations for which selected in-car information is valuable based on drivers' actual experience and verified the situations with a large number of drivers through online surveys.

5.1 Driving Situations for Which the Selected in-Car Information Is Valuable

Sixteen drivers participated in our interview. For both the novice (age average = 26.6, age range = 23–33) and experienced groups (age average = 28.3, age range = 23–35), we recruited 8 participants. The average driving frequency and years of driving were 2–3 day/week and 1.42 years for the novice group, and 4–5 day/week and 6.43 years for the experienced group.

The sets of highlighted context factors and the list of selected in-car information were provided to the participants. The participants were asked to apply each of the highlighted context factors to the relevant in-car information item to specify driving situations when they think the selected in-car information item is important. We asked them to start with the context factors that are close to drivers and then move outward to explore the external environment. Through this process, we collected 452 driving situations for 16 in-car information items (9 for novice driver group; 7 for experienced driver group).

5.2 *Representative Driving Situations Derived by a Large Number of Drivers*

Although we collected a large set of driving situations, the value of selected in-car information varied among those situations. To derive consensus from a large number of drivers, we conducted an online survey. Responses were collected from 96 drivers: Thirty-three novice drivers (age average = 27.0, age range = 18–53) and sixty-three experienced drivers (age average = 31.8, age range = 20–71). The novice group reported an average driving frequency of 1–2 day/week over an average of 1.86 years of driving. The experienced group reported an average driving frequency of 4–5 days/week over an average of 9.90 years. On our online survey, those drivers rated the value of selected in-car information for the selected driving situations using a 7-point scale (score from 0 to 6). Verbal descriptions for the scale are as followed: point 0: I absolutely do not need this information at the described driving situation, point 3: This information is somewhat important but not crucial for the described situation, and point 6: This information is the most valuable for the described driving situation.

Tables 2 and 3 show the driving situations for which the value of the provided information is high (top 3 situations; only two situations were derived for the ‘engine overheating’ and ‘Real-end Collision Warning RCW’). Driving situations derived from the survey helped us to better understand how context sensitive (depending on drivers’ experience and driving situation) the value that each in-car information item is. For example, when exiting a highway or approaching a curved road, novice drivers want to check the car speed to determine if they are driving slow enough for the situation (and ranked the speedometer as being important, see Table 2), while experienced drivers can sense their speed without checking the speedometer. Many novice drivers feel nervous when the fuel reserve indicator turns on, which they want to know as soon as they start the engine so that they can drive immediately to a gas station. As previously noted, novice drivers tend to miss road signs or find it difficult to check their details. One frequently reported case occurs when novice drivers encounter a road sign with time variant information (such as effective hour information written on a traffic signal or a parking sign). In those instances, they become nervous and are afraid of shifting their attention from

Table 2 Top ranked driving situations when the value of the selected in-car information is high for the novice drivers

In-car information	Rank (average score)	Driving situation
Speedo-meter	**1** (4.89)	**When driving on a road which is regulated by a speed limit**
	2 (4.50)	**When I need to slow down because the road is in a bad condition (frozen, slippery, bumpy…)**
	3 (4.43)	When I'm in a road where I need to slow down (Highway exit, curve ahead …)
Fuel gauge	**1** (5.39)	**When I am about to leave for a long distance travel. (To see if the amount of fuel is enough)**
	2 (4.07)	When I'm about to pass a gas station that offers a good deal (cheap price or additional promotion)
	3 (3.43)	When I am not driving in a fuel efficient way (heater On, full trunk…)
Fuel reserve indicator	**1** (4.43)	When starting the engine
	2 (4.11)	**When I am not driving in a fuel efficient way (heater on, fully trunk…)**
	3 (3.93)	**When I am about to pass a gas station that is the only one that is nearby**
Engine overheating	**1** (5.00)	**When my car seems to have a mechanical problem (sound, vibration…)**
	2 (4.07)	When I have been driving for a while but still need to drive a long distance (to reach destination or repair shop)
	3 (-)	–
Road signs	**1** (5.29)	When I'm searching for a place to park (parking hours, fare…)
	2 (5.21)	When I'm about to pass a point at which speed limit changes
	3 (5.11)	Whenever there emerges a sudden change on the road (sharp curve, highway exit, merging road…)
ETA	**1** (4.57)	When the fuel gauge hits low
	2 (4.50)	When I am in a hurry to meet my schedule
	3 (4.32)	When I take a wrong route (e.g. miss highway exit)
FCW	**1** (4.75)	When I try parking into a narrow space
	2 (4.75)	When my sight is instantly interrupted (bright sunshine, heavy rain…)
	3 (4.64)	When I am conducting a secondary task (texting, phone call…) and my attention is away from road
BSW	**1** (5.14)	When I change lanes
	2 (5.04)	When I park my car (parallel parking)
	3 (4.82)	When two roads merge into one
RCW	**1** (5.32)	When I park my car (parallel parking)
	2 (4.68)	Whenever a car approaches from behind (difficult to shift attention to room mirror)
	3 (-)	–

The bolded situations are also selected by the experienced drivers

Table 3 Top ranked driving situations when the value of the selected in-car information is high for the experienced drivers

In-car information	Rank (Average score)	Driving situation
Speedo-meter	1 (4.80)	**When driving on a road which is regulated by a speed limit**
	2 (4.10)	When I recognize that my current speed is quite different from other cars
	3 (3.86)	**When I need to slow down because the road is in a bad condition (frozen, slippery, bumpy, …)**
Fuel gauge	1 (5.43)	When the fuel gauge level is significantly low
	2 (5.04)	**When I am about to leave for a long distance travel. (To see if the amount of fuel is enough)**
	3 (4.43)	When the time is late (evening time) and gas stations are about to close
Fuel reserve indicator	1 (5.11)	When I am driving to a new place (do not really know the location of a nearby gas station)
	2 (5.04)	**When I am about to pass a gas station that is the only one that is nearby**
	3 (3.78)	**When I am not driving in a fuel efficient way (heater On, full trunk…)**
Engine overheating	1 (5.41)	**When my car seems to have a mechanical problem (strange sound, vibration…)**
	2 (4.98)	When I'm about to go on a long distance travel
	3 (4.70)	When the weather is very warm/hot
Brake warning lamp	1 (5.74)	When my car seems to have a mechanical problem (sound, vibration…)
	2 (5.57)	When I ride on a road in a risky condition (dark, slippery, paved, downhill…)
	3 (5.33)	When I am riding at a high speed following a car ahead of me
Check engine light	1 (5.47)	When my car seems to have a mechanical problem (sound, vibration…)
	2 (5.04)	When I am about to go on a long-distance travel
	3 (4.62)	When I am away from a town and thus cannot directly repair my car
Clock	1 (5.00)	When I am late for a schedule
	2 (4.61)	When I want to estimate the remaining time needed to reach my destination
	3 (4.43)	When the effect of road sign changes with time (parking hours, no turn on red, school zone,…)

The bolded situations are also selected by the novice drivers

the road to a sign. Safety related information (such as collision warnings) are thought to be the most important information for novice drivers—many of whom wanted to have the information when they need to parallel park.

5.3 Context Awareness Technologies Required to Detect the Representative Driving Situations

From the results obtained from each research stage, we defined the sets of driving situations at which representative in-car information items are valuable. Combining our results with those from prior studies [3], we can now identify which information to deliver at interruptible moments.

In the final step of this study we examined the availability of sensor-based technologies that care available to detect the important context factors that we derived. If a context factor is both important and not currently detectable, it points for the need for technologies to be developed in the future. For each driving situation, we listed currently available technologies and technologies that needs to be developed or improved (Table 4 for the driving situations of novice driver and Table 5 for that of experienced driver).

Table 4 Currently available context awareness technologies (technologies that needs to be further improved are shown in bold)

Driving situation	Context information to detect	Required technologies
SD1	Current location, current speed, road signs	GPS, OBD, **road sign detection**
SD2	Whether, road condition	Weather information, temp/humid sensors, **road condition detection**
SD3	Current location, the road ahead (downhill, curve…)	GPS systems
FG1	Destination, current location, distance/time needed to reach the destination	Fuel gauge
FG2	Current location, closest gas station, deals provided by the gas stations	GPS, fuel gauge
FG3	Vehicle status (HVAC, trunk …)	**Vehicle status**
FR1	Vehicle status (engine …)	OBD
FR2	Vehicle status (HVAC, trunk …)	**Vehicle status**
FR3	Current location, location of gas stations, deals provided by the gas stations	GPS, Internet services
EO1	Driver identified vehicle status (sound, vibration, smoke …)	Accelerometers, **smoke/sound sensors**
EO2	Destination, current location, distance/time to destination	GPS systems
EO3	–	–
RS1	Current location, empty parking lot, parking information (hours, fare…)	Parking info through Internet systems (exists in research)
RS2	Current location, location of the sign, speed limit information	GPS systems

(continued)

Table 4 (continued)

Driving situation	Context information to detect	Required technologies
RS3	Current location, condition of the road ahead (downhill, curve…), current speed	GPS systems
ETA1	Fuel gauge	Fuel gauge
ETA2	Current time, my schedule	Clock, scheduler apps
ETA3	Current location, destination location, route to the destination	GPS systems
FCW1	Current location, distance between car and other surrounding objectives	GPS, proximity sensors
FCW2	Weather, brightness, windshield transparency (by rain, snow, ETC)	Illuminometer, weather information on internet
FCW3	User state (secondary task)	**Secondary interaction detection**
BSW1	Handle movement, lane change information	OBD, lane detection system
BSW2	Handle movement, distance between car and other surrounding objectives	OBD, proximity sensors
BSW3	Current location, type of road ahead (interchange, merging …), driver route	GPS systems
RCW1	Handle movement, distance between car and other surrounding objectives	OBD, proximity sensors
RCW2	speed of other car, current speed, distance bet cars	Partially exists (warning if car in blind spot), speed of other cars
RCW3	–	–

Labels provided in the first column denotes driving situations by the type of in-car information and the rank: SD1 means ranked 1 driving situation in ‘Speedometer’ (FG: fuel gauge, FR: fuel reserve lamp, EO: engine overheating, RS: road sign)

Table 5 Currently available context awareness technologies (technologies that needs to be further improved are shown in bold)

Driving situation	Context information to detect	Required technologies
SD1	Current location, current speed, road sign	GPS, OBD, **road sign detection**
SD2	Current speed, current speeds of other cars	GPS, OBD, speed of other cars
SD3	Weather, road condition	Weather information, temp/humid sensors, **road condition detection**
FG1	Fuel gauge	Fuel gauge
FG2	Destination, current location, distance/time needed to reach the destination	GPS, fuel gauge
FG3	Current time, hours of the gas station	GPS, Internet services (Yelp, Waze, Maps)
FR1	Driver profile (driving history)	GPS history
FR2	Vehicle status (HVAC, trunk …)	**Vehicle status**

(continued)

Table 5 (continued)

Driving situation	Context information to detect	Required technologies
FR3	Current location, locations of gas station near by me, driver profile (which gas station did the driver frequently visited)	GPS, Internet services
EO1	Whether	Weather information, temperature sensors
EO2	Driver identified vehicle status (sound, vibration, smoke …)	Accelerometers, **smoke/sound sensors**
EO3	Destination, current location, distance/time needed to reach the destination	GPS
BW1	Type of the road (highway, downhill …), road condition	GPS services, **road condition**
BW2	Current speed, speeds of other cars, distance between cars	Distance between cars, **speed of other cars**
BW3	Driver identified vehicle status (sound, vibration, smoke …)	Accelerometers, **smoke/sound sensors**
CE1	Current location, time (late night), date (holidays)	GPS, Internet services, calendar
CE2	Driver identified vehicle status (sound, vibration, smoke …)	Accelerometers, **smoke/sound sensors**
CE3	Destination, current location, distance/time needed to reach the destination	GPS, navigation systems
CL1	Current time, current location, my schedule	GPS, scheduler apps
CL2	Destination, current location, distance/time needed to reach the destination	GPS, navigation systems
CL3	Road signs, traffic sign, driver profile (commute route)	**Road sign detection**

Labels provided in the first column denotes driving situations by the type of in-car information and the rank: SD1 means ranked 1 driving situation in 'Speedometer' (*FG* fuel gauge, *FR* fuel reserve lamp, *EO* engine overheating, *BW* brake warning, *CE* check engine lamp, *CL* clock)

6 Conclusion

We aimed to derive sets of driving situations for which representative in-car information items are valuable to drivers. By considering the effect of the driving experience on the value of in-car information, we recruited two groups of drivers (novice and experienced). Throughout the study, we used a series of surveys and interviews because drivers assess information value internally, which thus cannot be measured directly. To conduct this research, we first identified representative in-car information that is considered important in general driving contexts and identified context factors that significantly affect its value. By combining those selected context factors, we collected a variety of driving situations for which the value of the selected in-car information is high. Then, through a large scale online survey, we estimated value of in-car information in those driving scenarios.

The results obtained in each research steps helped us to better understand the value of information provided in-car. We now understand that drivers could require different types of information in similar situations depending on their experience. This finding helps us to select which information to provide to drivers at interruptible moments. Also, by identifying current technologies for the derived driving situations, we identified situations that are currently detectable (more specifically, which context factors are identifiable) and for which factors technologies need to be developed.

In our future work, we will develop an intelligent information delivery system aimed to deliver appropriate information at appropriate times, in an actual driving environment.

Acknowledgments This project is funded in part by Carnegie Mellon University's Technologies for Safe and Efficient Transportation, The National USDOT University Transportation Center for Safety (T-SET UTC) which is sponsored by the US Department of Transportation.

References

1. Simon, H.A.: Designing organizations for an information rich world. Comput. Commun. Public Interest, 37–72 (1971)
2. Ranney, T.A., Mazzae, E., Garrott, R., Goodman, M.J.: NHTSA driver distraction research: past, present, and future. Transp. Res. Center Inc., East Liberty, OH, Tech. Rep. (2000)
3. Kim, S., Chun, J., Dey, A.K.: Sensors know when to interrupt you in the car: detecting driver interruptibility through monitoring of peripheral interactions. In: Proceedings of SIGCHI Conference Human Factors in Computing Systems (CHI '15), Seoul, Korea (2015)
4. National Highway Traffic Safety Administration.: Distracted Driving 2011. Traffic Safety Facts. DOT HS 811 737. US Department of Transportation, Washington (2013)
5. Barr, L., Popkin, S., Howarth, H.: An Evaluation of Emerging Driver Fatigue Detection Measures and Technologies. Volpe Nat, Cambridge, MA (2009)
6. www.media.volvocars.com
7. www.emercedesbenz.com
8. Schneegass, S., Pfleging, B., Broy, N., Heinrich, F., Schmidt, A.: A data set of real world driving to assess driver workload. In: Proceedings of Automotive UI '13, pp. 150–157. ACM (2013)
9. Solovey, E.T., Zec, M., Perez, E.A.G., Reimer, B., Mehler, B.: Classifying driver workload using physiological and driving performance data: two field studies. In: Proceedings of CHI '14, pp. 4057–4066 (2014)
10. Kwahk, J., Han, S.H.: A methodology for evaluating the usability of audiovisual consumer electronic products. Appl. Ergon. **33**, 419–431 (2002)
11. Fastrez, P., Haué, J.-B.: Editorial: designing and evaluating driver support systems with the user in mind. Int. J. Hum. Comput. Stud. **66**, 125–131 (2008)

A Typology of In-Vehicle Eco-Driving Feedback

Angela Sanguinetti, Hannah Park, Suhaila Sikand and Ken Kurani

Abstract Eco-driving is a promising strategy for reductions in fossil fuel consumption and carbon emissions. Eco-driving is most frequently promoted via in-vehicle feedback. Eco-driving feedback studies demonstrate fuel economy improvements up to 18 %, but results are widely variable—partly due to the wide variation in feedback design. This paper addresses the need for a greater understanding of how variations in eco-driving feedback design are related to its effectiveness. We identified characteristics of feedback with implications for behavior change based on behavioral theory and evaluation of a large sample of in-vehicle eco-driving feedback interfaces. We developed a typology of in-vehicle eco-driving feedback interfaces based on these characteristics. We identified 15 distinct types of in-vehicle eco-driving feedback interfaces. We describe each feedback type and discuss implications for feedback design. Our typology provides a foundation for subsequent research to determine most effective feedback types for particular behaviors, drivers, and driving conditions.

Keywords Eco-driving · Green driving · Smart driving · Automotive interface

A. Sanguinetti (✉) · H. Park · S. Sikand · K. Kurani
Institute of Transportation Studies at University of California,
Davis 1605 Tilia, St. Davis, CA 95616, USA
e-mail: asanguinetti@ucdavis.edu

H. Park
e-mail: hnpark@ucdavis.edu

S. Sikand
e-mail: sksikand@ucdavis.edu

K. Kurani
e-mail: knkurani@ucdavis.edu

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_80

1 Introduction

The Monroney sticker posted on every new vehicle in the US states, "Actual results will vary for many reasons, including driving conditions, and how you drive and maintain your vehicle." Eco-driving is a means of strategically taking advantage of this variability by operating one's vehicle in the most efficient manner [1]; for example, maintaining an even driving pace and minimizing use of cabin heating and air conditioning. In this paper, we define eco-driving as behavior a driver engages in while driving that reduces the vehicle's fuel consumption and/or polluting emissions.

The most common strategy to promote eco-driving is the provision of feedback, i.e., information provided to the driver about the effects of driving behavior on fuel economy and/or emissions after the behavior occurs. Eco-driving feedback includes instantaneous and average fuel economy information, as well as information about how efficiently a driver performs more specific behaviors, such as accelerating. Eco-driving feedback may be provided via manufacturer-supplied instrumentation, or after-market devices and/or apps. Our focus in this paper is on the former, which we will refer to as in-vehicle eco-driving feedback for the remainder of the paper.

1.1 In-Vehicle Eco-Driving Feedback

No policies exist requiring manufacturers to provide eco-driving feedback, yet feedback systems of increasing variety are appearing in vehicles, especially hybrid (HEVs), plug-in hybrid (PHEVs) and electric vehicles (EVs). One reason for this differential attention is that fuel economy in efficient vehicles is actually more sensitive to driver behavior. Manufacturers have deployed many different designs, reflecting various driver behaviors and vehicle states. This wide variation could indicate a belief in competitive advantage or a lack of evidence-based design and consistent assumptions about human behavior. The rapidly increasing prevalence and complexity of in-vehicle displays and concern for driver distraction [2] suggest standardization of eco-driving feedback may be warranted in the near future.

1.2 Eco-Driving Feedback Studies

A recent comprehensive review of in-vehicle feedback interventions to promote eco-driving in conventional internal combustion engine (gas) vehicles (ICEVs) calculated average fuel use reduction of 5.6 %, ranging from −6.8 to 18.4 % [1]. Some of this variation in effectiveness is undoubtedly due to the extreme variation in feedback provided (e.g., numeric indicators, haptic pedal feedback). Relatively few of these studies compared different types of feedback, but those that did identified some characteristics of feedback that make it more or less effective. For example, feedback is more effective when metrics align with the driver's goals, e.g., to save money [3] and when feedback is adaptive to performance [4]. There have been some attempts to classify eco-driving feedback interfaces according to these and other important characteristics [5–8].

1.3 Previous Eco-Driving Feedback Typologies

Tulusan et al. [5] categorized eco-driving feedback based on the timing of its presentation and/or the duration of behaviors it reflects: (a) feedback on momentary driving behavior (real-time, reflecting instantaneous behavior), (b) accumulated feedback (reflecting accumulated behavior), and (c) offline feedback (provided outside the context of driving). They suggested the mode of interface for feedback on momentary driving behavior is typically ambient (e.g., conveyed via changing light intensity or color), yet the most common form is a numeric instantaneous fuel economy indicator.

Stillwater et al. [8] classified in-vehicle eco-driving feedback in electric and plug-in hybrid vehicles (together, PEVs) based on the specific behaviors reflected, the duration of those behaviors, and the metrics used (e.g., mileage or range): (a) current efficiency, (b) deceleration efficiency (including regenerative braking), (c) trip efficiency, and (d) range comparison. They describe trip efficiency and range comparison as useful for goal-setting: trip efficiency provides drivers with energy usage of a given trip and range comparison juxtaposes current estimated range with EPA-estimated range or distance to a programmed destination.

Extensive research for the United States Department of Transportation [6, 7] provides the most granular analysis of in-vehicle eco-driving feedback, which they termed FEDI (fuel economy driver interface). They analyzed in-vehicle feedback systems in terms of "components"—discrete visual elements within a given FEDI, classifying them first by mode of interface (intensity-changing light, representative pictures, graph, single dial, single bar, text, or other), then characterizing them in terms of the metric and duration of data presented (instantaneous fuel economy, trip average fuel economy, overall average fuel economy, and miles to empty). They did not consider the specific behaviors targeted by feedback or feedback timing.

These previous typologies are insufficient. Each is based on only a subset of important feedback characteristics, the selection and description of which is neither explicit nor grounded in behavioral theory. The current research develops a typology that distinguishes among the myriad of interfaces based on an understanding of the behavioral mechanisms by which eco-driving feedback affects driver behavior.

1.4 Theoretical Framework

Kluger and DeNisi [9] developed a useful theory for understanding the behavioral mechanisms involved in feedback interventions. According to Feedback Intervention Theory (FIT), “behavior is regulated by comparisons of feedback to goals or standards” (p. 259). There may be one or more standards to which feedback is compared, including comparisons to self (e.g., past performance) and comparisons to others. A discrepancy between feedback and its standard is called the feedback-standard gap. A key tenet is that only feedback-standard gaps receiving attention will be addressed; thus, a critical aspect of any feedback intervention is controlling the locus of attention.

Principles established by the science of behavior analysis [10] can also aid in understanding the underlying behavioral mechanisms of effective eco-driving feedback. In general terms, feedback may affect behavior by reinforcing or punishing particular responses, prompting behavior, or creating conditions of motivation (e.g., via comparison to social norms or past performance). These behavioral theories guided the development of our eco-driving feedback typology.

2 Methodology

2.1 Sampling Eco-Driving Feedback Systems

Our first goal was to obtain a representative sample of in-vehicle eco-driving feedback systems. We identified systems via previous research [5–8] and Internet searches using Google; search terms “eco-driving system”, “dashboard”, “multi-information display”, “instrument panel”, “instrument cluster”, “audio-navigation display”. After identifying a relevant system, we gathered further information from vehicle manufacturer websites, YouTube videos, and owner’s manuals. We focused on the US market, although two systems in our sample are unavailable in the US. When we discovered the same or similar systems in multiple models from the same manufacturer or multiple years of the same model, we included only one model or the most recent model, respectively. We identified 14 eco-driving feedback systems—all in alternatively-fueled vehicles (HEVs: Ford Fusion 2016, Honda Insight 2014, Lexus 300H 2014, Nissan Sentra 2014, Toyota Camry 2015, Volkswagen Jetta Hybrid 2013; PHEVs: Cadillac ELR 2014, Chevy Volt 2013, Toyota Prius 2014, Toyota Prius C 2015; EVs: Ford Focus 2015, Kia Ray 2014, Nissan Leaf 2013, Toyota RAV4 2013).

2.2 Defining the Feedback Stream

Eco-driving feedback systems can be composed of multiple distinct feedback types, therefore the system as a whole is not an appropriate unit of analysis. Manser et al. [7] articulated the feedback component as their unit of analysis based on mode of interface as the defining parameter (e.g., a graph is one component, a colored light is another). In some in-vehicle feedback systems, visual interface components are interdependent and complimentary, so separating them would compromise the validity of our analysis. We therefore defined a new unit of analysis: the feedback stream.

A feedback stream consists of at least one interface component corresponding to at least one driver behavior related to fuel consumption. Multiple interface components are considered the same feedback stream if they (a) are presented together, reflect the same behavior(s), and differ only in terms of data or design, or (b) if they reflect different behaviors but are similar in terms of feedback attributes and are visually integrated. We excluded interfaces with only forecast information (e.g., fuel remaining) or information specific to maintenance behaviors (e.g., check engine and tire pressure). We identified 116 distinct feedback streams among the 14 feedback systems.

2.3 Defining Behaviorally Relevant Feedback Attributes

Based on our review of the literature and examination of these 116 in-vehicle eco-driving feedback streams, we inventoried and defined in-vehicle eco-driving feedback attributes that have implications for behavior change according to FIT [9] and behavior analytic principles [10]. The attributes we identified correspond to ten feedback parameters that fall into three categories: data, timing, and design (Table 1).

Table 1 Feedback parameters and attributes

	Parameter	Attributes
Data	Behavioral granularity	Aggregate—all affecting mileage; all affecting power source. Specific—accelerating; cruising; decelerating; driving mode selection; cabin comfort (climate control, auxiliary electronics) [16]
	Data granularity	Low (≤3 levels). Mod. (4–10 levels). High (≥10 levels)
	Temporal granularity	Instantaneous. Accumulated—event within a trip; intervals during trip; trip level; tank/charge/reset; lifetime
	Metric	Fuel economy (MPG or mi/kWh). Environmental impact (CO_2 saved/emitted). Money ($ saved/spent). Distance traveled or range

(continued)

Table 1 (continued)

	Parameter	Attributes
		(mi). Power/energy required/regenerated (kW, %). Fuel (gallons). Points/score
	Feedback standard	Optimal zone. Score. Past performance (average, best). EPA-estimated mileage. Expected range
Timing	Immediacy	Immediate. Retrospective
	Frequency	Continuous. Discrete
Design	Mode of interface	Numeric. Histogram. Pie chart. Meter (bar). Pictorial. Light (color change and/or blinking). Movement. Haptic
	Gameful design	Yes. No
	Biophilic design	Yes. No

2.3.1 Data

In-vehicle eco-driving feedback data vary in terms of the behaviors reflected (behavioral granularity), the duration of behavior reflected (temporal granularity), and the magnitude of behavior required to produce changes (data granularity). More granular feedback on all these levels (i.e., reflecting specific behaviors, reflecting momentary behaviors, and sensitive to small changes in behavior, respectively) is useful for learning because it can function to reinforce or punish behavior [10] and the driver can more readily comprehend the connection between behavior and consequences.

Numeric feedback corresponds to a metric (e.g., dollars, MPG, miles), which can be framed positively or negatively (e.g., money saved or spent). Feedback expressed in meaningful metrics can increase drivers' awareness of consequences of their behavior, which may activate moral norms and motivate behavior [11]. Feedback is sometimes accompanied by a standard, or referent (e.g., optimal zone of acceleration). According to FIT [9], feedback-standard comparison is the mechanism by which feedback regulates behavior. Standards may be adaptive based on driver performance to support incremental improvement [4], i.e., the process of behavioral shaping [10].

2.3.2 Timing

Parameters of eco-driving feedback timing include immediacy and frequency. Immediacy refers to the immediacy of feedback with respect to the behavior it reflects. Feedback frequency refers to the frequency with which feedback is presented (e.g., continuously or at discrete times). Similar to granular feedback, immediate and frequent feedback is useful for training new behavior. Immediacy is critical for feedback to function as a reinforcement. Retrospective feedback may motivate drivers to engage with the feedback, but may be too far removed to function as reinforcement.

2.3.3 Design

Design attributes include mode of interface and whether the feedback employs gameful or biophilic design. Broad categories of mode of interface are visual, auditory, or haptic. Mode of interface has implications for feedback saliency and driver distraction; for example, ambient interfaces [12] are potentially less distracting compared to numbers or graphs [13]. Gameful design refers to the use of game design elements (e.g., points, levels, leaderboards, and badges) that may promote engagement with eco-driving feedback [14, 15]. Biophilic design [16] refers to the integration of elements of nature (e.g., plants, animals) to promote a sense of connection to nature. Biophilic design in energy feedback has been effective in promoting conservation behaviors [17].

2.4 Data Analysis

We coded the 116 feedback streams in terms of each parameter of attributes (Table 1). Attributes within each parameter were not mutually exclusive, with the exception of gameful and biophilic design which were coded as binary (yes or no). We sorted the feedback streams by behavioral granularity to give it more weight in the two-step cluster analyses, which we performed with SPSS software. In an iterative process, we refined feedback attribute definitions and codes, ran multiple cluster analyses, and adjusted based on prioritization of target behaviors. An example of our prioritization of target behaviors is that we did not include driving mode selection indicators (Fig. 3) in statistical cluster analyses; we judged them to be a unique type of feedback based on the specific target behavior (pressing an ECO mode button). This approach has aspects of taxonomic classification (hierarchical approach of prioritizing target behaviors) and some aspects of typological classification (all other attributes were weighted equally).

3 Results

Our analysis yielded 15 types of in-vehicle eco-driving feedback, including basic interfaces for alternatively-fueled vehicles (Figs. 1, 2), feedback specifically targeting accelerating, cruising, and/or decelerating (Figs. 4, 5, 6), and feedback reflecting fuel economy, inclusive of an aggregate of behaviors (Figs. 7, 8).

Fig. 1 Power Meter. The power meter in a hybrid or electric vehicle is analogous to a tachometer in an ICEV, but often includes an optimal zone standard and regenerative braking feedback

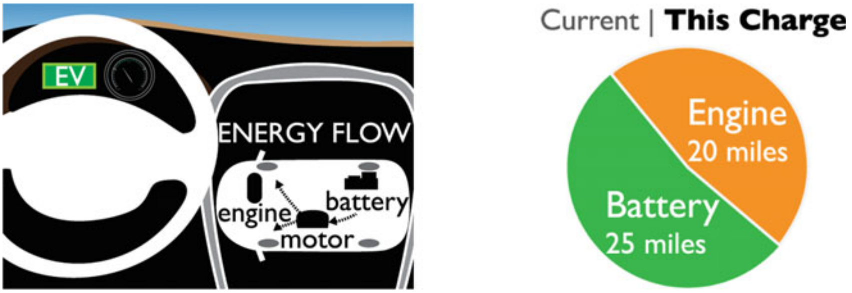


Fig. 2 Hybrid system indicators. Hybrids indicate the drive train(s) currently in operation with a single lamp or pictorial diagram (*left*), as well as a breakdown of current power demand or accumulated miles traveled with each fuel source (*right*)

Fig. 3 Driving mode selection indicator. The driver can press a button to select a driving mode that will improve fuel economy (ECO mode) or control fuel sources (EV mode), reflected by a single lamp with icon or colored glow surrounding elements of the instrument cluster

Fig. 4 Eco-accelerometer (*left*) and eco indicator (*right*). Feedback on accelerating and cruising speed is provided by a meter (Eco-accelerometer) or single lamp (Eco Indicator)

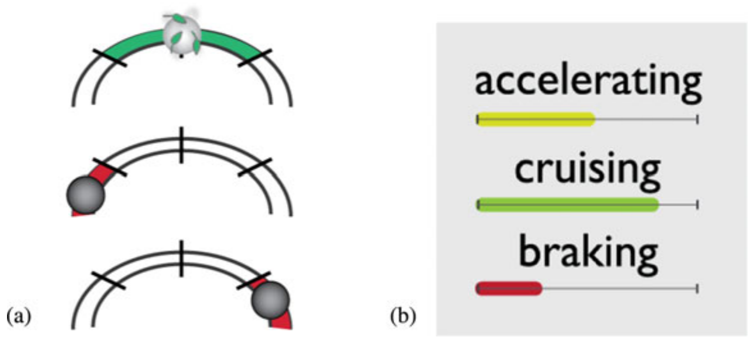


Fig. 5 **a** Pedal feed (*left*). Efficient accelerating, cruising, and decelerating are reflected, with the center as the feedback standard. **b** Eco-driving coach (*right*). Behavior-specific feedback is provided either in real-time and/or at the end of each accelerating, cruising, and braking event

Fig. 6 **a** Haptic pedal (*left*). Haptic feedback was found in only one on-market vehicle, not available in the US. Pressure exerted by the driver on the throttle is met with corresponding counter-pressure to teach the driver most efficient throttle position during acceleration events. **b** Regenerative braking feedback (*right*). A display dedicated to regenerative braking feedback depicts total percentage of lost mechanical energy recaptured during a braking event

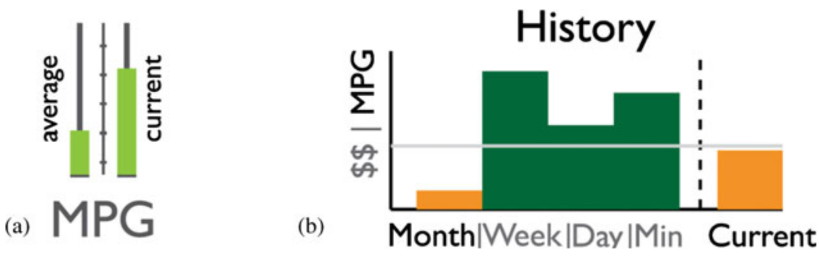


Fig. 7 **a** Mileage meter (*left*). Meters are a common mode of interface for current and average aggregate efficiency, sometimes presented together as pictured here, where average MPG serves as a feedback standard for current performance. **b** Efficiency history (*right*). Past performance is reflected in a histogram with intervals of some duration (e.g., minute, day), providing a standard for recent or current performance

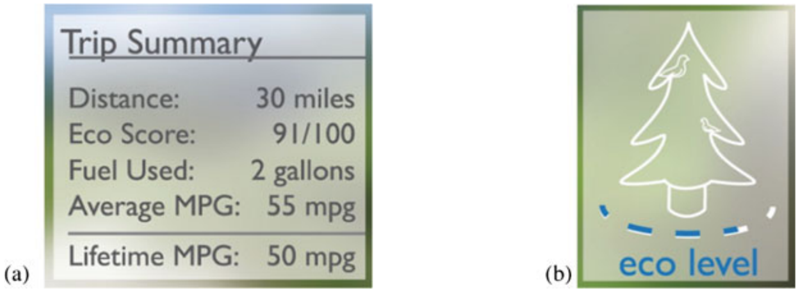


Fig. 8 **a** Summary statistics (*left*). Eco-driving behavior is reflected in numeric indicators of mileage and other metrics, such as distance traveled, fuel consumed, or efficiency score. **b** Biophilic rewards (*right*). Eco-driving behavior is rewarded with the accumulation of nature imagery, often paired with graduated levels or scores

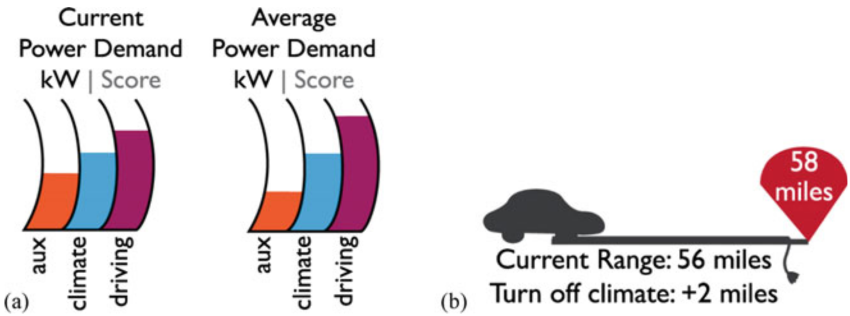


Fig. 9 **a** End-use analysis (*left*). Current and average power demand for different end-uses/driver behaviors may be presented in terms of power demand or an efficiency score. **b** Relative range (*right*). A remaining range estimate based on current driving style and/or use of climate settings may be presented juxtaposed with a feedback standard such as the potential maximum range, distance to destination, and/or additional range gained or lost if climate settings are changed

4 Discussion

This typology allows us to compare within and among feedback types and consider implications for maximizing the effectiveness of feedback. In this discussion, we focus on the implications of six feedback parameters, as manifested in identified feedback types, on three general behavioral mechanisms underlying feedback effectiveness: attention, learning, and motivation. We find that mode of interface has significant implications for driver attention; behavioral granularity for driver learning; and metrics, standards, gameful design, and biophilic design for driver motivation.

4.1 Salient Feedback for Driver Attention

The Haptic Pedal (Fig. 6a) is the only non-visual mode of interface we found and it was only present in one vehicle, not available in the US. Research suggests that Haptic Pedal can promote eco-driving with less potential for driver distraction than visual feedback [13]. One might hypothesize that it is also less susceptible to a novelty effect whereby feedback is attended to initially then tuned out. The same may be true for other non-visual modes of feedback, such as auditory feedback, analogous to the beep of a seat belt reminder. Auditory feedback is altogether absent from the on-market systems we analyzed, but examples can be found in after-market apps. It seems that non-visual eco-feedback warrants further consideration from vehicle manufacturers.

There may sometimes be a trade-off between salience and data granularity, which has implications for feedback sensitivity to behavior. For example, numeric feedback is more granular but less salient than a meter, which is more granular but less salient than a colored light. One strategy we saw to maximize salience and sensitivity in visual feedback was integration of colored lights and blinking or spinning movement with a meter (Figs. 4 and 5a).

4.2 Precise Feedback for Driver Learning

In some cases, we had trouble determining which eco-driving behaviors were reflected by Eco Indicators (Fig. 4). For example, some manufacturers note that their Eco Indicator reflects only acceleration, whereas others claim it reflects both acceleration and deceleration, and for yet others we were unable to rule out whether they included all engine loads, such as climate control and auxiliary electronics. Given the similarity of the interface and the ambiguity from the driver's perspective, we considered all these as the same type of feedback. It is notable that even owner's manuals do not always sufficiently define the eco-driving behaviors targeted by their feedback systems. Lack of precision in the behavioral granularity of feedback is sub-optimal for learning.

Other types of feedback do a much better job of precisely conveying the target behaviors. For example, Eco-driving Coach (Fig. 5b) targets three key driving behaviors: accelerating, cruising, and braking. Such precision supports learning. End-use Analysis (Fig. 9a) and Relative Range (Fig. 9b) disaggregate feedback in a similar way, e.g., calling out the impacts of climate settings, use of auxiliary electronics, and/or driving behaviors (an aggregate of accelerating, cruising, braking). However, some cases of End-use Analysis showed "Motor" and "Other" as end-uses, which are less useful in teaching a driver what to do compared to "Driving Style" and "Aux" (for the latter, "Accessories" or "Electronics" might be more meaningful still). Climate settings and use of electronics are unique targets in that they are available to passengers as well as drivers.

In cases where feedback on aggregate behavior reflects actual fuel economy, a myriad of behaviors become relevant, i.e., those classified by Kurani et al. [1] as driving behavior (accelerating, cruising, decelerating, waiting, parking, and driving mode selection), cabin comfort (behaviors related to thermal comfort, communications, and entertainment), trip planning (e.g., route selection, timing), load management (interior and exterior cargo weight and aerodynamics), fueling (e.g., fuel selection), and maintenance (e.g., tire pressure, engine oil). This lack of precision could be sub-optimal for learning and motivation as the effects of any one behavior are washed out. Aggregate feedback could also be perceived as specific when it is highly sensitive to a particular behavior, e.g., accelerating; when that is the case, that particular behavior might be the only behavior affected by the feedback.

4.3 Meaningful Feedback for Driver Motivation

Efficiency expressed as a number, meter, or graph, is typically in terms of MPG or kWh/mi, but several tactics are employed to make metrics more meaningful for some drivers. For example, points, scores, levels, and progress bars are applications of gameful design (e.g., Fig. 5b, Fig. 8a and b). Translating metrics is also useful when the actual metric is unfamiliar, abstract, or poorly understood; e.g., power demand can be displayed as percentage points or score rather than kilowatts (e.g., Fig. 9a).

One feedback system displayed money mileage ($/mi). According to Dogan et al. [18], if financial costs are negligible this could actually be a disincentive to eco-driving. This suggests that money mileage may be most effective when reflecting accumulated behavior (e.g., weekly or monthly) and/or juxtaposed with a feedback standard that reflects accumulated behavior (Fig. 7b). On repeated commutes, fuel cost on prior trips could serve as a meaningful feedback standard.

In the case of Biophilic Rewards (Fig. 8b), metrics are sometimes abstracted away altogether—replaced by imagery. Biophilic Rewards range from a more quantitative, gameful design that includes numeric or metered scores or levels, to a more qualitative, aesthetic design where the only quantitative data is number of trees, leaves, or flowers. The visualization may be labeled as fuel efficiency, remaining range, or CO_2 reduction. Aesthetic designs that prompt reactions similar to reactions to actual nature and use preferred nature imagery should be most successful in priming consideration and care for the natural world, which might motivate eco-driving.

4.4 Limitations

Stated implications and recommendations regarding feedback types should be considered in a broader context when applied to feedback design and related policy-making. For example, we did not consider feedback location (e.g., instrument cluster, center console, or head-up display), which has implications for saliency and distraction. Implications of the holistic composition of a feedback system with multiple streams should also be considered in future research. Finally, our identified feedback types are not entirely orthogonal, but we contend that our method of classification yielded a typology that best reflects the current distinctions among in-vehicle feedback systems.

4.5 Conclusion

This research presents an in-vehicle eco-driving feedback typology based on behavioral theory. Our discussion of 15 distinct types of feedback points to opportunities for optimized and novel designs. We highlighted the importance of saliency, precision, meaningful metrics, and multiple feedback standards to promote eco-driving learning and motivation. Future research is needed to systematically evaluate these feedback types in relation to specific eco-driving behaviors, driver characteristics, and driving conditions in order to determine the most effective configurations.

Acknowledgments This research was supported by the National Center for Sustainable Transportation (NCST), University of California Multicampus Research Programs and Initiatives (MRPI), and University of California, Davis, Research Investments in the Sciences and Engineering (RISE).

References

1. Kurani, K.S., Sanguinetti, A., Park, H.: Actual results may vary: a behavioral review of eco-driving for policy makers. A white paper for the National Center for Sustainable Transportation (2015)
2. Rouzikhah, H., King, M., Rakotonirainy, A.: Examining the effects of an eco-driving message on driver distraction. Accid. Anal. Prev. **50**, 975–983 (2013)
3. Kurani, K.S., Stillwater, T., Jones, M.: Ecodrive I-80: A Large Sample Fuel Economy Feedback Field Test Final Report. Institute of Transportation Studies Report ITS-RR-13-15, University of California, Davis (2013)
4. Wada, T., Yoshimura, K., Youhata, H., Tomiyama, K.: Proposal of an eco-driving assist system adoptive to driver's skill. In: 14th International IEEE Conference on Intelligent Transportation System, pp. 1880–1885. ITSC (2011)

5. Tulusan, J., Soi, L., Paefgen, J., Brogle, M., Staake, T.: Eco-efficient feedback technologies: which eco-feedback types prefer drivers most?. In: 12th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, pp. 1–8. WoWMoM, Lucca (2011)
6. Jenness, J.W., Singer, J., Walrath, J., Lubar, E.: Fuel Economy Driver Interfaces: Design Range and Driver Opinions (Report on Task 1 and Task2). DOT HS 811 092, National Highway and Traffic Safety Administration, Washington, D.C. (2009)
7. Manser, M.P., Rakauskas, M., Graving, J., Jenness, J.W.: Fuel economy driver interfaces: develop interface recommendations (Report on Task 3). DOT HS 811 319, National Highway Traffic Safety Administration, Washington, D.C. (2010)
8. Stillwater, T., Davies, J., Kurani, K.S.: EV and PHEV Energy HMI Function and Usability. Unpublished report for CEC, Public Interest Energy Research (PIER) Program (2013)
9. Kluger, A.N., DeNisi, A.: The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychol. Bull. **119**, 254–284 (1996)
10. Cooper, J.O., Heron, T.E., Heward, W.L.: Applied Behavior Analysis, (2nd. Ed.). Pearson, UK (2007)
11. Ahmed, S., Sanguinetti, A.: OBDEnergy: Making metrics meaningful in eco-driving feedback. In: Human Computer Interactions International Conference Proceedings, pp. 395–401. Springer International Publishing, Switzerland (2015)
12. Pousman, Z., Stasko, J.: A taxonomy of ambient information systems: four patterns of design. In: Proceedings of the Working Conference on Advanced Visual Interfaces, pp. 67–74 (2006)
13. Seewald, P., Ahlström, C., Aleksic, M., Fors, C., García, E., Hibberd, D., Iviglia, A., Jamson, S., Jamson, H., Josten, J., Kircher, K.: Deliverable D13.2.: Results of HMI and feedback solutions evaluations. Version 10, ecoDriver (2013)
14. Zichermann, G., Cunningham, C.: Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps. O'Reilly Media, Sebastopol (2011)
15. Diewald, S., Möller, A., Roalter, L., Stockinger, T., Kranz, M.: Gameful design in the automotive domain: review, outlook and challenges. In: Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 262–265. ACM (2013)
16. Kellert, S.R., Heerwagen, J., Mador, M.: Biophilic Design: The Theory. Science and Practice of Bringing Buildings to Life. Wiley, New York (2011)
17. Loeb, L., Loeb, G., Tice, E., Tregubov, T.: Emotionally engaging students to change behaviors and conserve resources: unplug or the polar bear gets it! Int. J. Environ. Cult. Econ. Soc. Sustain. **6**(2) (2010)
18. Dogan, E., Bolderdijk, J.W., Steg, L.: Making small numbers count: environmental and financial feedback in promoting eco-driving behaviours. J Consum. Policy. **37**, 413–422 (2014)

Part XII
Road and Rail—Education and Hazard Perception

Education and Training of Problematic Drivers and Drivers of the Integrated Emergency System

Pavel Řezáč, Veronika Kurečková, Petr Zámečník and Aleš Zaoral

Abstract The study focuses on outputs specifics on how to implement education and training of specific groups of drivers in the Czech Republic (problematic drivers and drivers of integrated emergency system). Both groups are among the risk group of drivers. First group of drivers lost their driving license as a result of repeated traffic offences. A group of drivers to practice driving problem is prone to relapse of problem behavior in traffic. Targeted interventions contribute significantly to reducing recidivism risk behavior and increase traffic safety. Methodology for drivers of vehicles of integrated emergency system also includes tools increasing traffic safety. These drivers are under pressure (time etc.). In the process of teaching and training of future rescuers, policemen and firemen not issue driving devoted ample space (only marginally or not at all). Results of this study shows, how can we improve education and training of these groups of drivers.

Keywords Education of problematic drivers · Education of the drivers of the integrated emergency system · Rehabilitation courses for problematic drivers

1 Introduction

Human factor is involved in 90 % of accidents [1]. Reason for this percentage is not a driver's mistakes but as a deliberate breaking rules. Violations of traffic laws are much more common than, for example insufficient information processing [2]. Lack or absence of negative feedback reinforces beliefs low probability of accidents [3] and subjectively felt a greater chance of avoiding a traffic accident than others [4]. Lack of feedback, along with a perceived improvement in driving skills during the first years of management helps to change the perception of the consequences of risky behavior [5], overestimation of their own driving skills [6] and a negative

P. Řezáč (✉) · V. Kurečková · P. Zámečník · A. Zaoral
CDV—Transport Research Centre,
Lisenska 33a, Brno 636 00, Czech Republic
e-mail: pavel.rezac@cdv.cz

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_81

impact on risk perception [7]. A big step in the right direction was the introduction of a demerit point system that needs feedback can provide. It is shown that the introduction of a demerit point system leads to a reduction in the number of accidents and a large reduction in traffic accidents [8], however, step improvement of traffic safety after the first year of operation reduces and has a lower effect on mortality on highways [8]. Although the effectiveness of the points system generally highest in over one to two years of implementation [9], with the use of seat belts and child seats has a long term effect, not only to drivers but also to the passenger [10]. This is very important as the use of seat belts and child seats is the most effective way to prevent serious injuries and deaths in traffic accidents [11]. The demerit point system thus has two basic benefits. Longterm positive effect on the reduction of certain risk behaviors and identify those drivers who proves unreliable and dangerous.

Need to focus on regular training and teaching of these drivers is therefore evident. Problematic drivers occupy the Czech population is relatively small proportion, but the danger is very high. It is estimated that in about 25 % of accident committed by 1 % of the most problematic drivers. This group consists of drivers who repeatedly don't respect traffic rules (they removed a license on the basis of more minor offenses, or even cause a traffic offense). A group of drivers with problematic driving practice usually repeat their problematic behavior in traffic. It is thus necessary to solve, how to prevent the repeatedly recurrence of committing traffic offences. The other group of drivers is the drivers of vehicles with right of way. This group consists of drivers of integrated rescue system. These drivers are under huge pressure (reduction of the time arrival, taking into account the behavior of other road users when driving a car with activated beacon, etc.). Besides high demands on the driver's competence they are also forced to cope with the enormous acute and chronic stress. Continuous improvement of their driving competence is crucial not only for safety but also to increase efficiency and productivity. To meet the high demands on professional drivers of vehicles with right of way should contribute to an effective system of continuous education and training, which is currently in the Czech Republic unsystematic. As a contribution, we propose a form of education and training for problem drivers and drivers of the integrated rescue system. Taking account of the broader context of GDE (guidelines for driver's education) [12]. The internationally recommended driver education model takes into account several levels: 1. driver's ability to control the vehicle; 2. Ability to reflection traffic and maneuvering in traffic; Area 3 motives and path planning; 4. Lifestyle and aspects affecting driving behavior (culture, age, social status, etc.). The model reflects not only the actual level of skills and knowledge of drivers, but often neglected area of their attitudes and lifestyles compatible with safe driving.

2 Data Collection

Data were collected from both groups of drivers (problematic drivers and drivers of the integrated rescue system). In the first part we were conducted focus group, which focused on the detection of problematic situations in traffic. These were the areas with which the driver frequently encounters. It was so identified several areas that were named as problematic. Subsequently drivers were sampled traffic psychological test battery VTS (Vienna Test System)—intellect, perception test, attention and concentration, determination test, peripheral perception (among drivers integrated rescue system), a reaction test, aggressive driving behavior questionnaire personality traits related to management, test of Viennese tendency to take risks in traffic).

3 Research Sample

The survey included a total of 30 drivers integrated rescue system and 53 problematic drivers. Drivers of integrated rescue system were included in the research based on the implemented training for drivers of integrated rescue system. Problem drivers were included in the research on the recommendation of staff of the Probation and Mediation Service, or order of a court.

4 Methods

Within traffic psychological batteries were used in these tests. (a) Perception test focuses on diagnosing the extent of visual perception and perception speed. Stimulus material consists of photographs depicting various traffic situations. (b) The test of attention and concentration is built on the basis of the comparison between the figures and their conformity. The test is based on the theory Reulecky attention and concentration. (c) Determination test is aimed at monitoring the reactions of test persons in a stressful situation. The task of the person being tested is to quickly and accurately respond to visual and audio cues in a particular way (by pressing the corresponding keys and pedals). The test measures the resistance to stress, attention deficit, and the reaction time. (d) The test is intended for the diagnosis ability to perceive and process the peripheral image information. His administration needed special panels manned plenty of LEDs and ultrasonic sensor for accurate detection of the position of the head of the tested person. (e) Reaction test measures reaction time to visual and auditory stimuli. In addition to measuring the reaction time of the test is aimed at diagnosing attention, the ability to suppress faulty or inadequate response, assessing the level of alertness and focused attention. (f) Test of aggressive driving behavior through self-assessment determines the subjective level of risk behavior in traffic situations. The assay is characterized in that it allows comparing the behavior of the examined person in normal situations with its behavior in stressful situations. (g) Multidimensional Personality Questionnaire detecting the level of those personality characteristics that significantly affect a person's behavior in traffic situations (ability to self-control, emotional stability, accountability, etc.). (h) Objective personality test designed to measure specific tendencies toward risky behavior in traffic situations. This test is highly resistant to deliberate distortion of the above result from the examined person.

5 Results

On the basis of focus group and the results of the traffic psychological examination of drivers were designed way of education and training for problematic drivers and drivers of integrated rescue system. The way of the best training course for drivers of integrated rescue system involves training emergency braking and training on the polygon and simulator training. Training for problem involves driving simulator training and inclusion in a rehabilitation program for problematic drivers. Below we will specificity of drivers training on the simulator and the specifics of the rehabilitation program for drivers. As part of a focus group were described risk situations in traffic that have been programmed into the scenarios on simulator of the 3rd generation. Simulator lets tilt and vibration in the cabin. The platform offers 6 degrees of movement. The equipment also includes a dynamic mathematical model for cars and trucks, which defines their physical behavior in traffic. Thanks to the possibility of modifying the mathematical models can be selected physical properties of any type of vehicle, thus ensuring its adequate response analogous to a real traffic environment. It is also possible to select the type of gearbox—automatic or manual, driving modes with different number of routes and difficulty levels (e.g. a test drive, city traffic, extra urban driving and motorway driving or riding in a non-operating terrain). It is also possible to set different road surfaces, optional adhesion in the daytime, at night, in fog, rain and other weather conditions, including wind strength and direction. Simulator software provides the flexibility to modify the virtual traffic environments. Simulator offers variable traffic signs, interaction with other road users, various types of traffic behavior of other road users, critical scenarios that can be run automatically or manually instructor who can actively intervene in the current scene with the help of so called a virtual subscriber. Simulator also offers the possibility of reevaluating traveled routes including a comprehensive record of adequate variables of experimental ride. Based on analysis of risk traffic situation four scenarios have been created.

(a) *Urban scenario* and the behavior of other road users will be *defined as usual easygoing*. In this scenario, other traffic participants behave normally; most vehicles will observe traffic rules and with the active beacon blue car with the right of way adjust his driving lane or releasing slowdown. Even here there are slightly non serious risk situation analogous to normal traffic flow.

(b) *Rural scenario* and the behavior of other road users will be defined *as usual easygoing*. In this scenario, other traffic participants behave normally; most vehicles will observe traffic rules and with the active beacon blue car with the right of way adjust his driving lane or releasing slowdown. Even here there are slightly non serious risk situation analogous to normal traffic flow. In comparison with the urban traffic environment, there appear wild animals, which in this "normal conflict free" scenario will occur in close proximity to the roadway without jump directly into the road.

(c) *Urban scenario* and the behavior of other road users will be defined *as the risk of conflict*. Under this scenario, the driver will be exposed to the risk situation as unexpected (ex. pedestrian or child jump into the road, sudden braking of the vehicle before the driver, giving way, crossing the intersection on a red light, traffic congestion, which will be the driver with active blue beacon forced to use area off the road or driving in the opposite direction, etc.).

(d) *Rural scenario* and the behavior of other road users will be defined is defined *as the risk of conflict*. In this scenario, the driver will be exposed to risky situations such as unexpected pedestrian or child jump in the road, sudden braking of the vehicle before the driver, giving way, crossing the intersection on a red light, traffic congestion, which will be the driver with active blue beacon forced to use space outside road or driving in the opposite direction like in comparison with urban traffic environment, there occurs wildlife that is in this "risk of conflict" scenario will occur in close proximity to the roadway and into the road also runs.

These situations should be used for drivers training. This simulators part of training is common to both groups of drivers (problematic drivers and drivers of integrated rescue system). Subsequent driver training has been different. For drivers of integrated rescue system is appropriate to implement training aimed to increasing the ability of control vehicles (training emergency braking, etc.). In contrary the training for problematic drivers focuses on group psychotherapy. During this rehabilitation program should be solved a few areas. Firstly the analysis of the offense (the most serious traffic violations or accidents of participants or offenses that were the immediate cause of the withdrawal of driving license) should be solved. Further areas is using of alcohol and substance abuse with efforts to separate the use of driving. It is also appropriate to take the perceived risks in transport and self-perception of the consequences of accidents from the perspective of the victims of these accidents. Other topics should be—tiredness and stress while driving and creating personal strategies to avoid recurrence of problem behavior while driving.

6 Discussion

On the basis of the research show the different education needs for groups of drivers (integrated rescue system and problem drivers). Rehabilitation programs are effective and well tested tool for reducing recidivism of problem drivers. They are intended primarily for drivers who repeatedly violate traffic rules. Generally, rehabilitation programs can be defined as "the systematic measures for drivers who violate traffic rules—especially drivers under the influence of alcohol and drugs and the driver who crosses considerably speed limits" [13]. The main objective of these programs is to achieve a change in attitudes towards traffic regulations and to improve social responsibility during driving. Drivers, who attended the program, usually do not have problems with driving motor vehicles of traffic situations, but have problems in planning trips (e.g. With regard to their alcohol consumption), responsibility, risk perception and access to transport regulations or general rules. These areas would greatly affect how drivers perceive traffic rules and other safety recommendations [14–16]. Most drivers believe that the likelihood of an accident or other adverse events in transport is they themselves smaller than the other. This unrealistic optimism or comparative optimism is typical for driving behavior [5, 17]. Most drivers consider their driving abilities as above average [5, 17, 18]. The rehabilitation courses accented these topics. Suitable element in driver training also appears to use the simulator for driver training. This simulator should accent the problematic situations during driving.

Acknowledgments 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. Najeeb, M.A.: A study of the psychological factors that influence the rule violation behavior of drivers. In: Conference Roadsafety in a Globalised and More Sustainable World—Current Issues and Future Challenges (2012)
2. Reason, J.T., Manstead, A.S.R., Stradling, S.G., Baxter, J.S., Campbell, K.: Errors and violations on the road: a real distinction? Ergonomics **33**(10–11), 1315–1332 (1990)
3. Fuller, R.: Behavior analysis and unsafe driving: warning learning trap ahead! J. Appl. Behav. Anal. **24**(1), 73–75 (1991)
4. Cestac, J., Paran, F., Delhomme, P.: Young drivers' sensation seeking, subjective norms, and perceived behavioral control and thein roles in predicting speeding intention: how risk-taking motivation evolve with gender and driving experience. Saf. Sci. **49**(3), 424–432 (2011)
5. Delhomme, P., Verlhiac, J.-F., Martha, C.: Are drivers' comparative risk judgments about speeding realistic? J. Saf. Res. **40**(5), 333–339 (2009)
6. Whissell, R., Bigelow, B.: The speeding attitude scale and the role of sensation seeking in profilig young drivers at risk. Risk Anal. Int. J. **23**(4), 811–820 (2003)

7. Rosenbloom, T.: Risk evaluation and risky behavior of high and low sensation seekers. Soc. Behav. Pers. **31**(4), 375–386 (2003)
8. Farchi, S., Molino, N., Giorgi Rossi, P., Borgia, P., Krzyzanowski, M., Dalbokova, D., Kim, R.: Defining a common set of indicators to monitor road accidents in the European Union. BMC Public Health. **6**, 183 (2006)
9. 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. Accidents Anal. Prevent. **42**(4), 1310–1319 (2010)
10. Zambon, F., Fedeli, U., Visentin, C., Avossa, F., Brocco, S., Spolare, P.: Evidence-based policy on road safety: the effect of the demerit points system on seat belt use and health outcomes. Public Health **23**, 93–113 (2007)
11. Waller, P.F.: Challenges in motor vehicle safety. Annu. Rev. Public Health **23**, 93–113 (2002)
12. Hatakka, M., Keskinen, E., Gregersen, N.P., Glad, A., Hemetkoski, K.: From control of the vehicle to personal self-control; broadening the perspectives to driver education. Transp. Res. Part F **5**, 201–215 (2002)
13. Bartl, G.: EU-Project Andrea. Analysis of Driver Rehabilitation Programmes. KFV, Wien (2002)
14. Helweg-Larsen, M., Shepperd, J.A.: Do moderators of the optimistic bias affect personal or target risk estimates? A review of the literature. Pers. Soc. Psychol. Rev. **5**, 74–95 (2001)
15. Rundmo, T., Iversen, H.H.: Risk perception and driving behaviour among adolescents in two Norwegian counties before and after a traffic safety campaign. Saf. Sci. **42**, 1–21 (2004)
16. Ulleberg, P., Rundmo, T.: Personality, attitudes and risk perception as predictors of risky driving behavior among young drivers. Saf. Sci. **41**, 427–443(2003)
17. Harré, N., Susan, F., O'Neill, M.: Self-enhancement, crash-risk optimism and the impact of safety advertisements on young drivers. Br. J. Psychol. **96**, 215–230 (2005)
18. Horswill, M.S., Waylen, A.E., Tofield, M.I.: Drivers' ratings of different components of their own driving skill: a greater illusion of superiority for skills that relate to accident involvement. J. Appl. Soc. Psychol. **34**, 177–195 (2004)

Development and Trial of an Item Set for Testing Traffic Perception and Hazard Avoidance

Annika Dreßler, Bianca Bredow, Lars Rößger, Mathias Rüdel and Dietmar Sturzbecher

Abstract We describe the development of a set of items intended to measure a broad range of traffic perception and hazard avoidance skills for use in diagnosis of skill level in driver training and testing. The item set involves simulated dynamic traffic scenarios and contains three different task types. Type 1 relates to basic abilities of traffic observation, type 2 additionally requires the anticipation and identification of potential hazards, and type 3 further requires a decision (go/no-go) to act in order to prevent a potential hazard. We introduce the design of a training study with learner drivers in which the item set is currently being trialed. Finally, we give information about the status of assessment, intermediate results and further plans to evaluate the item set towards the development of a hazard perception test for the German system of novice driver preparation.

Keywords Hazard perception · Young drivers · Novice drivers · Driver licencing

A. Dreßler (✉) · B. Bredow
Institute for Prevention and Traffic Safety IPV, Staffelder Dorfstraße 19, 16766 Kremmen, Germany
e-mail: annika.dressler@ipv-ok.de

B. Bredow
e-mail: bianca.bredow@ipv-ok.de

L. Rößger · M. Rüdel
TÜV|Dekra Arge Tp21, Wintergartenstraße 4, 01307 Dresden, Germany
e-mail: lars.roessger@argetp21.de

M. Rüdel
e-mail: mathias.ruedel@argetp21.de

D. Sturzbecher
University of Potsdam, August-Bebel-Straße 89, 14482 Potsdam, Germany
e-mail: dietmar@sturzbecher.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_82

1 Introduction

1.1 The Impact of Traffic Perception and Hazard Avoidance Skills

Numerous studies indicate that young novice drivers are at an elevated risk of getting involved in road accidents [1–3]. Driver inexperience, e.g. lack of crucial skills, and age-related factors, such as youth-specific risk-taking patterns, have been identified as two key factors for the increased accident rate in young novice drivers. Analyses focusing on a separation of experience- versus age-related effects [4, 5] revealed that the initial crash risk drastically declines during the first months of solo driving and that this decrease is primarily attributable to effects of experience. Multivariate analyses indicate that the gains in experience within the first year of solo driving lead to a 35 % decrease in the accident rate [6].

With respect to effects of experience, a particularly crucial role is played by drivers' skills related to traffic perception and hazard avoidance [7, 8]. Accordingly, in several countries, instructional concepts for improving these skills in driver training have been implemented. Furthermore, a number of countries have developed test concepts ("hazard perception tests") to assess candidates' abilities concerning traffic perception and hazard avoidance and have made these tests a part of the statutory exams for acquiring a driver's licence [9, 10]. In the context of these promising approaches in driver training and licencing, however, there is still ongoing scientific discourse on the underlying cognitive components of traffic perception and hazard avoidance, and beyond, on suitable approaches to measuring these components.

1.2 Aims of the Current Study

The research presented here aims at making a contribution to the enhancement of novice drivers' road safety by advancing concepts for teaching and testing competencies of traffic perception and hazard avoidance. For this purpose, at a first step, published hazard cognition models were comparatively analyzed in order to derive a comprehensive description of relevant competencies (demand components) in the area of traffic perception and hazard avoidance. At a second step, an educational concept was developed for imparting these competencies in the course of regular driver training. Finally, at a third step, appropriate task formats and task contents to measure relevant skills were identified, and, on this basis, a specific set of test items was devised. The item set is currently being trialed in a pilot study with learner drivers: Progress in traffic perception and hazard avoidance skills is monitored in a control group receiving conventional driver training and an experimental group whose driver training is enriched by units on traffic observation and hazard perception. Findings from the study are expected to provide deeper insights for the future development of a reliable and valid testing tool.

2 Structural Components of Traffic Perception and Hazard Avoidance

The ability to adequately perceive traffic situations and efficiently avoid hazards involves several cognitive processes, which have been outlined in a number of hazard cognition models. Based on a comparative analysis of recent models, we identified eight specific demand components involved in traffic perception and hazard avoidance [11]. Firstly, the demand component of *observation* refers to the whole process of acquiring visual information. The limited capacity of cognitive and temporal resources calls for efficient observation strategies ("scanning" of the environment, use of mirrors), so as to allocate available resources to those situational attributes with a high information content concerning hazard avoidance. Secondly, the component of *localization* denotes the registration of relevant objects within the traffic environment based on ambient vision and the predetermination of worthwhile target spots to allocate foveal attention [12]. Thirdly, *identification* comprises the (mostly foveal) visual processing of objects, their comparison to and matching with driver's stored mental representations. These three basic demands describe the perceptual part of hazard cognition, which is commonly termed hazard detection.

The fourth and fifth demand component refer to the *assessment of the hazard* regarding to its urgency and intensity, and to the *self-assessment of abilities* in coping with the hazard. The greater an individual perceives their own abilities to cope with a given hazardous situation, the less the situation will be perceived as dangerous by that individual. The enclosure of both appraisal components underlines the obvious parallels that can be drawn to control theories of driving behaviour [13]. The same holds for the sixth component, the *weighing-up of risk:* This process covers the appraisal of the perceived risk in relation to one's own risk acceptance level. Thus, this component of hazard avoidance does not solely depend on the driver's self-perception with respect to skill level, but also on individual risk-taking attitudes. Moreover, from an activation perspective, it has to be taken into account that the preferred level of risk (as a correlative of preferred task difficulty) is generally rather at a medium than at a particularly low or high level (implying, respectively, monotony and boredom or distress and fear).

Decision-making, as the seventh demand component, is based on the aforementioned processes of risk assessment and, beyond, on drivers' behavioural repertoire and anticipated behavioural consequences. Especially during hazardous situations, decisions are made under critical conditions (e.g. increased arousal, temporal urgency, adverse environmental circumstances). Thus, reliable decision-making, not requiring high cognitive effort or deliberating—i.e. the activation and retrieval of automated, routinized behavioural scripts—becomes essential for successfully coping with imminent hazards. The eighth demand component, *action,* covers the actual execution of a driving behaviour selected from the available options. This component of hazard avoidance thus refers to the level of vehicle operation and mainly includes psychomotor aspects. Failures in this component are commonly described as slips and lapses [14].

As pointed out before, this framework of structural components is a result of a comparative analysis of hazard cognition models and, as such, is expected to cover the main aspects of traffic perception and hazard avoidance. However, it has to be noted that the linear way in which components are depicted serves systematization, but the notion of processes running concurrently and influencing each other both bottom-up and top-down must not be neglected when thinking about hazard cognition in traffic.

3 Development of an Item Set for Testing Traffic Perception and Hazard Avoidance

3.1 Selection of Task Types

In the context of the multi-faceted nature of hazard avoidance, a vast body of research has compared novice and expert drivers with respect to the skill components mentioned above. In many studies, differences were found between the two groups concerning their visual scanning patterns and attentional focus [15–17], their hazard detection performance and anticipatory performance in hazard prediction [18–20], and their self-enhancement bias with respect to perceived driving capabilities [21]. Some studies also revealed somewhat mixed results and partly failed to identify substantial differences between novice and experienced drivers in hazard perception tasks [22, 23]. In order to explain inconsistent results, some authors [24] argued that different types of traffic hazards must be distinguished according to the relationship between a hazard and its precursors. Performance differences between novice drivers and experienced drivers are assumed to be elicited in particular by hazards that are indirectly linked with preceding precursors. Furthermore, in most of the tasks, subjects' responses towards hazards reflect the confounded impact of multiple processes [18] and some of these processes might be more susceptible to experiential influences than others [25].

In this context, further research has examined several task formats to identify appropriate formats for isolated measurement of distinguishable skill components. For instance, categorization tasks and rating tasks (rating of situations' hazardousness) were proposed to particularly display processes of risk appraisal [26], whereas real-time hazard identification tasks have been linked to the perceptual components of hazard cognition. Recently, derived from the Situational Awareness Global Assessment Technique (SAGAT, [27, 28]), a task format has been proposed that explicitly targets the prediction component within hazard cognition [18]. Experimental results provided support for this approach with respect to the discrimination between experts and novices and also revealed an effect of hazard type on discriminative power.

In order to find appropriate task formats for our purpose, it seemed worthwhile at a first stage to look for formats proposed in the research literature or already applied in practice. Based on an extensive literature review [29], 12 task formats were identified, described and evaluated according to criteria of existing empirical support, relevance concerning the identified demand components, suitability for driver training and testing, and economy regarding test implementation and interpretation. As a result of this assessment process, three different task formats were selected which promised to be most suitable for our objectives.

The first format—*Task Type 1*—refers to the localization and identification of objects in the traffic environment. A variant of this format has been described by [30]. The task requires adequate scanning of the traffic scene and the ability to discriminate objects in the environment that are relevant to the driving task from those that are not. Thus, this format does not explicitly focus on the detection of hazards, but is conceived to assess general traffic observation skills. Within the hazard cognition process, it relates to the sub task of monitoring, for which differences between novice and expert drivers have been identified in eye-tracking studies (e.g., novice drivers produce longer fixations on objects; their visual search is restricted to a smaller area that is closer to the front of the vehicle [31]). In a task of type 1, participants are presented with a simulated dynamic traffic scenario, which they view from a car driver's ego-perspective. They are requested to mark all objects they have to pay attention to as a driver (in real time, by finger-touch on the screen). From the number of objects correctly identified as relevant in relation to the total number of markings, a score can be derived as a measure of scanning skills.

The second format—*Task Type 2*—refers to the anticipation of events in the ongoing situation and the identification of potential hazards. This format is closely related to the aforementioned task format derived from the SAGAT and has been used in a similar version by [18], among others. In a task of type 2, the dynamic traffic scenario contains the development of a potential hazard, defined as an event that requires the driver to react (by decelerating, braking or steering, e.g.). At a predefined moment, in which the potential hazard source is already perceivable, the traffic scenario freezes. Participants then have five seconds to mark at most two objects as hazard sources or to indicate "no hazard present" by touching the respective button (they are informed that scenarios can contain either no, one or two hazard source/s). Performance is measured using the number of correct identifications.

The third format—*Task Type 3*—also contains a developing hazard, but a decisional component is added on the demand side: Participants are requested to touch a stop button whenever a dynamically presented traffic scenario becomes potentially hazardous and a regulative action (steering, decelerating, braking) is required from the driver. Then, the scenario freezes, and participants are requested to indicate the source of the hazard by finger-touch on the screen. Performance measures are derived from the latency of stops in relation to the moment the hazard first becomes perceivable and correct indications of the hazard source.

3.2 Selection of Task Contents

Besides the identification of suitable task formats for assessing traffic perception and hazard avoidance, the question had to be answered what demands should be represented in the traffic scenarios with regard to contents. To obtain items of high validity, ideally, the situational demands imposed in the scenarios should be relevant to novice driver safety and at the same time call for the competencies described above. Furthermore, task contents should reflect demands drivers have to cope with in real traffic.

A comprehensive description of driving task demands has been worked out by [32–35]. These demands were identified using a task analytical approach and have also been evaluated with respect to their relevance to novice driver safety [32–35]. Based on this work, a catalogue of driving tasks has been developed by [36]. This catalogue contains, among others, eight categories of driving tasks with high relevance for road safety, i.e. eight prototypical classes of traffic situations (e.g. changing lanes, maneuvering in curves, crossing intersections). Moreover, it specifies five overarching fields of competence (e.g. traffic observation, vehicle positioning, speed adjustment) and provides detailed descriptions of behavioral demands and examples of poor as well as above-average performance in each field of competence for each driving task.

We selected task contents so as to cover the eight categories of driving tasks and used the behavioral descriptions with respect to the five fields of competence, particularly *traffic observation*, to elaborate relevant details of the specific traffic scenarios to be presented. Additionally, we considered the aforementioned findings on the predictability of hazards based on certain precursors and, if applicable, findings on situational risk factors associated with severe crashes in novice drivers.

3.3 Implementation and Pretest

Based on the traffic scenarios developed, five items of task type 1, ten items of task type 2 and ten items of task type 3 were implemented for further testing. Stimulus materials were created using the VICOM-Editor, a software of TÜV | DEKRA arge tp 21. This software enables the construction of complex virtual traffic situations in great perceptual accordance to real-world sceneries and provides the utility to systematically vary specific factors (e.g. weather, daytime, sight conditions) while other conditions can be held constant. 25 virtual movie sequences were produced, representing the eight driving tasks from a driver's point of view (for an example see Fig. 1).

Fig. 1 Virtual traffic scenario created in VICOM-Editor

Before being employed in the training study, the items constructed were subjected to a pretest with 53 participants. One objective of this pretest was to obtain information on item difficulty in order to detect and avoid ceiling or floor effects. Another objective was to examine whether experienced drivers would perform better in the tasks than inexperienced drivers. Therefore, participants with varying levels of driving experience were recurited: 28 were learner drivers (beginners without a valid driving license), while 25 held a valid driver license and reported a total mileage assigned to one of three categories (<30.000 km, 30.000–50.000 km, and >50.000 km). The items were presented on a 10.1-in touchscreen; participants' responses were recorded via the touchscreen functionality.

Results of the pretest revealed that, for the majority of the items, item difficulty was in the range of 0.2–0.8, suggesting a suitable degree of difficulty. There was a tendency towards better performance in the group of experienced drivers, which suggests some degree of criterion validity, taking into account that the proportion of experienced drivers with very high mileage ("expert" drivers as often recruited in group comparisons to demonstrate criterion validity) was rather low.

Items with insufficient values in either difficulty or discrimination were subjected to a content-related reanalysis and expert discussion in order to identify features that could be modified to improve the item and were then revised accordingly. The revised item set is currently being trialed in a larger-scale pilot study, described hereafter.

4 Trial Study

4.1 Study Design

For the purpose of further testing, the item set will be employed, among other measures of driving competence, to assess traffic perception and hazard avoidance skills in a training study with learner drivers. Participants of the control group take part in a conventional driver training, consisting of 14 theory lessons (90 min each) and about 25–35 practical driving lessons (45 min each) that does not entail special training of hazard perception skills. For participants of the experimental group, driver training extends over the same time; however, it comprises innovative training methods designed to improve competencies of traffic perception and hazard avoidance. The materials include a 90-min unit for theory classes designed to impart the basics of traffic perception (e.g. developing effective strategies for traffic observation) and another 90-min training unit dealing with specific hazardous roads in the learner drivers' local environment (roads with high rates of young novice driver accidents). Using road videos and accident descriptions, the learner drivers virtually travel on the hazardous roads and analyze potential hazards. Afterwards, in the practical driving lessons, the roads are purposefully used to apply and stabilize the traffic observation and hazard avoidance strategies acquired.

The study is conducted in nine driving schools in various federal states of Germany. Assessment starts with the control group in all participating driving schools. When this first phase is completed, the involved driving instructors take part in a further training, in which they get to know the innovative training units and implementation methods. After that, in a second phase, the experimental group will run through the driver training and assessment process. Sample sizes of c. 150 participants are targeted for both the control and the experimental condition.

In order to monitor the development of traffic perception and hazard avoidance skills, all participants are to be tested with the item set described above at three times of measurement: (1) prior to the beginning of driver training, (2) upon completion of theory classes, and (3) upon completion of practice classes. Furthermore, knowledge tests (on road traffic hazards and characteristics of novice driver skill deficits) will be conducted at all times of measurement. Additionally, at t_3, driving performance in real traffic will be assessed in order to test for transfer of potential learning effects into real traffic environments. Driving performance in real traffic will be measured by means of a systematic observation, based on an electronic test protocol that is to be prospectively used in the practical driving test in Germany.

No differences between the experimental and the control group are expected at t_1 with regard to traffic perception and hazard avoidance skills measured with the item set. In the course of theory lessons, it can be assumed that the participants of both groups experience gains in their knowledge and skills concerning traffic perception and hazard avoidance. However, the growth in competence at t_2 should be higher in the experimental group than in the control group. The advantage should also be obvious at t_3 and be additionally reflected in the outcomes of the observation drive in real traffic.

4.2 Status of Assessment

Data collection started in November 2015 and is expected to continue until October 2016. Up to now, data of 77 control group participants, aged 16–39 years (*M* = 18.7, *SD* = 5.1), have been obtained for measurement time t_1. Of these participants, 23 have also contributed data for measurement time t_2 already (10 men and 13 women; *M* = 19.3, *SD* = 5.8 for age).

Intermediate results concerning the traffic perception item set provide initial evidence that a medium overall level of task difficulty could be established for the target group in all three of the task types: Overall, 48 % of predefined targets for driver attention were correctly identified in task type 1, 53 % of hazards were found in task type 2, and 41 % of hazards were correctly indicated in task type 3. Preliminary comparisons of test performance at t_1 and t_2, i.e. prior to the beginning of driver training versus upon completion of theory classes, show no signs of increase in the overall score of targets detected for task types 1 and 3. There appears to be a small increase in test performance with regard to task type 2. However, we refrain from reporting further statistics here, as these intermediate results only provide initial and preliminary information and must be reconsidered in due course as the collection of the entire data set is completed.

5 Outlook

The study outlined here is expected to yield valuable hints on the feasibility of the three task types and the item set applied for assessing traffic perception and hazard avoidance. Thus, it is to advance the overall state of knowledge on the measurement of hazard perception abilities on the one hand; on the other hand, it will be useful in the development of a respective test for the German system of driver preparation. It will also be informative with regard to possible improvements in driver training in Germany, especially concerning the instruction and training of knowledge and skills in traffic observation, hazard awareness and hazard avoidance.

Data collection is to continue until results will be available for about 150 learner drivers in both the control group and the experimental group. Test performance (knowledge tests, traffic perception and hazard avoidance test data) will be analyzed by group and time of measurement. In addition, correlations of test performance with observed driving performance in real traffic at measurement time t_3 are to be examined. The results of the training study will provide information as to the practicability and validity of the item set and task types applied.

If the tasks and items prove valid in being sensitive to differences in driver training with regard to traffic perception and hazard avoidance, still, a few more questions can be asked concerning validity. In order to further investigate these issues, we plan to extend the study to the probation phase that follows the completion of formal driver training and to collect data on critical incidents encountered by our participants in their driving career during one year after passing the practical driving test. Despite the challenges of predictive validity studies in this area, it will be interesting to see if a relation can be found between performance in the traffic perception test and the frequency and kind of accidents and other critical situations in traffic that participants are involved in.

Finally, towards the development of a hazard perception test for the German system of novice driver preparation, a number of other questions need to be resolved, including the definition of a criterion for passing and the determination of the optimal time at which to embed the test in the overall licencing process. The study introduced here represents one important step in this development process and will also make its contribution to the international state of knowledge on hazard perception testing.

References

1. Mayhew, D.R., Simpson, H.M., Pak, A.: Changes in collision rates among novice drivers during the first months of driving. Accid. Anal. Prev. **35**, 683–691 (2003)
2. OECD: Young Drivers, the Road to Safety. Transportation Research Centre, OECD Publishing, Paris (2006)
3. Curry, A.E., Pfeiffer, M.R., Durbin, D.R., Elliot, M.R.: Young driver crash rates by licencing age, driving experience, and license phase. Accid. Anal. Prev. **80**, 243–250 (2015)
4. Maycock, G., Lockwood, C., Lester, J.: The Accident Liability of Car Drivers. Transport and Road Research Laboratory, Crowthorne (1991)
5. Forsyth, E., Maycock, G., Sexton, B.: Cohort Study of Learner and Novice Drivers: Part 3, Accidents, Offences and Driving Experiences in the First Three Years of Driving. Transportation Research Laboratory, Crowthorne (1995)
6. Elvik, R., Hoye, A., Vaa, T., Sorensen, M.: The Handbook of Road Safety Measures. Emerald Group Publishing Ltd., Bingley (2009)
7. Wells, P., Tong, S., Sexton, B., Grayson, G., Jones, E.: Cohort II: A Study of Learner and New Drivers, vol. 1. Main Report. Department for Transport, London (2008)
8. Helman, S., Grayson, G.B., Parkes, A.M.: How Can We Produce Safer New Drivers?. Transportation Research Laboratory, Crowthrone (2010)
9. Wetton, M.A., Hill, A., Horswill, M.S.: The development and validation of a hazard perception test for use in driver licencing. Accid. Anal. Prev. **43**(5), 1759–1770 (2011)
10. Sexton, B.F.: Implementation of Hazard Perception Testing for Learner Drivers. Transportation Research Laboratory Ltd., Crowthorne (2002)
11. Genschow, J., Sturzbecher, D.: Traffic perception tests as innovative examination concept during novice driver preparation. In: Traffic Perception and Hazard Avoidance. Basic Concepts and Scope of Implementation in Novice Driver Preparation, pp. 13–57. TÜV| DEKRA arge tp21, Dresden (2015)
12. Velichkovsky, B., Rothert, A., Kopf, M., Dornhoefer, S., Joos, M.: Towards an exprees-diagnostics for level of processing and hazard perception. Transp. Res. Part F Traffic Psychol. Behav. **5**(2), 145–156 (2002)
13. Fuller, R.: Driver control theory, from task difficulty homeostasis to risk allostasis. In: Handbook of Traffic Psychology, pp. 13–26. Elsevier, Amsterdam (2011)
14. Reason, J.: Human Error. Cambridge University Press, Cambridge (1990)

15. Underwood, G., Chapman, P., Berger, Z., Crundall, D.: Driving experience, attentional focus, and the recall of recently inspected events. Trasnp. Res. Part F Traffic Trans. Psychol. **6**(4), 289–304 (2003)
16. Young, K., Lenne, M.G., Beanland, V., Salmon, P., Stanton, N.A.: Where do novice and experienced drivers direct their attention on approach to urban rail level crossings? Accid. Anal. Prev. **77**, 1–11 (2015)
17. Konstantopoulos, P., Chapman, P., Crundall, D.: Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. Accid. Anal. Prev. **42**, 827–834 (2010)
18. Crundall, D.: Hazard prediction discriminates between novice and experienced drivers. Accid. Anal. Prev. **86**, 47–58 (2016)
19. Kinnear, N., Kelly, S.W., Stradling, S., Thomson, J.: Understanding how drivers learn to anticipate risk on the road: a laboratory experiment of affective anticipation of road hazards. Accid. Anal. Prev. **50**, 1025–1033 (2013)
20. Lim, P.C., Sheppard, E., Crundall, D.: A predictive hazard perception paradigm differentiates driving experience cross-culturally. Trasnp. Res. Part F **26**, 210–217 (2014)
21. De Craen, S., Twisk, D.A.M., Hagenzieker, M.P., Elffers, H., Brookhuis, K.A.: Do young drivers overestimate their skills more than experienced drivers? Different methods lead to different conclusions. Accid. Anal. Prev. **43**(5), 1660–1665 (2011)
22. Sagberg, F., Bjornskau, T.: Hazard perception and driving experience among novice drivers. Accid. Anal. Prev. **38**, 407–414 (2006)
23. Underwood, G., Ngai, A., Underwood, J.: Driving experience and situation awareness in hazard detection. Saf. Sci. **56**, 29–35 (2013)
24. Crundall, D., Chapman, P., Trawley, S., Collins, L., Van Loon, E.: 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)
25. Wallis, T.S.A., Horswill, M.S.: Using fuzzy signal detection theory to determine why experienced and trained drivers respond faster than novices in a hazard perception test. Accid. Anal. Prev. **39**, 1177–1185 (2007)
26. Borowsky, A., Oron-Gilad, T.: Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. Accid. Anal. Prev. **59**, 548–565 (2013)
27. Endsley, M.: Situation Awareness Global Assesment Technique (SAGAT): Air-to Air Tactical Version User Guide. Northrop Corp, Hawthorne (1990)
28. Endsley, M.: Toward a theory of situation awarness in dynamic systems. Hum. Fact. **37**(1), 32–64 (1995)
29. TÜV|DEKRA arge tp21: Traffic perception and hazard avoidance. In: Foundations and Possibilities in Novice Driver Preparation. TÜV|DEKRA arge tp 21, Dresden (2015)
30. Hampel, B.: Evaluation of an auditiv-visual test system of the TÜV Rheinland. In: Human–Vehicle—Environment, pp. 57–91. TÜV Rheinland, Cologne (1977)
31. Chapman, P., Underwood, G.: Visual search of dynamic scenes: event types and the role of experience in viewing driving situations. Eye Guid. Read. Scene. Percept. 369–393 (1998)
32. McKnight, A., Adams, B.: Driver education analysis. Volume II: task analysis methods. In: Department of Transportation. National Highway Traffic Safety Administration, Washington D.C. (1970)
33. McKnight, A., Hundt, A.: Education Task Analysis, Vol. 4: The Development of Instructional Objectives. Human Resources Research Organization, Alexandria, VA (1971)
34. McKnight, A., Adams, B.: Driver Education Task Analysis, Vol. 1: Task Description. Department of Transportation, National Highway Administration, Washington D.C. (1970)
35. McKnight, A., Hundt, A.: Driver Education Analysis, Vol. 3: Instructional Objectives. Human Resources Research Organization, Alexandria, VA (1971)
36. Sturzbecher, D., Moerl, S., Kaltenbaek, J.: Optimisation of the Practical Driving Test, Report of the Federal Highway Research Institute M243b. Fachverlag NW, Bremen (2014)

Driver Distraction and Advanced Vehicle Assistive Systems (ADAS): Investigating Effects on Driver Behavior

David M. Cades, Caroline Crump, Benjamin D. Lester and Douglas Young

Abstract The component technologies of Advanced Driver Assistive Systems (ADAS) are becoming increasingly automated, with systems capable of operating in concert in multiple driving environments. However, how these systems affect a driver's ability to safely, efficiently, and comfortably operate a vehicle remains unclear. We investigated the effects of ADAS [specifically Lane Departure Warning (LDW)] on driving performance while participants performed a secondary task (mental math) designed to simulate cognitive effort while driving. The experiment was conducted on a closed-course test track in an instrumented vehicle. Results suggest that cognitive engagement influenced driver control of the vehicle. Effects of cognitive engagement in a secondary task were not mitigated by the presence of LDW. We discuss our results in the framework of a continued need for active input and control from the human operator in vehicles with assistive technologies.

Keywords Human factors · ADAS · LDW · Distraction

D.M. Cades (✉)
Exponent, Human Factors, Chicago, IL, USA
e-mail: dcades@exponent.com

C. Crump · D. Young
Exponent, Human Factors, Los Angeles, CA, USA
e-mail: ccrump@exponent.com

D. Young
e-mail: dyoung@exponent.com

B.D. Lester
Exponent, Human Factors, Phoenix, AZ, USA
e-mail: blester@exponent.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_83

1 Introduction

With the advent of advanced driver assistive systems (ADAS), such as lane departure warning, blind spot indication, forward collision warning and mitigation, and autonomous braking, drivers will be exposed to novel situations in which they must simultaneously prioritize and respond to vehicle assistive technologies and on-road conditions. ADAS technologies were designed to reduce driver error and enhance the overall safety of ground transportation [1]. Currently available statistics from the Insurance Institute of Highway Safety (IIHS) indicate certain ADAS systems appear to reduce property damage and liability claims [2]. However, understanding the impact of vehicle assistance technologies on driver behavior and vehicle control has lagged behind the evolution of the technology. Much of the current understanding on the behavioral impacts of ADAS has been gained from naïve drivers' initial exposure to systems in driving simulators. These studies show that the presence of assistive technologies has been associated with behavioral changes in drivers including increased complacency, reduced human supervision of the technologies, and increased likelihood of engaging in secondary tasks [1, 3–6]. As a result, there is growing concern that drivers are unable to take control of the vehicle in emergent, safety-critical situations which automated systems are typically unable to handle.

More specifically, scientific studies have investigated the effects of various ADAS technologies on driver performance. For example, several studies have focused Lane Departure Warning (LDW) technologies and their effect on driver behavior. LDW typically consists of a passive audio and/or visual warning that the driver is departing the lane of travel, and requires that the driver maintain both speed and lane position. Studies of driver behavior when utilizing LDW present a generally optimistic view of driver interaction, including that drivers reduce the number and frequency of lane departures [7, 8], increase the use of turn signals [7], and respond more quickly to events even when drowsy [9]. Furthermore, one study showed that intermittent removal of the LDW did not lead to increases in near-misses or collisions [10], suggesting that drivers maintain awareness of lane position independently when LDW is present, though not consistently active.

The performance of secondary tasks while driving typically results in negative effects on performance of the primary driving task. For example, secondary task performance has been associated with increased deviations in lane position with some tasks [8, 11], decreased deviations in lane position with other tasks [12], and reduced or more variable speeds [13, 14]. While drivers have been shown to be significantly more likely to participate in secondary tasks when utilizing ADAS technologies [15–17], it remains to be seen whether ADAS technologies will mitigate the negative effects of such tasks. Furthermore, secondary tasks represent an interesting method by which to study not only the effects of the tasks themselves, such as texting while driving, but more general effects, such as visuomanual distraction while driving. This is important because tasks such as navigation can be considered cognitive or visual distractions from the primary task of controlling the vehicle. Distractions can be categorized as visual, manual, or cognitive [18]. These categories of distractions may differentially affect aspects of driving performance. Similarly, different ADAS technologies may support driver performance in the presence of specific types of distractions by secondary tasks.

Driver response to ADAS has not been thoroughly studied in dynamic, naturalistic conditions where drivers must interact with ADAS systems in the presence of common driver tasks or challenges (e.gs., vehicle following, curve negotiation, maintaining lane position, and performing secondary tasks). Additionally, the impact of assistive technologies on driver input and control of the vehicle (e.g., accelerator position, steering input, and speed) during on-road driving is lacking. Accordingly, of the current study evaluated driver-vehicle interactions under common on-road conditions in the safety of a closed-course test track. We specifically examined how drivers respond to the LDW system while simultaneously performing a task designed to simulate cognitive engagement, and looked at the differential impact on vehicle control metrics. Consistent with existing findings in simulators, the use of LDW was expected to reduce the variance observed in steering wheel position. Similarly, we hypothesized that participants performing a secondary task would exhibit changes in lateral vehicle control and speed. Finally, we predicted that if the LDW system was sufficiently salient and effective, changes in the variance of steering wheel position and potentially speed would be observed when drivers were performing the cognitively demanding secondary task.

The analyses focused on data from participants' initial exposure to LDW, as studies suggest increased experience may affect the way drivers interact with the systems [19, 20]. Overall, this study serves as a field test of LDW on driver behavior and vehicle control during common driving behavior including, lane maintenance, curve negotiation, car following, and secondary task performance.

2 Experiment

To investigate how LDW may affect driver behavior with and without cognitive engagement, we conducted an on-road, closed-course experiment at Exponent's Testing and Engineering Center in Phoenix, Arizona. The experiment was completed over the course of three visits; for this analysis, we examined driver performance metrics in the first visit only.

2.1 Methods

Participants. Twenty adults (15 female; median age 46.5 years; age range: 32–57) recruited from the Phoenix area participated in this study. Participants all had a valid driver's license, normal or corrected-to-normal vision and were screened to ensure that they did not currently drive a vehicle equipped with ADAS technologies.

Facility and Equipment. The experiment was conducted on a two-mile oval track and a 10-acre skid pad at Exponent's Test and Engineering Center. A single lane was painted on the track for testing. On each straightaway the lane shifted at the halfway point and transitioned either from wide to narrow or narrow to wide (see Fig. 1). The narrow sections of the lane markings were approximately seven feet in width and the wide sections were approximately 11 feet in width. To increase similarity to the natural driving environment, the test track environment included roadway signage and a simulated bus stop placed around the course of the track.

The test vehicle was a 2014 luxury sedan equipped with several driver assistance technologies, including an LDW system. The LDW system consisted of both auditory (series of beeps) and visual (digital display of a vehicle leaving a lane on the instrument cluster when sensors detected the vehicle was moving outside of the lane lines) warnings. The variable width and shifting of the lane was used to maximize the likelihood of participants receiving LDWs. LDW was activated or deactivated by the experimenter on the appropriate laps using the infotainment system in the center stack.

Data Collection Equipment. The test-vehicle was instrumented with four lipstick cameras and a laptop computer equipped with data recording software collecting information from the vehicle's Controller Area Network (CAN) bus. The CANbus recording software was utilized to record status information of the relevant state of various vehicle systems including, but not limited to, activation of ADAS, steering wheel angle, accelerator position, vehicle speed, engine RPMs and brake activation. Cameras were mounted over each side-view mirror, pointed at the ground to provide visual indications of lane position and instances of lane departure. Additional cameras were mounted inside the vehicle and captured the instrument cluster and an over-the-shoulder (driver's left shoulder) view of the roadway. A lead vehicle was driven by an experimenter for all testing laps.

Procedure. The full experiment was completed across three visits, each separated by approximately one week. On each visit, participants completed 12 laps around the 2-mile test track following the lead vehicle. ADAS conditions were varied throughout the laps such that equal numbers of laps were completed with no ADAS, with LDW only, with Adaptive Cruise Control (ACC), and with both LDW and ACC. In the present analysis, we examined only the effects of LDW alone, relative to the effects of no ADAS. Analyses of the effects of ACC or the effects of the combination of various ADAS technologies will be discussed in future publications. During each lap, participants followed a lead vehicle that controlled the pace and driving route. The lead vehicle varied its speed from between 40 and 50 mph during all laps.

Some participants were randomly assigned to perform a secondary task (i.e., cognitive or manual tasks) on every third testing lap. Participants who completed a mental math task, meant to induce cognitive distraction, were examined here here; data from three participants who completed other tasks were not considered for this

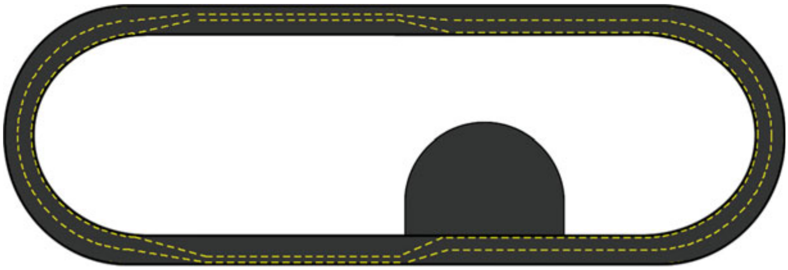


Fig. 1 General layout of test track and lane markings

analysis. The mental math task required participants to verbally answer multiplication problems administered by an experimenter seated in the passenger seat of the test vehicle. The presentation rate of the problems was controlled by how quickly each participant answered them. Mental math was performed only in the straightaways of the testing track.

Analysis. Data from the vehicle's CANbus were extracted for each participant. Data for each lap were manually separated into individual laps. Accuracy of the data separation was confirmed by comparing the data to the videos of each lap. The dependent variables of interest in the current analyses were steering wheel angle and vehicle speed. Vehicle speed was initially recorded in kilometers per hour (kph), and was converted to miles per hour (mph). Data from one subject was removed due to technical difficulties during data collection. Data for two additional participants were removed due to data collection in unusually dark lighting conditions. Data from the remaining 14 participants were submitted for this analysis.

The variance in steering wheel angle and in vehicle speed was quantified using the standard deviation (SD) for each metric, separately for each testing lap. Standard deviations were then analyzed using a 2 (ADAS Condition) × 2 (Secondary Task) mixed measures ANOVA. T-tests were conducted to further examine the differential effects of mental math with and without LDW.

3 Results

3.1 ADAS Condition

Presence of LDW did not significantly affect SD of steering wheel angle ($p = 0.75$), or SD of vehicle speed ($p = 0.37$).

3.2 Mental Math

Performing mental math marginally affected SD of steering wheel angle ($p = 0.076$), such that SDs were reduced during mental math (M = 8.019°) as compared to SDs when driving only (M = 8.48°). However, mental math also significantly affected SD of vehicle speed ($p = 0.045$), such that SD of vehicle speed was *increased* when performing mental math and driving (M = 15.83 mph) when compared to driving only (M = 12.14 mph).

3.3 Interaction Between LDW and Mental Math

Performing mental math did not interact with LDW to affect SDs of steering wheel angle ($p = 0.99$). Mental math also did not significantly interact with LDW to affect SDs of vehicle speed ($p = 0.57$), although examination of the data indicates that SDs when performing mental math while utilizing LDW (M = 17.36 mph) were numerically higher than SDs when driving only and utilizing LDW (M = 12.63 mph; $p = 0.10$); see Fig. 2.

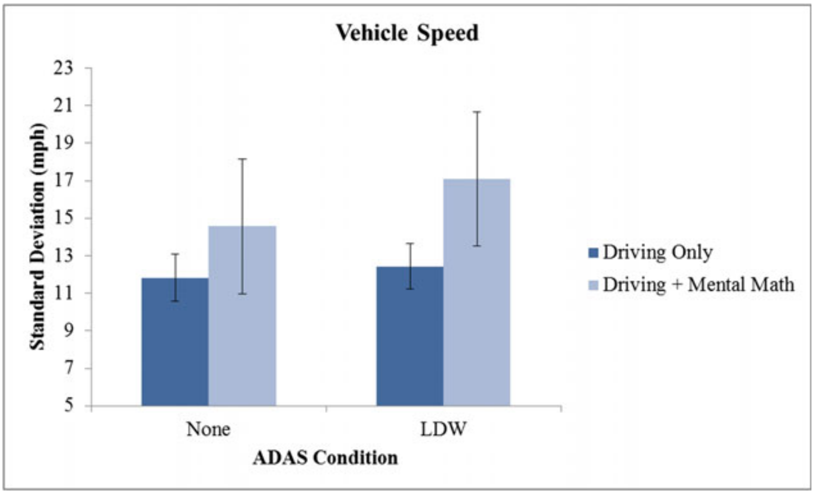


Fig. 2 Standard deviation in vehicle speed

4 Discussion

The present results indicate that LDW did not affect variation in the vehicle control metrics we examined. Consistent with existing scientific literature, we found that performance of a secondary tasks did affect vehicle control, specifically variations in steering wheel angle and speed control. We anticipated that LDW may reduce the effect of the secondary task. This hypothesis was supported by the null effect of mental math on variation in steering wheel angle in the presence of LDW, despite a marginally significant effect of mental math on variation in steering wheel angle overall. However, as shown in Fig. 2, the difference between SDs in speed when performing mental math and when driving only were actually greater in the presence of LDW than without any ADAS. Taken together, our results suggest that passive ADAS technologies such as LDW may not fully compensate for the effects of cognitive distractions on vehicle control. Our results further suggest discrete effects of LDW as a reminder to maintain lateral control, rather than a more general reminder to attend to the vehicle's position.

The current analyses focus on the first day drivers were exposed to LDW. Existing literature suggests that effects of ADAS may change over time and with increasing experience with the technologies. Future analyses will examine how these patterns might change as drivers gain experience and familiarity with ACC and LDW. It should also be noted that we examined relatively few laps with mental math when compared to laps without a secondary task. Thus, while our results indicate several marginal effects or trends, it is likely that with increased observations, the results would show statistical significance. Regardless, the effects of cognitive distraction, such as that experienced when engaged in conversation while driving, on both lateral and longitudinal control when following a lead vehicle are highly relevant in terms of everyday situations.

5 Conclusion

These preliminary results suggest that cognitive engagement affects vehicle control in on-road situations and is not fully compensated by the presence of ADAS. While it may be hoped that ADAS technologies will help reduce driver workload, our results support a growing body of evidence that such technologies cannot replace an alert, attentive driver when considering safe operation and control of a motor vehicle in naturalistic on-road conditions.

References

1. Brookhuis, K.A., de Waard, D., Janssen, W.H.: Behavioral impacts of advanced driver assistance systems—an overview. Eur. J. Transp. Infrastruct. Res. **1**(3), 245–253 (2001)
2. Lund, A.: Drivers and driver assistance systems: how well do they match? In: Driving Assessment Conference—Lake George, NY (2013)
3. 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)
4. Merat, N., Jamson, A.H., Lai, F.C., Carsten, O.: Highly automated driving, secondary task performance, and driver state. Hum. Fact. J. Hum. Fact. Ergon. Soc. **54**(5), 562–771 (2012)
5. Carsten, O., Lai, F., Barnard, Y., Jamson, A.H., Merat, N.: Control task substitution in semiautomated driving: does it matter what aspects are automated? Hum. Fact. J. Hum. Fact. Ergon. Soc. **54**(5), 747–761 (2012)
6. Muhrer, E., Reinprecht, K., Vollrath, M.: Driving with a partially autonomous forward collision warning system: how do drivers react? Hum. Fact. J. Hum. Fact. Ergon. Soc. **54**(5), 698–708 (2012)
7. LeBlanc, D., Sayer, J., Winkler, C., Ervin, R., Bogard, S., et al.: Road departure crash warning system field operational test methodology and results (Report No. UMTRI-2006-9-1) (2006)
8. Blaschke, C., Breyer, F., Farber, B., Freyer, J., Limbacher, R.: Driver distraction based lane-keeping assistance. Transp. Res. Part F **12**, 288–299 (2009)
9. Kozak, K., Pohl, J., Birk, W., Greenberg, J., Artz, B., et al.: Evaluation of lane departure warnings for drowsy drivers. In: Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, pp. 2400–2404 (2006)
10. Breyer, F., Blaschke, C., Farber, B., Freyer, J., Limbacher, R.: Negative behavioral adaptation to lane-keeping assistance systems. IEEE Intell. Transp. Syst. 21–32 (2010, Summer)
11. Peng, Y., Boyle, L.N., Hallmark, S.L.: Driver's lane keeping ability with eyes off road: insights from a naturalistic study. Accid. Anal. Prev. **50**, 628–634 (2013)
12. He, J., McCarley, J.S., Kramer, A.F.: Lane keeping under cognitive load: performance changes and mechanisms. Hum. Fact. **56**(2), 414–426 (2013)
13. Horberry, T., Anderson, J., Regan, M.A., Triggs, T.J., Brown, J.: Driver distraction: the effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. Accid. Anal. Prev. **38**, 185–191 (2006)
14. Rakauskas, M.E., Gugerty, L.J., Ward, N.J.: Effects of naturalistic cell phone conversations on driving performance. J. Saf. Res. **35**, 453–464 (2004)
15. Jamson, A.H., Merat, N., Carsten, O.M., Lai, F.C.: Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. Transp. Res. Part C Emerg. Technol. **30**, 116–125 (2013)
16. Carsten, O., Lai, F., Barnard, Y., Jamson, A.H., Merat, N.: Control task substitution in semiautomated driving does it matter what aspects are automated? Hum. Fact. **54**, 747–761 (2012)
17. De Winter, J.C., Happee, R., Martens, M.H., Stanton, N.A.: Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. Transp. Res. Part F Traffic Psychol. Behav. **27**, 196–217 (2014)
18. Cades, D.M., Arndt, S.R., Kwasniak, A.M.: Driver distraction is more than just taking eyes off the road. Inst. Transp. Eng. J. 81.7 (2011)
19. Xiong, H., Boyle, L.N., Moeckli, J., Dow, B.R., Brown, T.L.: Use patterns among early adopters of adaptive cruise control. Hum. Fact. **54**(5), 722–733 (2012)
20. Kidd, D.G., Nelson, E.T., Baldwin, C.L.: The effects of repeated exposures to collision warnings on drivers' willingness to engage in a distracting secondary task. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 54, pp. 2086–2090 (2010)

Part XIII
Road and Rail—Infrastructure

Generating a Lane-Specific Transportation Network Based on Floating-Car Data

Robert Neuhold, Michael Haberl, Martin Fellendorf, Gernot Pucher, Mario Dolancic, Martin Rudigier and Jörg Pfister

Abstract Future applications in ITS and automated driving require high precise digital maps including a lane-specific transportation network. The paper presents a method for estimating lane center lines based on vehicle trajectories from floating-car data. Kernel density estimation was applied for estimating lane center lines. The floating-car dataset is based on measurements on three different road types (urban 3-lane freeway, urban arterial, rural 2-lane freeway) using different low-cost GNSS receivers (GPS data logger and several smartphone GPS positioning apps). As reference, some test runs were conducted with high precise D-GPS measurement equipment. The longitudinal and lateral positioning errors were analyzed within a roadway and trip based distance analysis. The final results show deviations less than 0.14 m in median between measured and estimated lane center lines. This accurate estimation of lane center lines allows a generation of lane-specific transportation networks based on common floating-car data.

R. Neuhold (✉) · M. Haberl · M. Fellendorf
Institute of Highway Engineering and Transport Planning,
Graz University of Technology, Rechbauerstr. 12, 8010 Graz, Austria
e-mail: Robert.Neuhold@tugraz.at

M. Haberl
e-mail: Michael.Haberl@tugraz.at

M. Fellendorf
e-mail: Martin.Fellendorf@tugraz.at

G. Pucher · M. Dolancic
TraffiCon—Traffic Consultants GmbH, Strubergasse 26, 5020 Salzburg, Austria
e-mail: Pucher@trafficon.eu

M. Dolancic
e-mail: Dolancic@trafficon.eu

M. Rudigier
Virtual Vehicle Research Center, Inffeldgasse 21a, 8010 Graz, Austria
e-mail: Martin.Rudigier@v2c2.at

J. Pfister
pwp-systems GmbH, Prießnitzstr. 11, 65520 Bad Camberg, Germany
e-mail: Pfister@pwp-systems.de

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_84

Keywords Lane-specific transportation network · Floating-car data · GPS measurement devices · Distance analysis · Kernel density estimation

1 Introduction

Transport planners usually model roads as one single edge between two nodes (e.g. intersections) in transportation networks, irrespective of the number of lanes. Therefore, single lines as part of an entire road graph represent the road sections. Often lane-specific information like the number of lanes is included in additional attributes of the graph. This generalization of road geometries reduces the resolution of the data as well as costs in the development and maintenance of a transportation network [1].

However, existing and emerging ITS services might require digital road network graphs with a higher level of detail and accuracy regarding the representation of lane center lines. Cooperative services, for instance, often either need the lane specific localization of messages or provide information for specific lanes [2]. Examples would be lane departure warnings, local hazard messages (e.g. road bumps, accidents, congestion) or lane specific route information (speed limit, turn relations, curvature).

In the context of (highly) automated driving, transportation networks acts as a priori basic information, so that a vehicle can localize itself on the road using its own position relative to the road geometry. For this purpose, highly detailed maps, which include among others the lane center positions, the exact lane widths, associations between neighboring lanes and road hierarchy of single lanes, are required [3].

The development of such detailed maps, which contain a lane-specific transportation network, needs extensive measurement campaigns using highly accurate localization equipment or technologies. This is a costly and time-consuming process especially for wide areas or spacious transportation networks. On the other hand positioning data from moving observers (vehicles) so called floating-car data (FCD) is a GNSS based data source and is often available for wide areas.

Methodologies to derive geometries and topologies for digital street maps using GNSS-based FCD has been the focus of several research projects and studies in recent years. Davies et al. [4] focused on determining road center lines by assigning GNSS-positions to raster cells and creating histograms. Cells with high sums of allocated GNSS-points were assumed to represent the road center line. Sato et al. [5] also observed the frequency distribution of GNSS-points in raster cells, but focused on identifying the correct number of lanes. While they could reliably identify the correct number of lanes, they did not evaluate their exact center line positions.

This was the aim of a study by Knoop et al. [6], who introduced the Precise Point Positioning (PPP) technique in order to determine the lane a vehicle is travelling on and to create a self-learning street map in real-time. Uduwaragoda et al. [7] also focused on identifying the number of lanes and their center lines using GNSS data. They analyzed the probability density distribution of vehicle trajectories at road cross sections using a non-parametric Kernel Density Estimation. Results showed that lane center lines can be computed accurately enough if a minimum of 150 trajectories are available, independently of road type and characteristic.

Traffic management operators often use FCD for different applications like traffic monitoring and forecast [8]. Generally particular vehicle fleets (e.g. city taxi fleet) are equipped with GNSS positioning systems (e.g. GPS receiver) and provides FCD for traffic management centers in different forms, either as raw positioning data (vehicle trajectories) or as processed and map-matched data (e.g. link related travel times). The quality and accuracy of raw FCD in terms of positioning depends strongly on the measurement equipment. In general, low-cost GPS receivers are used which are installed either fixed in the vehicle itself or within other devices inside the vehicle (smartphone, route guidance system, GPS data logger).

Herrera et al. [9] analyzed traffic data obtained via GPS-enabled phones for purposes of traffic management applications and found out that FCD is suitable for average speed estimation on roads if 2–3 % of all vehicles are equipped with GPS-enabled phones. Zheng et al. [10] evaluated the accuracy of GPS-based taxi trajectory records in Guangzhou, China. Zheng et al. identified different types of erroneous data using a four filter criteria. Most outliers were detected by the low accurate signal criterion. Zheng et al. conclude that 65 % of records seem valid, so GPS often fail in positioning correct coordinates.

The development of a lane-specific transportation network based on vehicle trajectories from FCD is the key objective in the research project “LaneS”, funded by the Austrian Federal Ministry for Transport, Innovation and Technology. The idea is to estimate the center lines of each lane based on a wide set of lane-specific trajectories obtained from measurements with low-cost GPS devices. The quality of measured vehicle trajectories is evaluated already in advance by comparing them with trajectories from high accurate positioning measurements.

2 Methodology

2.1 General Approach

The general approach in this study within the research project “LaneS” is summarized in Fig. 1. The basis is a broad data collection of vehicle trajectories (VT) from test runs with different GNSS-based positioning technologies on various road sections. For positioning a high accurate differential GPS (D-GPS) measurement equipment is used as well as common GPS-devices like smartphones and data loggers.

The quality of VT was evaluated afterwards within a roadway based (lateral deviations of several VT) and a trip based (longitudinal and lateral deviations of

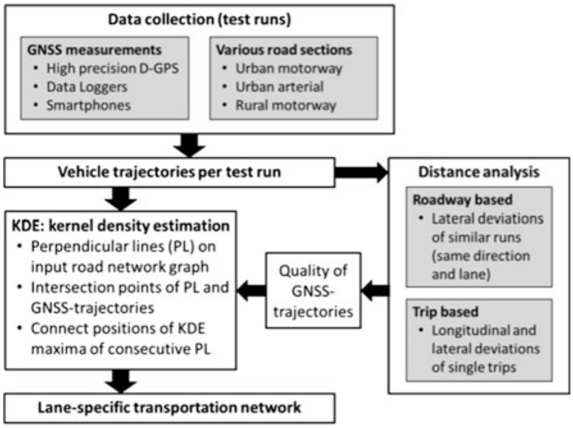


Fig. 1 General methodology to generate a lane-specific transportation network

single VT) distance analysis. Therefore, the VT were compared with the high accurate D-GPS measurements to identify outliers and erroneous trajectories.

The generation of a lane-specific graph was realized with kernel density estimation (KDE), which is a non-parametric probability density function. First perpendicular lines (PL) on an input graph (e.g. from Open-Street-Map) of the considered road section were created every 5 m. Then the VT were cut with all PL to establish intersection points. Applying KDE the position of lane center lines (maximum of probability density function) were estimated for each PL. Connecting every center point per lane over all PL achieves finally a lane-specific transportation network.

2.2 *Study Area and Measurement Systems*

The measurements of the floating-car data (FCD) took place at three different measurement sites (section A, B, C) near Graz (Austria) to cover various road categories. Section A is an urban 3-lane section on the freeway A2 near the city of Graz with a length of 14 km (8.7 mi). Section B is a 2-lane section on the urban arterial road Triesterstrasse in the city of Graz. A characteristic of urban sections is that shadowing effects caused by buildings may occur when measuring the vehicle position with a GPS receiver. Section C is a rural 2-lane section on the freeway S35 in the north of Graz with a length of 12 km (7.5 mi). Within the choice of these sections, we paid attention to avoid tunnels and bridges, because these sites can disturb the sensitive GPS receivers.

In total 369 test runs over more than 4000 km (2500 mi) were conducted on the three measurement sections. Within some of these trips, one vehicle was equipped with a differential GPS measurement system (D-GPS), which consists of an inertial measurement unit (IMU) combined with a GPS receiver. The correction data of a reference station are received with a GSM antenna. The achieved positioning accuracy is about 0.02 m at 100 Hz recording rate. During the measurements, we installed several low-cost GPS receivers in the vehicles. Therefore, several Qstarz GPS data logger with an update rate of 1 or 5 Hz and some smartphones with different GPS logging applications were used. Four apps for Android and one for iPhone was tested, all of them recording with an update rate of 1 Hz.

At all three measurement sites each lane was surveyed separately at several constant vehicle speeds. First, there were no lane changes within the test runs. The vehicle moved as close as possible to the center of the road lane. This is necessary especially for the generation of the reference trajectory based on D-GPS. After that, we also performed measurements with lane changes, because common and available FCD in real, used for generating map data, will contain irregular distributed lane changes and will not contain data only from one defined lane.

2.3 Distance Analysis of GNSS Based Vehicle Trajectories

The position accuracy of the measured GNSS based vehicle trajectories (VT) was evaluated within two different approaches of distance analysis. Therefore, we choose only test runs without lane change. In the roadway based distance analysis, only lateral deviations to a reference graph of similar test runs (trips) were analyzed to get results for spatial positioning errors. Therefore, VTs of same lane, direction and GPS device were considered separately. Afterwards the results were compared between different lanes and other GPS devices. Additionally in the trip based distance analysis, lateral and longitudinal deviations of VTs from the same trip (test run) but from different devices were analyzed to achieve also results for time-based positioning errors.

Roadway Based Distance Analysis. At first a reference graph was generated which models the center line of each lane in the study area. This was realized with the open source statistics program R-project. This reference graph is the result of smoothing several VTs from the high accurate D-GPS measurements per lane. The smoothing uses spline curve estimation in R-project. Then we calculated the Euclidean distances of each GNSS based VT, which are the nearest distances from each point of trajectory perpendicular to the reference graph. All distances of similar test runs (same lane, direction and device) were merged. For evaluating the quality of lateral positioning of the VTs we established two different graphical analysis: a boxplot to get the distribution of distances and a barplot where all distance measures were classified in different groups of positioning accuracy. To compare different measurement devices and road characteristics, average distance and deviation measures for each measurement device were calculated over all lanes and both directions per measurement section (section A, B and C).

Trip Based Distance Analysis. This type of analysis sets its focus on the total two-dimensional error of position fixes contained in typical VTs. Thus, the complete horizontal position error will be determined for each instance in time, for which a respective test receiver provides a valid position fix. In order to quantify the contained error of all accumulated fixes during the test runs, the "true" trajectory—which the vehicle was actually driving—has to be known with high precision. This "true" trajectory of the vehicle has been determined on the basis of the D-GPS measurement equipment. Due to the combination of dual frequency GNSS and an inertial navigation unit, the accuracy of these "true" reference trajectories (RT), are in the range of a few centimeters for all RT position fixes. The superior quality of the RTs are perfectly suited to determine the contained position errors in all valid fixes of the VTs, which are expected to be in the range of a few meters. While the roadway based distance analysis cannot distinguish between position errors of the test receiver and the deviation of the vehicle from the exact center line due to the driver, the current analysis is capturing the horizontal position error with high precision. In the course of error determination, the location of the GNSS antenna of the test receiver inside the vehicle has to be known accurately with respect to the reference point of the high performance equipment. In the current test setup, the respective lever arms have been determined a priori to the conducted test runs. These body offsets between test receiver and reference equipment inside the test vehicle are taken into account and the RTs are transformed to the exact location of the VTs before the residuals are drawn. While the position error along the driving trajectory does not harm the process of center line determination, only the deviation perpendicular to the driving direction contributes errors into the algorithm of this study. Thus, the distinction between longitudinal and lateral position error has been made. The determination of both parts of the position error, require the knowledge of the exact driving direction, which is also provided with the RT-coordinates. For the goal of this study, the lateral part of the determined position error is relevant and it was analyzed whether this part of the VTs is accurate enough to support the developed approach.

2.4 Lane-Specific Transportation Network Based on Kernel Density Estimation

The central assumption for estimating lane center lines from a set of standard GNSS vehicle trajectories (VT) is that the probability to determine a vehicle's position on a lane is highest along its central axis. Thus, the density of a population of vehicle trajectories is highest in the center of a lane and lowest around the edges. It follows that the density maxima of vehicle positions at a road cross section should correspond to the positions of lane center lines of a road. Moreover, the number of estimated density maxima indicates the number of lanes on a road.

For the computation of density distributions of GNSS based VTs, a Kernel Density Estimation (KDE) is applied. It is a non-parametric probability density function, which centers a smooth kernel function at each data point and sums them to estimate densities. Deng and Wickham [11] defines it as follows in Eq. (1),

$$\hat{f}\,kde(x) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \quad (1)$$

where K is the kernel function and h is the bandwidth. In this work, a Gaussian kernel function is applied. In order to find an appropriate bandwidth as smoothing factor, a data-driven "solve-the-equation" plug-in approach developed by Sheather et al. [12] is applied. To further deal with distinct data outliers, confidence intervals of 5 % from the median later vehicle position are introduced. Trajectories outside these confidence intervals are not considered in the computation of the KDE.

Systematically erroneous GNSS trajectories within the underlying input data can lead to wrong maxima estimations in the sense of not representing an actually existing lane. Thus, a geographic distance matrix is calculated which contains the distances taken pairwise between all elements within the found maxima set. If there are n maxima in the maxima set, the distance matrix is an n * n symmetric two-dimensional array with n * (n − 1)/2 distinct elements. The probability of distance relations within the maxima set is evaluated, so that potentially implausible lane center lines can be detected and omitted. In this way, the potential effects of accumulated erroneous GNSS trajectories and over smoothed bandwidths are minimized.

The developed algorithm is applied on equidistant road cross sections every 5 m along the observed road. For these road cross sections, perpendicular lines are drawn. The positions of intersections between GNSS based VTs and perpendicular lines are determined and assigned with IDs. As a result, the lateral positions of VTs at cross sections every 5 m along the observed road are obtained. Based on these positions, the KDE is computed. Then, the local maxima of the derived density distributions are estimated. For this, first and second derivative tests are conducted. The maxima of consecutive cross sections are connected with line strings using a shortest distance algorithm. In this way, the geometries and the basic topology of the lane-specific road network are constructed. The resulting lane numbers and geometries are then compared to the lane center geometries based on the highly precise D-GPS measurements.

3 Results and Discussion

3.1 Results of Roadway Based Distance Analysis

In the roadway based distance analysis, distribution and quantity of distances between measured GPS based vehicle trajectories and the reference graph (represents the centerline of each lane) were analyzed for each lane and direction of each section (A, B, C) in the study area. Exemplarily the results over all lanes and both directions for the section A (urban 3-lane freeway) are presented in Fig. 2

considering three different measurement devices (one data logger and two smartphone apps).

In the example of section A in Fig. 2 the lateral position accuracy is similar for Qstarz Data Logger and Android GPS Logger, although detection rate (5, 1 Hz) and number of trips (77, 26) is different. The characteristic of classes of positioning accuracy of both are similar (about 60 % of distances are less than 2 m to reference graph) as well as median (about 1.5 m) and mean (about 2 m) but standard deviation is higher for Android GPS Logger (3.21 m against 1.78 m). The quality is comparatively worse for iPhone GPS Logger (median 3.31 m); only 31 % of distances are less than 2 m to the reference graph.

Finally, average distance and distribution measures were calculated for all sections in the study area for different GPS devices (see Table 1). We achieved the best results with the Qstarz Logger 5 Hz (median 1.2–1.5 m). A detection rate of 1 Hz for the Qstarz Logger is not recommended here (median 2.5–3 m). Except the iPhone GPS Logger, the smartphone apps provide similar results to the Qstarz 5 Hz in terms of the median, but standard deviation is higher especially for the Android GPS Logger.

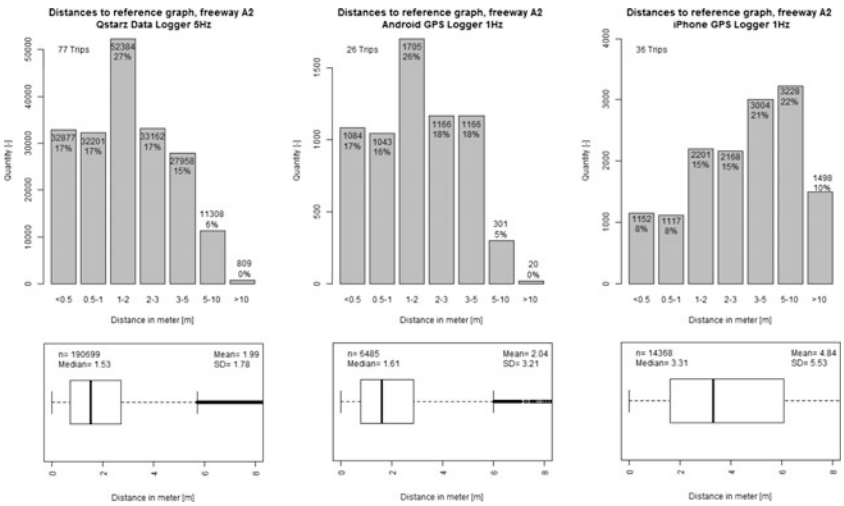


Fig. 2 Distances to reference graph on urban 3-lane freeway A2 (section A) for measured GPS vehicle trajectories based on Qstarz Data Logger (*left*), Android GPS Logger (*middle*) and iPhone GPS Logger (*right*). The barplot above shows the quantity of distances within classes of lateral positioning accuracy. The boxplot below shows distribution and statistics (quantity n, median, mean and standard deviation SD) of the distances

Table 1 Average distance and distribution measures for distances of different GPS devices for the three measurement sections A, B and C (SD means the standard deviation)

Average distance and distribution measures in meter (m)		Android GPS Logger	Android TopoNa-vigator	iPhone GPS Logger	Qstarz Logger 5 Hz	Qstarz Logger 1 Hz
Section A urban 3-lane freeway	1. Quartile	0.8	1.1	1.6	0.7	1.2
	Median	1.6	2.2	3.3	1.5	2.5
	Mean	2.0	2.4	4.8	2.0	3.0
	3. Quartile	2.9	3.5	6.1	2.7	4.3
	SD	3.2	1.7	5.5	1.8	2.5
Section B urban 2-lane arterial road	1. Quartile	0.8	0.6	–	0.6	–
	Median	1.7	1.2	–	1.3	–
	Mean	2.1	1.5	–	1.5	–
	3. Quartile	2.8	2.1	–	2.1	–
	SD	1.7	1.3	–	1.2	–
Section C rural 2-lane freeway	1. Quartile	0.8	0.7	0.8	0.6	1.4
	Median	1.8	1.5	1.8	1.2	3.0
	Mean	10.6	1.8	2.5	1.4	3.5
	3. Quartile	3.5	2.5	3.3	1.9	5.1
	SD	77.1	1.6	2.6	1.1	2.7

In section B iPhone GPS Logger and Qstarz Logger 1 Hz were not used, so there are no results here

3.2 Results of Trip Based Distance Analysis

With respect to the approach of this study, we expect that most vehicle trajectories would arise from the use of smartphones, as modern devices contain GNSS and data transmission to provide their tracks. In this regard, two different Android smartphones and a Qstarz data logger have been placed in the same vehicle equipped with the D-GPS measurement. The data logger is used as reference device, to check if the GNSS chip set inside the smartphone can achieve similar performance values of typical mass-market receiver. These tests have been conducted for different road categories, in order to capture the influence of environmental conditions on freeways and urban streets. All three devices have been analyzed individually for each of the three road sections, to detect whether vehicle speed or environmental conditions would have significant impact on the overall performance.

The test area and its surroundings show good GNSS reception conditions on the urban 3-lane freeway (section A), some influence from topography on the rural 2-lane freeway (section C) and minor urban challenges in the city (section B), since the buildings have mostly 4–6 floors and have some distance to the road. Exemplarily the results of trip based distance analysis on the urban 3-lane freeway (section A) for the Android GPS logger are shown in Fig. 3, which is a representative example of the performance that can be expected by using low-cost GPS receivers.

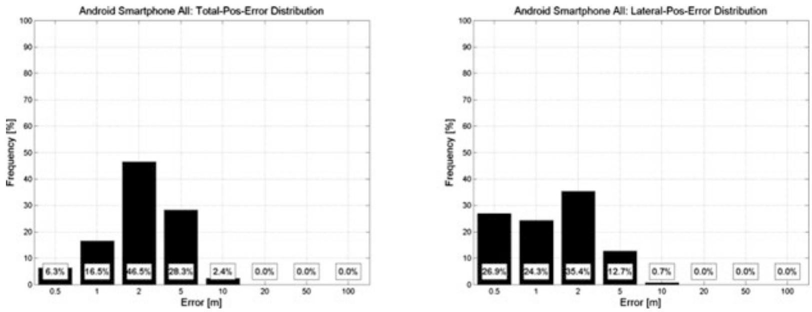


Fig. 3 Total (*left*) and lateral (*right*) position error distribution on the urban 3-lane freeway (section A) for the trajectories based on the Android GPS logger application for smartphones

The resulting position errors are presented as position error density, since this representation is most suitable for the current assessment objectives. In Fig. 3 such an error density is shown for the Android GPS logger over a sample of 25 test runs within one day. The resulting errors are distributed over eight error classes, from the half-meter class (the very left bin in both diagrams) to the hundred-meter class (the very right bin). The separation between two classes is the mean of both center values of each class. For example, all position errors greater than 1.5 m and smaller 3.5 m have been accumulated into the two-meter class, which represents the biggest bin in the left diagram of Fig. 3 with a share of 46.5 % of all determined position errors. The next bin with a high accumulation of error values is the five-meter class, which is a typical picture for mass-market receivers to have most hits in these two error classes.

With this understanding the left diagram in Fig. 3 shows that only 22.8 % of the errors are either contained in the half-meter class or the one-meter class, which would be sufficient, to be on the correct lane. Now looking at the right diagram in Fig. 3 the error density for the lateral part of the same Android receiver is depicted and it can be seen that the two left bins contain 51.2 % of the errors and thus smaller than 1.5 m. In other words, approximately half of all the valid fixes from smartphones are on the correct lane.

The results are not perfect, but they encourage the application of the kernel density estimation, since mass-market receiver would have the majority of all fixes on the correct lane. This approach also shows the limits of ordinary vehicle trajectories coming from mass-market devices, with respect to its applicability in other domains. The quality of the lateral position error is suitable for the purposes of this study, but it has to be noticed that the quality results cannot be assumed in the same way for other applications. The scientific analysis of mass-market receivers and the derivation of adequate parameters have been executed with respect to the specific requirements of the study.

3.3 *Results of Generating a Lane-Specific Transportation Network*

The algorithm described in Sect. 2.4 was applied for each of the three measurement sites in the study area (section A, B, C). In Fig. 4, exemplary cross sections of each section with the resulting KDEs are visualized. The punctuated lines show the x-coordinate positions of the detected maxima along the perpendicular lines of the road cross section. The light grey lines are equivalent to the positions of the highly precise D-GPS measurements. Furthermore, the respective derived lane geometries are depicted next to the diagrams. In these examples, the number of lanes was estimated correctly. The density distributions show distinctive maxima peaks corresponding to the estimated lane center lines. These estimated maxima are situated close to the lane center lines from the reference measurements. This indicates that the developed algorithm is capable for estimating the positions of lane center lines with high accuracy.

The overall performance of the estimation of lane center lines is evaluated based on the reference measurements. The following boxplot (Fig. 5) shows the distribution of distances of the estimated lane center line positions to the reference lane center lines which were measured using the highly precise D-GPS equipment. The

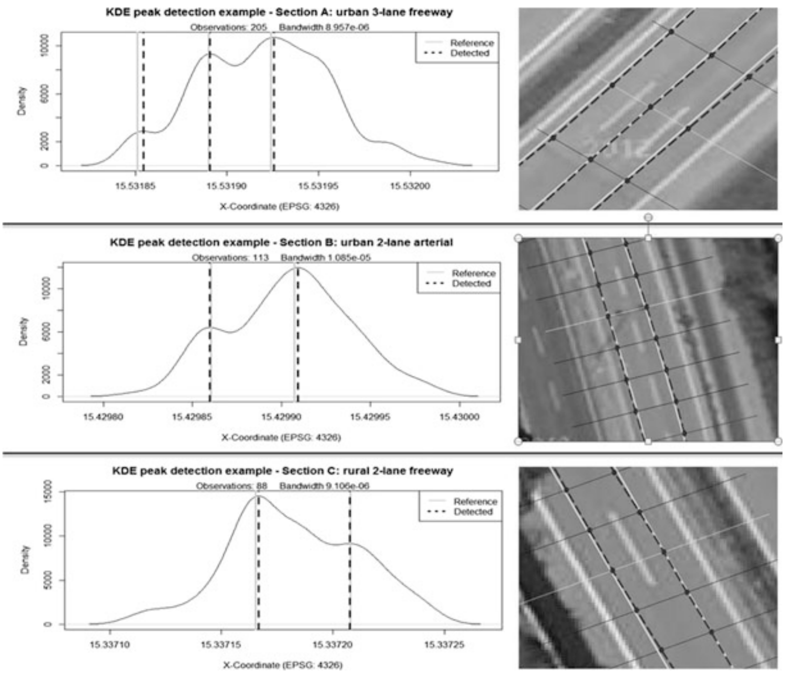


Fig. 4 Exemplary road sections in the three study areas, for which lane center lines are estimated with high accuracy (estimation close to reference). Background Map: basemap.at

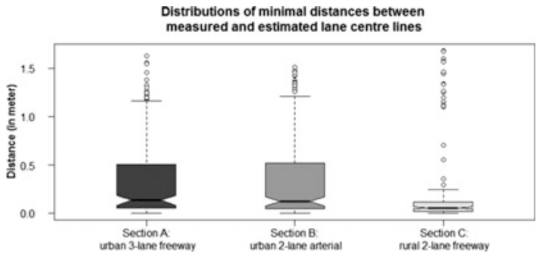


Fig. 5 Boxplot of the distribution of deviations of lane center lines between the algorithmic estimations and the reference measurements in the three study areas

median distance is 0.135 m for section A, 0.123 m for section B and 0.056 m for section C. The distribution of distances leans towards the upper quartile in all observed study areas, with outliers up to 1.882 m. Considering lane widths between 2.75 and 3.75 m, the estimation of lane center lines performs with high accuracy.

4 Conclusion

This study was carried out within the Austrian research project "LaneS" with the goal to generate a lane-specific transportation network as a basis for future ITS applications and automated driving. A wide set of test runs with different measurement equipment (high precise D-GPS and low-cost GPS receivers like Qstarz data loggers and different smartphone GPS positioning apps) were conducted on three road sections (urban 3-lane freeway, urban arterial, rural 2-lane freeway).

First, the position accuracy of the measured vehicle trajectories from the low-cost GPS receivers was checked against the reference trajectory from D-GPS within a roadway and trip based distance analysis. Lateral errors with a median of 1.5–3.3 m were determined in the roadway based distance analysis. Best results were achieved here with the Qstarz data logger with an update rate of 5 Hz. Some smartphone apps lead to similar good results, but having a higher standard deviation, especially on the freeway sections. In the trip based distance analysis, also the total positioning error (lateral and longitudinal) was checked for trajectories based on measurement devices inside the same vehicle. The exemplarily results for the Android GPS logger on the urban 3-lane freeway showed that looking at the total error, 22.8 % of distances are less than 1.5 m, but 51.2 % by considering only the lateral error. This means that the longitudinal error, which can also be time based, is an essential part of the total positioning error in the analysis.

The generation of the lane-specific transportation network was realized with kernel density estimation, which is a non-parametric probability density function. The idea here is that the density maxima of vehicle positions at a road cross section should correspond to the positions of lane center lines of a road. Moreover, the number of estimated density maxima indicates the number of lanes on a road. As a result the maxima of the density curve applied on several trajectories is the estimation of the center of the lane. Compared with the measured lane center line based on the D-GPS measurement, the distances are less than 0.14 m in median for all three road sections. Hence, the presented methodology for generating lane-specific transportation networks provides accurate estimations for lane center lines for several road characteristics in terms of different speeds, lane width and topology.

The accuracy of estimated lane center lines is diminished especially in areas with unfavorable environmental conditions or complex road situations. This underscores the dependency of the developed methodology on the positional accuracy of the input data. The assumption that the highest density of vehicle trajectories corresponds to lane center lines does not apply for erroneous input dataset with a significant accumulation of positional errors. Thus, further research is required with regard to dealing with a high level of distortion in the positions of GNSS vehicle trajectories in order to apply the developed methodology comprehensively on a road network, irrespective of road complexity and environmental conditions.

References

1. Gulgen, F., Gokgoz, T.: Selection of roads for cartographic generalization. In: Proceedings XXI ISPRS Congress 2008, Beijing (2008)
2. Böhm, M., Schneider, T.: Anforderungen an Positionierung und Referenzierung im Bereich kooperativer Systeme. In: Proceedings Symposium für Angewandte Geoinformatik, Salzburg (2008)
3. Czerwionka, P., Wang, M.: Optimized route network graph as map reference for autonomous cars operating on German Autobahn. In: Proceedings 5th International Conference on Automation, Robotics and Applications (ICARA), Wellington (2011)
4. Davies, J., Alastair, R., Hopper, A.: Scalable, distributed, real-time map generation. In: IEEE Transactions on Pervasive Computing, vol. 5(4). IEEE Computer Society, Los Alamitos (2006)
5. Sato, N., Takayama, T., Murata, Y.: Estimating the number of lanes on rapid road map survey system using GPS trajectories as collective intelligence. In: Proceedings 15th International Conference on Network-Based Information Systems, Melbourne (2012)
6. Knoop, V., Buist, P., Tiberius, Ch., Arem, B.: Automated lane identification using precise point positioning. An affordable and accurate GPS technique. In: Proceedings 15th International IEEE Conference on Intelligent Transportation Systems, Anchorage (2012)
7. Uduwaragoda, E., Perera, A., Dias, S.: Generating lane level road data from vehicle trajectories using Kernel density estimation. In: Proceedings 16th International IEEE Annual Conference on Intelligent Transportation Systems, The Hague (2013)
8. Arneodo, F., Botta, D., Gagliardi, G.: Floating car data for wide area traffic monitoring and forecast. In: Proceedings of 22nd ITS World Congress 2015. Paper no. ITS-2005, Bordeaux (2015)
9. Herrera, J.C., Work, D.B., Herring, R., Ban, X., Jacobson, Q., Bayen, A.: Evaluation of traffic data obtained via GPS-enabled mobile phones: the mobile century field experiment. Trans. Res. Part C **18**, 568–583 (2010) (Washington)
10. Zheng, Z., Rasouli, S., Timmermans, H.: Evaluating the accuracy of GPS-based taxi trajectory records. Procedia Environ. Sci. **22**, 186–198 (2014) (Elsevier)
11. Deng, H., Wickham, H.: Density estimation in R. Electronic publication (2011)
12. Sheather, S.J., Jones, M.C.: A reliable data-based bandwith selection method for Kernel density estimation. J. R. Stat. Soc. **53**(3) (1991) (Wiley, London)

Construction of PCT Girder Bridge Using an Overhead Movable Scaffolding System (MSS)

Chuan Chuntavan, Nuthaporn Nuttayasakul, Martin P. Bae and Huang Aiwu

Abstract Norwegian specialist, NRS AS, had been appointed by the contractor of Busan City, SK Construction Co. Ltd. and Kumho Industrial Co. Ltd., to undertake the design and supply of an Overhead Movable Scaffolding System (MSS) for the construction of the Busan-Yongdo bridge in Korea. The construction began at Pier P4 and end just beginning of the Main bridge at Pier YP1, total of 31 spans. The roadway bridge is a single viaduct, Pre-stressed Composite Truss (PCT) girder bridge. The PCT girder is composed of a PC slab lower deck, a RC slab upper deck and steel diagonals. For the light self-weight and relatively strong bending rigidity, the span length varies from 49.232–80.000 m with the average span weight of 17.0 t/m. There are eleven types of different pier head. The 2232 m long bridge is cast in the MSS. The 175 m MSS is mounted on the piers on each ends, cast the superstructure and moved toward the middle. The scope of the NRS AS work includes the design, steel fabrication, and supply of ancillary equipment for the MSS. This paper presents several design challenges due to site constraints and the bridge structures itself. These challenges include high impact loads on the MSS due to adverse wind load anticipated while working nearby the harbor and heavily traffics in the Busan City with limited clearance, special considerations to ensure efficient installation, bridge concreting and launching of the MSS in the sharp curve.

Keywords Movable scaffolding system · MSS · Pre-stressed composite truss girder bridge · PCT girder

C. Chuntavan (✉) · N. Nuttayasakul
Department of Civil Engineering, Academic Division,
Chulachomklao Royal Military Academy, Nakorn Nayok, Thailand
e-mail: cchuntavan@gmail.com

M.P. Bae · H. Aiwu
NRS Consulting Co. Ltd., 121/90 RS Tower, 32nd Floor, Ratchadaphisek Road,
Din Daeng, Bangkok, Thailand

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_85

1 Introduction

Bridge construction especially for a whole span cast in situ concrete bridge, over deep valleys, water crossings with steep slopes, over highway or railway, or environmentally protected regions can offer many challenges. Conventional Scaffolding or Formwork was formerly built in place and can only be used once. Because of high labor and material cost, the trend today is toward increasing prefabrication, assembly in large units, erection by mechanical means such as "movable forms" and continuing to modify and reuse the forms for other projects. The Movable Scaffolding System (MSS) for bridge construction may offer advantages over conventional method using conventional scaffolding, MSS offers minimal disturbance to surroundings, providing a more concentrated work area for superstructure assembly, and possibly increased worker safety.

The current system, developed by NRS AS, provides the new MSS with the self-launching system used for cast-in situ bridges which can offer many cost-saving advantages to the bridge construction project.

2 Bridge Configuration

The construction of the Busan-Yongdo bridge began at Pier P4 and end just beginning of the Main bridge at Pier YP1, total of 31 spans as shown in Fig. 1 below. The roadway bridge is a single viaduct, Pre-stressed Composite Truss (PCT) girder bridge. The PCT girder is composed of a PC slab lower deck, a RC slab upper deck and steel diagonals as shown in Fig. 2. For the light self-weight and relatively strong bending rigidity, the span length varies from 49.232–80.000 m with the average span weight of 17.0 t/m. The 2232 m long bridge is cast in the MSS.

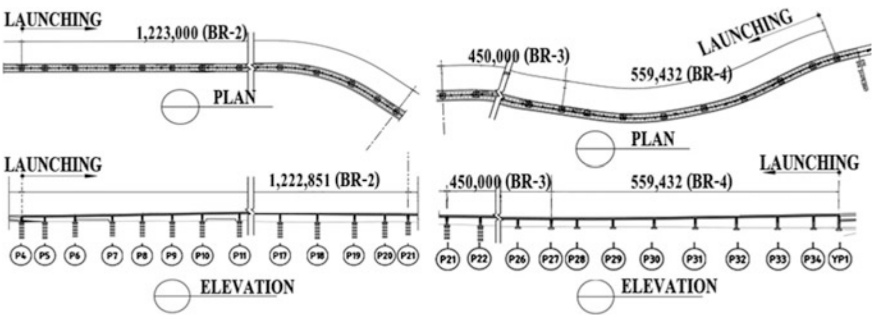


Fig. 1 Bridge layout

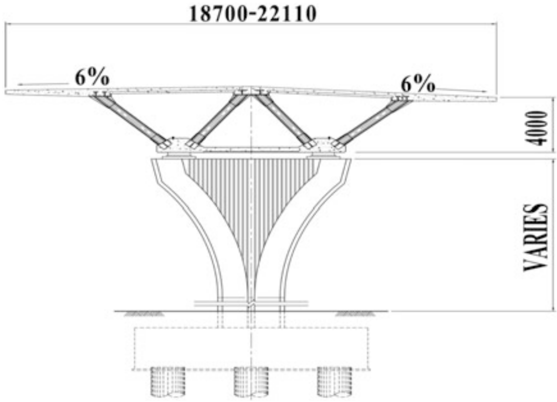


Fig. 2 Bridge superstructure and substructure

The bridge is constructed as a series of continuous spans with the expansion joints at Pier P21, P27 and YP1. It has a minimum horizontal radius of 347 m and a maximum longitudinal slope of 2.0 %. The deck width is 22.11 m for pier P21–P27 and 18.70 m for pier P27–YP1 with constant deck height of 4.00 m. The bridge substructure has pier widths vary from 3.00–5.00 m and pier lengths vary from 9.00–12.50 m. There are eleven types of different pier head combination of pier width and pier length mentioned above.

3 Design Criteria

3.1 Design Code

The structural calculation of the MSS was performed according to the following design rules and specifications: *Steel structure*: Eurocode 3 [1], AISC [2], NS 3472E [3], AASHTO [4], and BS 5950 [5]. *Wooden formwork*: NS 3470 [6]. *Lifting equipment*: F.E.M. 1.001 [7] and DnV rules [8].

3.2 Load

The assumed loads and density are as following: *Density of*: concrete 2650 kg/m^3, steel 7850 kg/m^3, wood 815 kg/m^3. *Working platforms*: upper-working platform of top panel formwork are designed for a uniform load of 2.5 kN/m^2. *Wind*: allowable wind speed for reinforcement and concreting ≤22 m/s, for launching ≤12 m/s. No operation is allowed (MSS parked) when wind speed is 23–30 m/s. The MSS should be secured in the position when wind speed is 31–68 m/s.

3.3 Deflection

Maximum deflection of formwork during concreting is less than or equal to L/400, when L is span length of formwork.

3.4 Material Data

Steel Quality. Since the MSS will be fabricated in China and shipped to Korea after complete fabrication, therefore, material properties are based on Chinese standard. Steel grades are Q345 for main member and Q235 for secondary member. Steel strength is depending on the thickness as shown in Table 1 where f_y is the yield strength and f_u is the ultimate strength.

Fastening Elements. The fastening elements are those for connecting each structural member together for example threaded bolts and pin bolts. The strength of these fastening elements depends on size, code, and grade of the elements as shown in Table 2.

Table 1 Steel properties

Plate thickness, t (mm)	t ≤ 16	16 < t ≤ 25	25 < t ≤ 36	36 < t ≤ 50	50 < t ≤ 100
Steel grade	fy/fu, MPa	fy/fu, MPa	fy/fu, MPa	fy/fu, MPa	fy/fu, MPa
Q345 (16Mn) (Profiles/plates)	345/510	325/490	315/470	295/470	275/470
Q235 (Profiles/plates)	235/340	225/340	225/340	215/340	205/340
Young's modulus of elasticity (MPa)				210,000	
Poisson's ratio				0.3	
Density (kg/m^3)				7850	

Table 2 Fastening elements properties

Bolt diameter, t (mm)		0–16	17–40	41–100	100–160	160–250
CODE	GRADE	fy/fu, MPa			fy/fu, MPa	fy/fu, MPa
NS-180898-1 and 2 (threaded bolts)	8.8	640/800			–	–
NS-180898-1 and 2 (threaded bolts)	10.9	900/1000			–	–
EN 10083-1 (pin bolts)	34CrNiMo6	980/1180	880/1080	780/980	690/880	590/780
	40 Cr	785/980				

Wooden Formwork. Design of wooden formwork is based on following: *Plywood*: thickness 21 mm, Dokaplex formwork standard. *Wooden material*: Norwegian, quality T18 or equivalent NDS. *Formwork ties*: ø15 mm, 900/1100 MPa and ø20 mm, 900/1100 MPa.

Material Factors (γ_m). Ultimate Limit State (ULS): All parts, $\gamma_m = 1.10$; Service Limit State (SLS): All parts, $\gamma_m = 1.00$.

3.5 Design Approach

The global anaylsis of the MSS was performed using Eurocode 3 [1]. The analysis of each part in the MSS system was performed using either Eurocode 3 [1] or AISC [2]. For the AISC, the Load and Resistance Factor Design (LRFD) method was used in the design of steel members.

4 Research Methodology and Structural Analysis

4.1 Research Methodology

Research on the design of the MSS was based on more than 30 years experience of NRS bridge construction equipment. All information from the site supervisors of the bridge construction team were collected, analyzed, scrutinized and finally summarized to build up the best equipment for this bridge in term of performance and efficiency. The research started from collecting all requirements from the client, studying data from past experience, building up the 3D finite element model, analyzing all load cases acting on the MSS, designing all parts of the MSS based on the worst load case, and finally constructing 3D model for both assembly and operation.

4.2 Structural Analysis

Three dimensional (3D) finite element model of the MSS was constructed based on the nominal geometric and material properties listed in Table 1. PC-based STAAD-PRO [9] was selected as the software for structural analysis. The analysis was carried out based on the situations that can occur during the MSS operation period which consists of the concreting stage, support relocation stage, launching stage and parking stage. The model composed of Main Girder, Front and Rear Noses, Transvere Trusses and Hangers generated by frame member in 3D model as shown in Fig. 3.

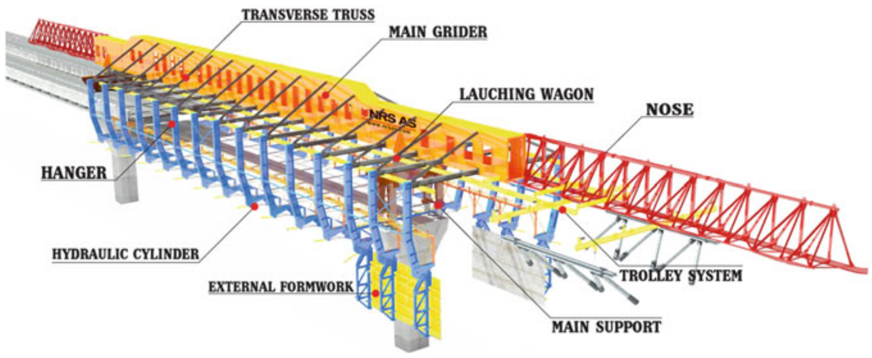


Fig. 3 3D Modeling of the MSS

After the 3D analytical model is completed, loads were applied on the MSS based on the actual weight of concreting bridge span, steel self-weight of MSS, external formworks, and live loads on working platforms. Wind load was also included in the analysis of the MSS for both concreting and launching positions.

In order to reduce the weight of main girder, web openings were utilized all along the length of main girder except over the support areas. A huge force transfers from the main girder to those supports may affect the shear capacity of the web openings, therefore, careful strengthening around the web openings are necessary. The method proposed by Hagen and Larsen [10] to calculate the shear capacity of steel girders with large web openings, based on the shear buckling capacity, as given by Eurocode 3, was used. The design of shear capacity of girder with opening given by

$$V_{bw,mod,Rd} = \left(1 - \frac{D_h}{h}\right)\chi_w c_2 \frac{f_{yw}}{\gamma_{M1}\sqrt{3}} ht \quad (1)$$

but not larger than

$$V_{bw,modcutoff,Rd} = \alpha_c \left(1 - \frac{D_h}{h}\right) \frac{f_{yw}}{\gamma_{M1}\sqrt{3}} ht \quad (2)$$

D_h is the height of the opening and h and t are the clear web height (depth) between flanges and the web thickness respectively. c_2 is an adjustment factor the secondary effects of the openings. α_c is a cut-off factor and χ_w is the buckling reduction factor for shear as given in EN 1993-1-5 [11].

The moment capacity of girder is given by

$$M_{mod,Rd} = M_{buckl,mod,Rd} = \frac{f_y}{\gamma_{M0}} W_{eff,mod} \quad (3)$$

$W_{eff,mod}$ is computed by neglecting all parts of the effective web area that fall within the opening. Horizontal reinforcement, if any, is not included.

Moreover, for girders with openings the verification of shear and primary moment interaction may be given by

$$\frac{V_{Ed}}{V_{bw,mod,Rd}} \leq 1 \tag{4}$$

$$\frac{M_{Ed}}{M_{el,mod,Rd}} + \left(1 - \frac{M_{f,mod,Rd}}{M_{pl,Rd}}\right)\left(2\frac{V_{Ed}}{V_{bw,mod,Rd}} - 1\right)^2 \leq 1 \quad \text{for } \frac{M_{Ed}}{M_{f,mod,Rd}} \geq 1.0 \tag{5}$$

$$\frac{M_{Ed}}{M_{buckl,mod,Rd}} \geq 1.0 \tag{6}$$

V_{Ed} and M_{Ed} are the shear force and primary moment, acting at a vertical section through the center of the opening. $M_{pl,Rd}$ is the plastic moment design capacity of the section, considering the effective area of the flanges and the gross area of the web. $M_{el,mod,Rd}$ is the elastic moment design capacity of the girder net section. $M_{buckl,mod,Rd}$ is the moment capacity design of the net section when only the modified effective areas are considered. $M_{f,mod,Rd}$ is the moment capacity design when only the effective flange areas are considered, and the areas required to support the additional flange forces are subtracted.

5 Design of MSS Main Parts

5.1 Main Girder

The Main Girder (MG) is the principle bearing component, which transfer the design loads to the supports. Concrete loads are transferred from the formwork into the hangers, which are supported on the Transverse Trusses. The Transverse Trusses are then bolted or pinned to the Main Girder. During launching, the Main Girders are supported on the Launching Wagon bogie. During concreting they are supported by the main jacks, 2 per launching wagon. At the front and rear ends there are connections for joining the Noses to the Main Girder. The Main Girder is divided into nine modules with 11.4 m long suitable for transport, connected together on site by bolted connections.

5.2 Transverse Truss

There are 15 Transverse Truss (TT) cantilevering on each side of Main Girder. The main purpose of the Transverse Truss is to transfer loads from hangers into the Main Girders. The Transverse Truss is pinned to the Main Girder on top and bolted on the bottom.

5.3 Nose

A Nose is connected to each end of the Main Girder and is designed to transfer loads to the Launching Wagon during launching. The lower chord of the nose is fitted with a rail on the outer beam. The tip of the Nose is curved to facilitate launching onto the forward Launching Wagon.

5.4 Launching Wagon

The MSS was supplied with three (3) Launching Wagons, one per support. These are the guiding supports for the Main Girder and Noses during the longitudinal- and side launching operations. During concreting, the hydraulic main jacks which located on both sides of the skid beam, will be activated. While in launching position, these jacks will be de-activated and the jacks including pads will be supported by chain blocks connected to the Main Girder supports. There is a braking device at the rear of the launching wagon. This braking device will be used to prevent rolling back of the MSS, when it's going up on the slope.

5.5 Main Support

The MSS was supplied with three (3) main supports i.e. Main support A, B and C. The support legs sit directly on top of the precast concrete blocks and jacks.

5.6 Hanger

The hanger arrangement is the main support for the bottom slab. Each cantilevered transverse truss is equipped with one hanger frame. A hydraulic cylinder is used to fold the bottom part of hanger for launching purpose.

5.7 Formwork

The formwork system consists of steel frames, which are covered with steel sheets to support the bottom slab concrete. The formwork is divided along the centre-line longitudinal axis. The panels are made up of various length from 4.395 m to 7.635 m long elements. The division along the longitudinal axis is necessary to separate the formwork in order to pass the bridge pier, when launching the MSS from a finished span to the next. The transversely division is among other things to obtain the curvature and horizontal radius of the bridge. Adjusting of the formwork system due to cross-fall and cambering is done by mechanical jacks on hangers.

6 Assembling and Erecting Sequence

In order to operate the MSS, the assembling and erecting sequence of the MSS can be described as shown in Fig. 4. The safety check lists need to be implemented before starting each operations. Contractor-engineering office shall carefully plan

Fig. 4 Assembling and erecting sequence of the MSS and PCT

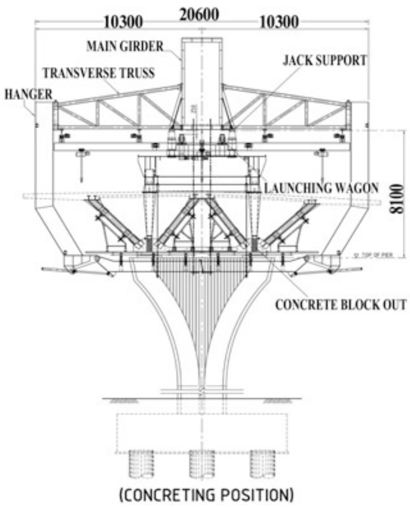


Fig. 5 MSS in concreting and launching positions

each single operation based on the input given in the operation manual and on the relevant conditions for each single operation. The planning shall include relevant obstructions for each span. Verify that all involved personnel are aware of the operations to be performed and that they are informed about the risks and safety measures involved in the different steps of the operation. Check that the hydraulic system is working properly. All personnel working on the formwork prior to securing with tension bars shall be supplied with a safety harness attached to the Hanger Truss. After bringing in and installing the PCT, reinforcing and prestressing steel, now the MSS is ready for casting the first span of the bridge as shown in Fig. 5.

7 Concreting Sequence and Launching Procedures

7.1 Concreting Sequence

Concreting starts above the front pier of the MSS and continues symmetric both in front and back of the pier, and then working back to the previous span. Based upon past experience, the maximum allowable unbalance during concreting is 100 kN (10 tons). This unbalance is defined as the difference of concrete between the right and left sides of the superstructure. By starting the concrete at the front end of the MSS, the Main Girder will have virtually reached its maximum deflection by the time concreting reaches the previous span.

7.2 Launching Procedures

The launching condition is the most extensive of all operations. It involves using almost all the facilities of the MSS. The launching method is affected by the obstructions below the formwork, obstructions on the sides of the formwork, already casted concrete sections, radius of spans and span lengths. The main problem when launching through a radius and adjacent to an existing structures will possibly be the launching rail getting jammed in the launching wagon or the hangers clashing with the existing structures or other obstructions as shown in Fig. 5.

8 Conclusion

The MSS for the Busan-Yongdo bridge in Korea, has many advantages over conventional scaffolding because of high efficiency in achieving rapid cycles (taken average 10 days per one cycle starting from finish concreting, prestressing of the bottom slab concrete, folding out the formwork, launching forward the MSS to new concreting position, folding up the formwork, bringing in & set-up the steel struts & reinforcement, set-up the formwork and finally concreting), lightweight (total weight of 895 tons), simple assembly, less manpower (One Surveyor, 1 MSS Supervisor and crew of 12 men), conveniently adaptable to different cross sections (allowing reuse elsewhere) and allow faster concrete casting operation. The self launching option allows the contractor to avoid using cranes for pier bracket relocation. Additionally, the prefabricated rebar cage and the PCT can be lifted and placed by the MSS.

References

1. Eurocode 3: Design of steel structures, Part 1–1. General rules and rules for buildings. European Committee for Standardisation, 1st edn., Comite Europeen de Normalisation (CEN), Europäisches Komitee für Normung, Central Secretariat: rue de Stassart 36. B-1050 Brussels (1993)
2. AISC: The American Institute of Steel Construction Inc., Load and resistance factor design specification for structural steel buildings (1999)
3. NS 3472E: Norwegian standard. Steel structures design rules, 2nd edn., published by Norges Standardise-ringsforbund (NSF). Norwegian standards Association (1984)
4. AASHTO: The American Association of State Highway and Transportation Officials. Guide design specification for bridge temporary works (1995)
5. BS 5950: Part 1, Structural use of steelwork in buildings. Code of practice for design in simple and continuous construction: hot rolled sections (1990)
6. NS 3470: Norwegian standard, 4th edn. (1989)
7. F.E.M.1.001: Rules for design of hoisting appliances, 3rd edn., published by Federation Europeenne De La Manutention (1987)

8. DnV Rules: Det Norske Veritas. Rules for planning and execution of Marine operations. Part 2, Chapter 5: Lifting (1996)
9. STAAD-PRO: Research Engineers, Inc.. US Corporate Headquarters. 22700 Savi Ranch, Yorba Linda, CA 92687, U.S.A.. Tel: (714) 974-2500. Fax: (714) 974-4771. E-Mail: info@reiusa.com. Internet:http://www.reiusa.com
10. Hagecenn, N.C., Larsen, P.K.: Shear capacity of steel plate girders with large web openings. Part II: design guidelines. J. Constr. Steel Res. **65**, 151–158 (2009)
11. EN 1993-1-5: Eurocode 3 Design of steel structure. Part 1–5: Plated structural elements. CEN (2006)

Part XIV
Maritime—Users, Tasks and Tools in the Maritime Domain

Prediction of Human Postural Response in Shipboard Environments

Nicholas Bourgeois and Robert Langlois

Abstract In order to measure the human response to ship motion, postural stability experiments were performed during an eight-day heavy-weather sea trial in the North Atlantic Ocean. Data were recorded for thirteen participants as they performed a cognitive task without using their hands to maintain balance. The collected data were used to determine what sensory inputs the human participants received from their environment, the control forces and torques their muscles produced to react to those inputs, and the final body motions that were used to maintain balance. These experimental data sets were used to tune the control system of a human postural stability model. Analysis of simulation results indicate that there was a consistent relationship between the accuracy of the model's motions and the direction of greatest ship angular motions, regardless of which direction the subject was facing. This result is significant because it demonstrates that an inverted pendulum model can be used to represent human motion both in the anterior/posterior direction and the lateral direction.

Keywords Human postural stability · Sea trial · Ship motion

1 Introduction

One of the applications of human postural stability modelling is to quantify the effects of moving environments on human performance. For instance, shipboard motion environments can hinder a person's ability to perform specific tasks that in turn can result in increased costs of commercial vessels and less effective warships [1]. Opportunities may exist to introduce postural stability analysis earlier in ship

N. Bourgeois (✉) · R. Langlois
Department of Mechanical and Aerospace Engineering, Carleton University,
1125 Colonel by Drive, Ottawa, ON, Canada
e-mail: Nicholas.Bourgeois@carleton.ca

R. Langlois
e-mail: Robert.Langlois@carleton.ca

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_86

design cycles thereby leading to more-relevant ship design and operations. This paper details the development of a human postural stability model that accurately predicts human response to ship motion. The human body is represented as a four-link spatial inverted pendulum model that allows for representative motion of ankles, knees, hips, and neck. The model's control system calculates the torques that must be applied at each of the joints in order to maintain balance. The parameters of the control system were tuned using biometric data gathered during a heavy-weather sea trial so that the motion behavior of the model and magnitude of the simulated muscular torques are similar to those measured at sea. This paper presents how this was accomplished and discusses the significance of the results.

2 Sea Trial Data Acquisition

In order to obtain heavy-weather experimental data, a series of postural stability experiments was carried out on board Canadian Forces Auxiliary Vessel Quest during the Q-348 sea trial that took place from November 20th through November 28th, 2012. Thirteen participants took part in postural stability and cognitive efficiency experiments carried out by researchers from Carleton University and Defence Research and Development Canada Atlantic. Participants took part in 90 min experimental sessions in which they were asked to maintain balance while performing a cognitive task, divided between 0°, 45°, and 90° standing orientations with respect to the ship centerline. The data measurement procedures in this experiment were unique in that they attempted to measure all of the sensory inputs that a person might use to maintain balance. Similar experiments were carried out aboard Quest trial Q-303 in 2007 [2]. Those experiments were different in that they incorporated a variety of tasks and measurement strategies, with the postural stability portion analyzed using only videotape recordings. For this project, much more precise data acquisition techniques were utilized. The sources that were used to predict human postural response included the following items.

- Two Microsoft Kinect sensors and iPi Studio software, which uses the Kinect's depth camera to determine the positions and orientations of human body segments in 3D-space.
- Tekscan F-Scan data acquisition system which is used to measure the distribution of pressure sensed by a person's feet.
- A Crossbow inertial sensor placed near the feet of the participants that measures ship motions.

The outputs of each of these data sources will be described in more detail in the following sections.

2.1 Kinect Motion Capture

The Kinect sensor is a motion tracking device produced by Microsoft that can be used to measure the distances between the sensor and objects in its field of view. A sample of two depth stream images is shown in Fig. 1 (left). The different colours represent varying distances from the camera. The image in Fig. 1 (right) shows a screenshot of the iPi Studio software, which maps the depth images from the two Kinect cameras onto a human skeleton. From this mapping, the positions of all of the joint positions indicated in Fig. 2 can be calculated.

2.2 Insole Pressure Measurements

The Tekscan F-Scan system consists of two instrumented insoles that are placed into a person's shoes. A photograph of an insole placed inside a shoe with the connection interface attached to the ankle is shown in Fig. 3 (left). Pressure measurements are made by pressure-sensitive ink that has an electrical resistance proportional to the compressive force applied to it. The resulting data sets provide a time history of the distribution of pressure sensed by each of the person's feet. From this pressure distribution, one can calculate the overall vertical force acting on each of the feet, and the location where this force acts. A screenshot of the Tekscan software showing this pressure distribution in shown in Fig. 3 (right).

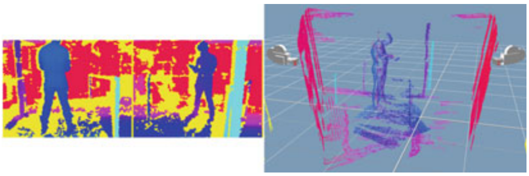
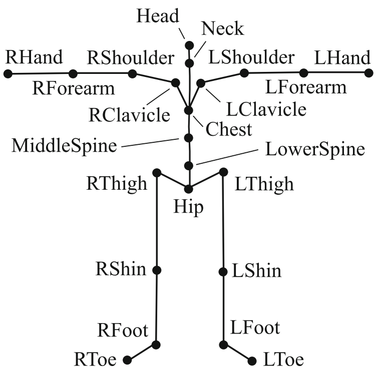


Fig. 1 Kinect depth data recorded simultaneously from two separate cameras (*left*) and Kinect depth data from two sensors arranged in 3D space in iPi Studio (*right*)

Fig. 2 Human skeletal structure output by iPi Studio software



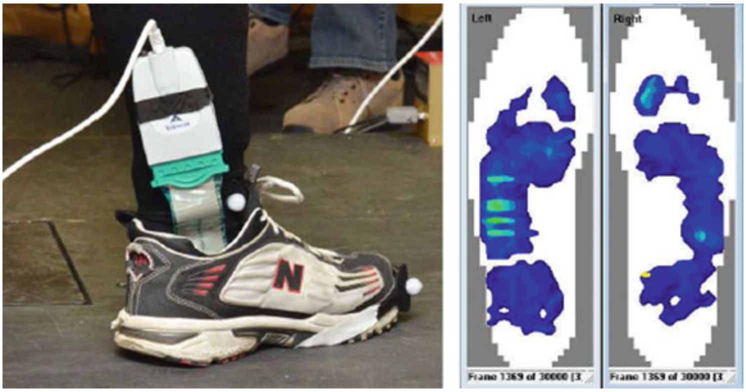


Fig. 3 Pressure-sensing insole with connection at ankles (*left*) and image from Tekscan software (*right*)

3 Ground Reaction Moments

One of the useful metrics that can be derived from the motion capture data is the location of the body centre of mass (CoM) at each time step. This is done by calculating the CoM of individual body segments and by combining them with a weighted average, where the weights are the percentage of the total body mass that each segment contains. The final result is projected onto the ground plane. The locations of the CoM of each individual body segment were taken from literature [3].

The centre of force (CoF) is the location in the ground plane where the vertical force acts to balance a person's weight. In quiet standing the locations of the CoM and CoF are co-located; however in a constantly moving environment, such as at sea, they are not. An image of the area of a person's stance with the CoM and CoF labelled is shown in Fig. 4 (left). The location of the CoF can be calculated from the sea trial data using the following steps: (1) sum together the pressure measurements beneath each foot to obtain the total vertical force acting on each foot; (2) determine the insole location where this forces acts; (3) using the motion capture data,

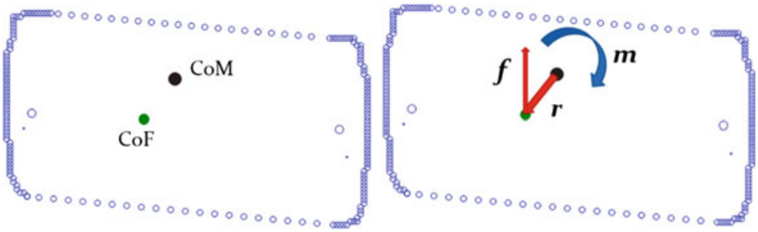


Fig. 4 Centre of mass (CoM) and centre of force (CoF) positioned within person's stance (*left*) and a diagram of person-ship reaction moment calculation (*right*)

determine where each foot force is located relative to the other; and (4) average the foot force locations using their magnitudes as weights to calculate the overall CoF location, and sum the forces to calculate the overall vertical reaction force.

Once the CoM and CoF have been calculated, the moment that the vertical force acting on a person's feet generates about their CoM can be calculated. This measure is representative of the effort that a person would have to exert in order to maintain balance in a dynamic environment. It can be calculated using the metrics labelled in Fig. 4 (right) as the cross product:

$$\boldsymbol{m} = \boldsymbol{r} \times \boldsymbol{f} \tag{1}$$

where $\boldsymbol{r}$ is the position vector from the CoM to the CoF and $\boldsymbol{f}$ is the total vertical force measured by the insoles.

4 Human Postural Stability Model

The objective of this project was to develop a model of a human standing in shipboard conditions that would be sufficiently complex that it would move similarly to an actual human in those conditions, yet be simple enough that the simulation computations be manageable. In this case, a four-link spatial inverted pendulum model was selected. This model allows for representative motions of ankles, knees, hips, and neck. A visual representation of the reduction from a full body to a four-link inverted pendulum is shown in Fig. 5. Each inverted pendulum segment has equivalent mass properties to the full-body segments that it is intended to represent.

The equations of motion for the system were derived based on an automated method of deriving the equations of motion for open chain-like mechanisms in a form that lends itself well to matrix mathematics [4]. The model was designed to allow for a total of 18 degrees of freedom of motion. This allowed for 3 rotations at each joint and for 3 translations and 3 rotations for the ship motion. The final form of the equations can be written as a function of the relative angular velocity components, $\dot{\boldsymbol{x}}$, of the system in the form of:

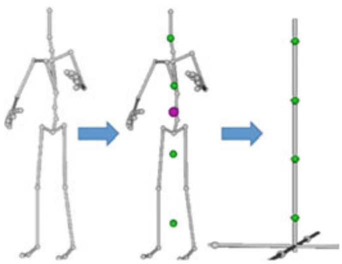


Fig. 5 Reduction of iPi Studio skeleton to four-link inverted pendulum model

$$A\ddot{x} + b(\dot{x}) = d \tag{2}$$

where A contains the mass-related terms, b contains the linear and nonlinear velocity-related terms, and d contains the external forces and moments acting on the system. This form of the equations can be rearranged to solve for the accelerations as:

$$\ddot{x} = A^{-1}[d - b(\dot{x})]. \tag{3}$$

which allows the system to be easily discretized and simulated with respect to time. The model was validated by comparing its performance to a number of existing shipboard postural stability models [5].

5 Control System Design

The objective of the control algorithm for the four-link inverted pendulum is to maintain stability of the inverted pendulum while it experiences six-degree-of-freedom ship motion, while also having a motion response that can be correlated to the forces and motions that were measured during the sea trial. Achieving this objective would imply that the model behaves similarly to actual human balance control. Figure 6 shows a visual representation of the design problem. It shows the control moments that are applied to each link, and the relative segment sizes represent the balance of mass between each of the links. Since the control moments are generated at the joints, they will contribute equal and opposite torques to each of the two links they are connected to. One can immediately observe that care must be taken to balance the moments within each individual link such that the desired orientation is reached, and to coordinate the moments between links such that the very large torso mass does not cause the lower links to buckle.

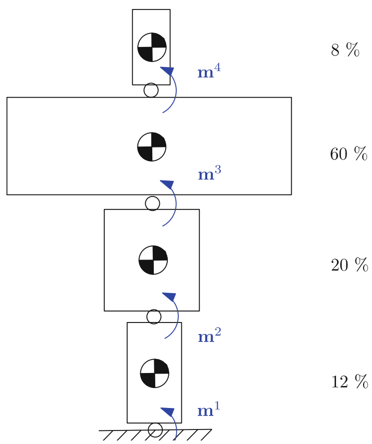


Fig. 6 Diagram of four-link inverted pendulum which shows the relative mass of each link and how the control moments are applied to the system

One of the sources of sensory information that humans use to maintain upright stance is the pressure sensed under each of the feet, otherwise known as somatosensory information. This provides information on the location of the CoM within the person's stance, and indirectly provides knowledge of the direction of the total acceleration vector acting on the person. In normal standing conditions, this acceleration is aligned with the direction of gravity, and humans tend to align themselves with it. In shipboard environments, however, a person would only be able to estimate the total acceleration due to gravity plus the translational acceleration contribution from the ship's motions. Figure 7 illustrates these two cases. Additionally, if there was no inertial visual reference for a person to align themselves with, as was the case during the sea trial experiments, this would be the primary method that a person would use to maintain upright stance.

A control algorithm for the human postural stability model was developed based on the total acceleration vector of the ship. While processing the motion capture data, it was observed that there were many cases where a vector from the subject's ankles to their CoM would align with the total acceleration vector. It was also observed that the subjects would tend to minimize the relative motion between their own body segments. This makes sense because it would minimize the energy required to maintain balance and facilitate completion of the cognitive task they were assigned during the experiments. A simple controller was designed based on these two premises. Its two primary objectives were to:

1. minimize the angle between the body and total acceleration vector; and
2. minimize the angular displacement between body segments.

In order to facilitate the controller development, some degrees of freedom of the inverted pendulum model were constrained. The yaw motion of each joint was constrained because it was determined that the very small amount of yaw motion experienced by the participants did not have a significant impact on their response. Additionally, the knee roll joint was constrained while tuning the controller in order to meet design objectives in both roll and pitch directions simultaneously. It was determined that this constraint did not have a significant effect on the model's

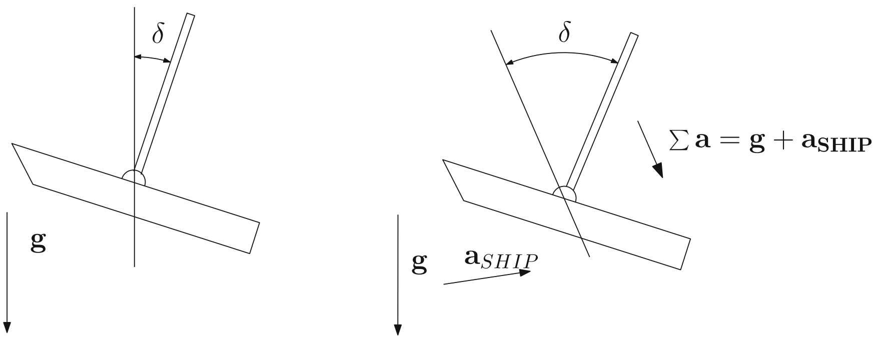


Fig. 7 Acceleration due to: gravity only (*left*) and gravity and ship motions (*right*)

ability to represent a human since human knees do not roll, and maintaining the pitch motion at the knee still allows the model to achieve relative motion between the ankle and hip joints.

6 Results

The success of the control algorithm was evaluated using two metrics: how well a vector directed from the person's ankles to their CoM matched the total acceleration vector, and how well the reaction moment at the ankle correlated between the model simulations and the experimental results (as calculated by Eq. 1).

Figure 8 shows sample results comparing the first metric. The graphed data are the roll and pitch angles that the vectors make with respect to the ship's vertical direction in the ship frame.

Figure 9 shows the results comparing the ship-person reaction moment between the simulation and experimental data.

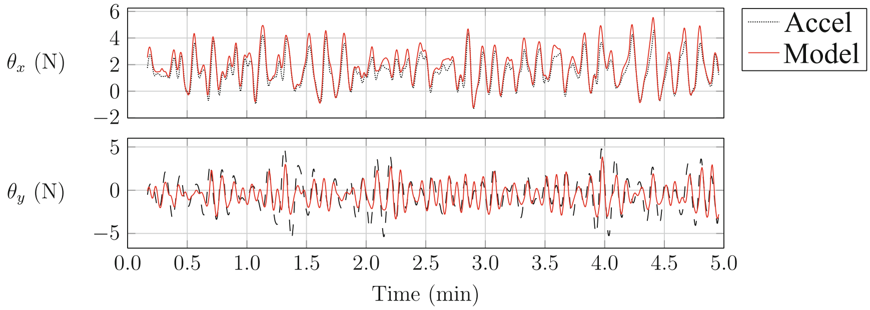


Fig. 8 Comparison of total ship acceleration angles (Accel) with the human model's CoM angles (Model) in the anterior/posterior (θ_x) and lateral (θ_y) directions

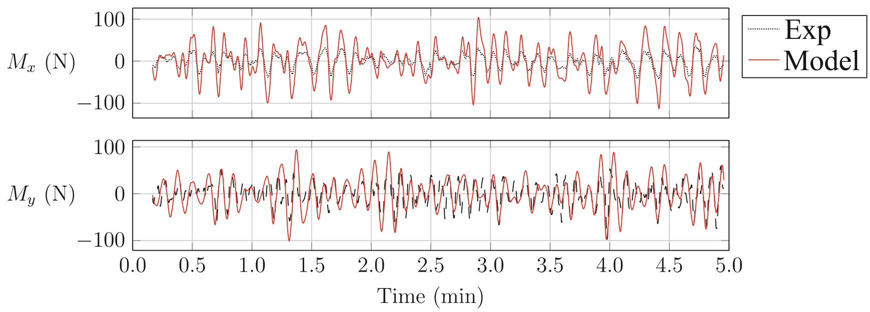


Fig. 9 Comparison of human-ship moments between experimental data (Exp) and simulation data (Model) in the anterior/posterior (M_x) and lateral (M_y) directions

Model simulation and experimental results were compiled for 49 test cases divided over the three standing orientations with respect to ship centreline. Data sets were compared by calculating the Pearson correlation coefficient between them. Figure 10 shows the correlation results for the ship-person reaction moments. The data points have been spread about their X-axis values (direction) for visualization and interpretation purposes.

These results show an interesting trend: when the subject was standing at 90°, the model best matched the experimental results in the pitch direction, and when the subject was rotated to 0°, the results best matched in the roll direction. This indicated that the performance of the model, in comparing it to how well it tracked the acceleration vector, was not related to strictly pitch or roll performance. This was an interesting result because the model was not symmetric since roll motion was constrained at the knee. This result suggested that the performance of the model was based on the ship's motion rather than the orientation of the subject. Metrics were derived to compare this result based on: (1) ship angular motion and (2) correlation of the acceleration vector with the simulated CoM vector discussed previously. Figure 11 shows the results of this comparison, where the X and Y axes represent roll and pitch magnitudes, respectively. The 'Ship Motion' data is determined by calculating the root mean square of the angular motions, and is scaled by their maximum range.

Figure 11 shows that there is a clear correlation between how well the model matched the behavior of the experiment participants, and the direction of greatest ship

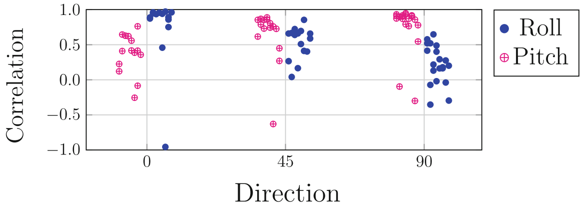


Fig. 10 Correlations for comparing experimental moments to simulation moments in the person's roll and pitch directions

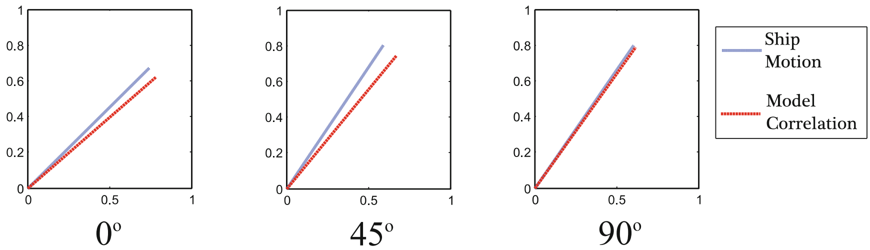


Fig. 11 Graphs that compare simulation correlations with ship and subject angular motion magnitudes

motion. The average angle difference between the ship angular motion and the simulation correlation for all 47 test cases was 8.85° with a standard deviation of 6.03°.

7 Discussion

The simulation results of the four-link spatial inverted pendulum model demonstrate that an inverted pendulum-based model can accurately reflect the motion of a person on board a ship, in the plane of greatest ship motion. This is an interesting result because it suggests that the appropriateness of the inverted pendulum model is based on the ship's motions rather than the physical properties of the person as is common in most models [6–8].

The Quest sea trial was one of few human postural stability experiments which has been carried out on a ship in the ocean rather than on a motion platform in the laboratory. The advantages of having actual ship motion are that the possible range of translational motions are much greater on the ship and the subjects have time to become acclimated to the ship motions in advance of the individual trials which provided a more accurate representation of shipboard working conditions. The disadvantages include the lack of repeatability of the motions and the difficulty of calibrating a variety of sensors in a constantly-moving environment.

Concerning the model's control system, there is much additional complexity that could be added to it. The control system presented here was only based on how well the person matched the orientation of the acceleration vector and was evaluated based on ship-person reaction forces (foot somatosensory) only. The model could be further developed to include vestibular and visual cues, as well as increased complexity in the control law itself by incorporating state estimation, time delays, and sensory input weighting. These sensory inputs and control strategies have been successfully incorporated into 2D postural stability models, although they become markedly nontrivial when applied to spatial dynamic models with multiple degrees of freedom.

The authors believe that the current implementation is a reasonably accurate representation of how the subjects would have maintained balance during the sea trial because of the cognitive task that the participants were required to perform. It was likely that since their attention was not focused on maintaining balance that most of their responses to the ship motion, aside from stepping events, were passively determined by the central nervous system, which is a standard procedure since people commonly stand, walk, run, and perform other motions without actively determining how their bodies will move from moment to moment. The combination of the task and the fact that the experiments took place in an enclosed area of the ship means that the usefulness of visual feedback in maintaining postural stability would have been minimized, as well as the contribution from vestibular information as research shows that it is primarily used to compensate for latency in visual feedback in maintaining a stable optical view of the environment [9].

It is certain that there are measurement errors associated with all of the data recorded during the sea trial that became magnified when data sources were combined and when noisy data was filtered. Of special note is that the unprocessed Kinect sensor measurements have an uncertainty of 1–2 cm at the distance they were used, and that the insole data had to be calibrated in post-processing due to not being able to be calibrated in a constantly-moving environment. That said, the motion capture CoM motions did independently correlate with the measurements made of total acceleration made by the inertial sensor, and it is more likely that if the experimental error had had a significant effect on the results, it would have prevented them from showing the clear correlation between the model and the experimental results that was observed.

8 Conclusion

This paper presents the development of a spatial human postural stability model that is used to model human postural response to ship motions. The model has been tuned with experimental data collected during a heavy-weather sea trial, and accurately mimics the motion of a human's centre of mass in the plane of greatest ship motions. This result is significant because it demonstrates that a multi-link inverted pendulum model can accurately model human motion, and that the performance of the model is a function of the motion environment, rather than how well an inverted pendulum model matches a human in the anterior/posterior or lateral directions.

References

1. Graham, R., Baitis, A.E., Meyers, W.G.: On the development of seakeeping criteria. Nav. Eng. J. **104**, 259–275 (1992)
2. Colwell, J.L., Allen, N., Bos, J.E., Bridger, R., Duncan, C., Elischer, P., Grech, M., Green, A., Hogervorst, M.A., MacKinnon, S.N., Munnoch, K., Perrault, D., Roger, W., Schwartz, R., Valk, P., Wright, D.: Human performance sea trial QUEST Q-303. In: International Martime Conference (2008)
3. Winter, D.A.: Biomechanics and Motor Control of Human Movement, 4th edn. John Wiley and Sons Inc., Hoboken, New Jersey (2009)
4. Amirouche, F.M.L.: Fundamentals of Multibody Dynamics. Birkhauser Boston (2006)
5. Bourgeois, N.R., Langlois, R.G.,: Derivation and validation of a spatial multi-link human postural stability model. In: CCToMM Symposium on Mechanisms, Machines, and Mechatronics (2015)
6. Sovol, A.W., Bustamante Valles, K.D., Riedel, S.A., Harris, G.F.: In: 32nd Annual International Conference of the IEEE IMBS (2010)

7. Mckee, J.: Simulating the Effects of Ship Motion on Postural Stability Using Articulated Dynamic Models. Master's thesis, Carleton University (2004)
8. Langlois, R.G.: Development of a spatial inverted pendulum shipboard postural stability model. In: International Conference on Human Performance at Sea, Glasgow, Scotland, UK (2010)
9. Carpenter, R.: Reddi, B: Neurophysiology: A Conceptual Approach, 5th edn. Hodder Arnold, London (2012)

Maritime and Port Activity Analysis Tool

Mahdi Safa and Brian N. Craig

Abstract Safety and productivity are crucial considerations when evaluating the effectiveness management has exhibited in their control of the operations involved with maritime transport and port management. Recent maritime and port management research has been concentrated primarily on the use of state-of-the-art technologies that result in the level of productivity becoming less dependent on the workforce. However, industry experts have consistently recognized and been concerned about challenges that remain in the enhancement of workforce productivity and safety. No readily available guide describing how to conduct activity analysis has existed until now. Hence, this research project involves the investigation and application of "activity analysis" methodologies. Activity analysis denotes the extension of simple work sampling so that it becomes a continuous productivity and safety improvement program. Such a program entails a cyclical process designed to quantify the activities of craft workers to identify barriers to safety and optimal productivity, to implement improvements for reducing or eliminating these barriers.

Keywords Activity analysis · Workforce productivity · Safety · Maritime and port operation

1 Introduction

The task of identifying port workforce productivity and safety challenges and opportunities will enable a port's operations to carry out the tasks necessary for the achievement of its goals is of fundamental importance. The primary objective of this research is to provide a comprehensive process for improving port workforce productivity and safety; the application of which will help port managers identify productivity and safety issues related to their operations and will provide a

M. Safa (✉) · B.N. Craig
Lamar University, Beaumont, TX, USA
e-mail: msafa@lamar.edu

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_87

framework for determining the causes of and potential solutions to any challenges. Because no single method or program can include consideration of every individual port operation/project scenario, a further objective is to develop the program in a generalized format that can be flexible as required so that port managers will be able to modify the method to accommodate their own site-specific needs. For the purposes of this study, the program has be developed for different port operations in Texas, with the goal of making the system flexible enough to be applicable for other ports.

Recent port management research has been concentrated primarily on the use of new methodologies that result in the level of productivity becoming less dependent on the port workforce [1]. However, port industry experts have concerns about challenges that remain in the areas of human resources and the enhancement of workforce productivity. New best practices could address the most prominently cited challenges that contribute to ineffective operation in many ports: overstaffing, inefficient work procedures, poor skills and training, and unreliability. The developed method in this study represents a comprehensive and updated definition of one of these practices: work sampling. Casual observation of typical ports reveals that historically successful productivity practices have not been implemented in most port projects and operations. The incorporation of such practices would result in an overall roadmap (methodology) for the effective management and improvement of port workforce productivity, which in and of itself would constitute a significant productivity-related innovation. To meet this objective, this research project will involve the investigation and application of “activity analysis” methodologies. As a productivity measurement tool, activity analysis denotes the extension of simple work sampling so that it becomes a continuous improvement program. Such a program entails a cyclical process designed to quantify the activities of port workers, to identify barriers to optimal productivity and safety, to implement improvements for reducing or eliminating these barriers, and to quantify any changes in activity percentages.

Without quantification prior to and following a change, any improvement cannot be considered effective since its value is unmeasurable. The benefit of an activity analysis methodology is that it enables port managers to understand real work percentages, implement improvements, and quantify changes to “direct work,” all with a degree of statistical accuracy. A port manager cannot successfully manage a port operation (project) without adequate information. Activity analysis not only provides such information, but also offers an effective means of applying it.

This research project has the potential to provide tangible benefits in the area of human resources management for port operations: enhanced accuracy of human resources planning and scheduling, clarification of roles and responsibilities, improvements in logistics and human traffic patterns, encouragement of a culture of performance and an environment focused on work process, ability to estimate required activity levels and minimum staffing levels by operational category, descriptions of new or modified job requirements, and the provision of clear and meaningful productivity objectives for craft workers. The validation of this work will require the acquisition of project- or activity-level productivity data, which will facilitate the evaluation of the strength of the relationship between the metric and the productivity of the port workers. The results of this study should not be expected to provide a silver bullet that will improve port workforce productivity and safety overnight, but they can instead offer an avenue for the creation of a process of cyclical improvement.

2 Background

Maritime and port management experienced a paradigm shift and fundamental changes in the past decades. New methodologies and technological advances have improved the productivity and the occupational health and safety of the maritime and port operation [2, 3]. Productivity is measured for two primary purposes: to control project (activities) costs and scheduling and to obtain data for planning future projects (activities) [4]. In a competitive port industry environment, a goal of all port organizations should be to decrease costs in an attempt to increase market competitiveness and profits. An integral feature of effective port management is the enhancement of workforce productivity and safety which is affected by logistical networks, port regulations, trade policies, changes in markets and services, the management culture, and technological advances. To come to grips with these myriad of human resources challenges, port managers must have access to accurate and updated information about the performance of their workforce because the collection and dissemination of such information are essential for improving the management of that workforce.

Port worker performance can be measured and reported in several formats, the most popular of which are direct productivity unit rates, rework percentages, and activity levels. Most productivity improvement processes are based on one or more of these metrics, each of which measures work in a distinct way and provides a useful contribution to an understanding of actual workforce performance. A statistical technique employed in relation to activity levels is work sampling, which involves an observer collecting a series of random observations from the worker population [5, 6]. For each observation, the observer instantaneously determines the activity of the worker, and then records it in one of several activity categories. It is acknowledged that research related to this type of activity analysis for port management has been limited, with no recent studies having been published in this area.

3 Methodology

The initial stages of this research will begin with the employment of practices that are either widely accepted on the basis of experience or for which there is strong statistical evidence of an impact on port worker productivity. For this study, the proposed activity analysis tool will be based on the 2010 guideline developed by the Construction Industry Institute (CII), according to which, activity analysis is considered an extension of the work sampling technique so that it becomes a continuous improvement process that includes two parts: (1) workforce assessment and (2) the continuous improvement process. Following a thorough literature review and detailed conversations with port experts, the CII activity analysis tool will be modified, adjusted, and customized to create a comprehensive "maritime and port activity analysis tool [7]. This tool will represent the coupling of the updated workforce assessment method with the continuous improvement process that is the strength of activity analysis and that differentiates it from work sampling. The proposed new tool will have the potential to improve port workforce productivity.

The port industry requires a continuous productivity improvement program based on a strong workforce assessment technique that should provide a method for quickly assessing productivity levels and that should be capable of identifying productivity issues, proposing and implementing improvements that will increase productivity, and reassessing productivity levels in order to quantify the effects of the improvements. The program must have a cyclical framework so that productivity improvements can be realized throughout the life of the facilities and the projects. No such program currently exists in the port industry. The suggested methodology for achieving these research objectives is as follows:

1. Complete a detailed review of the literature focused on construction productivity, productivity improvement methods, and work sampling.
2. Interview port industry experts who represent major Texan contracting firms, owner companies, and vendors.
3. Develop a comprehensive model (software-based) for a new workforce assessment methodology and continuous productivity improvement process, to be called the Port Activity Analysis Tool.
4. Conduct case studies at a Southeast Texas Port.
5. Improve the methodology according to the lessons learned from the case studies.
6. Validate the tool through testing on sample case projects.
7. Comment on the applicability of the model for other ports.
8. Provide conclusions and recommendations regarding future research.
9. Develop publications and possibilities for deploying the new method in the industry.

The absence of standardized processes for analyzing activities results in organizations, individually defining criteria and developing assessment tools internally for ensuring productive, safe practices are adhered by personnel throughout the job/task. For example, shown below in Fig. 1 is an internally developed method for evaluating safety and productivity during a project by consistently observing and analyzing criteria for safety and productivity in five different work areas, each tailored to this particular Southeast Texas Port.

Date:		Start Time:		Stop Time:			Comments:
Observation #:		Area 1	Area 2	Area 3	Area 4	Area 5	
Safety	Personal Protective Equipment						
	Ergonomic Risk Factors						
Productivity	Direct Work						
	Prep Work						
	Tool and Equip						
	Materials Handling						
	Waiting						
	Travel						
	Personal						

Fig. 1 Hourly observation worksheet in spreadsheet form; criterion are located to the left and divided between safety and productivity

Individuals who are responsible for activity analysis must have a common understanding of the tool, relying broadly on the basic definition of selected factors in the developed tools, narrowing in on circumstantial details unique to their projects that will require alterations, additions, or extra attention to develop tools and equipment that are more capable of capturing what is actually happening in terms of safety and productivity and help work towards a solution for improvement.

"Personal Protective Equipment (PPE)" specialized clothing or equipment worn to minimize the risk of harm to personnel associated with the specific workplace and project. There is no PPE equipment that can serve as a panacea for all risk management associated with a project; in the presence of task-specific risks, task-specific PPE equipment is necessary to address those risks involved that call for more or different safety concerns.

"Ergonomic risk factors" related to work activity and ergonomics which can increase the probability of "Musculoskeletal Disorder" (e.g. forceful exertions; and repetitive/sustained awkward postures).

"Direct Work" in this evaluation refers to the physical effort exerted that is directly initiated by the objectives identified in a project's work activities breakdown. For example, a manufacturing job may involve direct work such as designing, fabricating, welding, and other applications of skill sets; indirect work would be activities such as paid-idle time due to the late delivery of materials producing a bottleneck from the scheduling side of a project management. The more you can increase the amount of direct work used in a project, and decrease the need for indirect work, the higher the project will perform overall.

"Preparatory Work" the act of receiving assignment or determining requirements before starting a task. A vector of market demands and regulation increases is leading much of the indirect work to be automated or outsourced; leading to the discovery of solutions that streamline, automate, or at least simplify the time-consuming indirect work necessary to drastically improve the results of direct work. Precatory work is one of the most vital and least practiced patterns of human behavior; in professional settings it can make the difference between zero and a million dollars easy. A prepared mind foresees danger and has appropriate protection (PPE) ready in case a situation arises requiring potentially harmful engagement with it, ultimately preserving the interests of the individual and the organization by improving safety and, in turn, productivity.

"Tools and Equipment" related work is similar to preparatory work in that it requires almost clairvoyant-level foresight of upcoming task and activity requirements that will, may, or might not necessarily require the acquisition and preparation of tools related to the preparation for all work processes including enough flexibility for obtaining, transporting, and adjusting tools or equipment in time to counteract contingencies.

"Material Handling" is the transporting of materials from one area of the working area to another. The absence of materials on site at the time they are needed is often the main contributing factor to idle time spent by workers awaiting the resources necessary to move forward with productivity and meet objectives. Idle time is not only unproductive, it also contributes to the amount of on the job injuries because mental functions are in dormant anticipation, stress is often characteristic of unexpected delays in material arrival, as well as plane old boredom giving employees the motive to waste time or resources for mere personal amusement on the job.

"Travel entails" workers walking or riding from one area of the working area to another (when that worker is not carrying anything). Having this travel time accounted for to the appropriate supervisors also can reduce the likelihood of personal injury happening on the job and decrease travel times. A project site is no place to mosey or daydream, on many construction sites, this type of behavior can easily get an employee maimed or killed.

Figure 2 shows that a craft worker sits waiting on the necessary materials, tools, or equipment to do his job. While he complies with the PPE required for his head

Fig. 2 Construction site, southeast Texas, November 2015

protection, his boots are not up to safety standards, and lack of fall protection leaves him unprotected from a serious fall; coupled with a lackadaisical demeanor, this poses serious risks of injury or death.

4 Discussion and Analysis

The type of port area observed can affect activity percentages and should be considered when targets are set. In light of the complexity, port sites often have lower direct work rates and greater percentages of preparatory work than do many other jobs/tasks. In fact, any activity with very low tolerances typically have decreased direct work rates because more planning and attention to detail is required to avoid rework. Weather can also have significant impacts on activity percentages. During the summer, the high temperature/humidity on the coast regions could cause increased personal time (work/rest schedules) and decreased direct work rates. It is challenging to anticipate the full effect of a variable factor such as weather on activity percentages. Setting activity targets is essential for establishing activity analysis as an effective continuous productivity improvement process, hence it provides people at different levels of management goals to reach. However, setting targets is difficult because of the variances in project characteristics.

While craft workers under observation can change their behavior, there are some methodologies to minimize and to effectively manage this impact. The team can make random paths at random times throughout the process to avoid anticipation of the port workers. If the routes become routine, the workers could anticipate the observer and adjust their work practices accordingly. It is also important that these routes provide a representative sample of the entire site to gain an understanding of the behavior of all personnel on the project. To further mitigate observation impact, the team employed several techniques while making the walks through the project. The observers, when possible, made mental notes of the crew at a certain location before the crew become aware of an observer's presence. A further distance can make it difficult to discern the activities, while a close distance puts the observer at risk of being noticed. At no point was an observation taken in front of a suspecting crew in order to avoid observing the crew's reaction to being observed.

Another important element of activity analysis is determining an adequate sample size. Determining an adequate sample size is critical to the accuracy of the activity analysis study. As more samples are collected, the results become more accurate as sampling error is reduced. However there is a balance between statistical accuracy and the cost to collect samples. For the maritime and port operation, an error of ± 5 % at a confidence level of 95 % is generally acceptable [7]. It is essential to determine the time period for which the statistical accuracy is to be valid over. In general, the number of samples is for a one hour period. It may be difficult to obtain the total sample size in an hour period, and should be evenly distributed over the sampling-days. Many project management and engineering journals provide the following equation for determining sample size based on desired error, and anticipated category percentages:

$$n = \frac{(z_{\alpha/2})^2 p(1-p)}{s_a^2} \quad (1)$$

where $Z\alpha/2$ is the standard normal variable corresponding to a confidence level of α, p is the anticipated category percentage, and s_a is the error between the true percentage and the estimated. For a confidence level of 95 %, α = 0.05, and corresponds to a $Z\alpha/2 = 1.96$. If the anticipated percentage p is unknown, a value of 50 % (0.5) may be used as a worst case scenario ensuring the number of samples will be overestimated. As stated, the general acceptable values for these variables in the construction industry are p = 0.5, $Z\alpha/2 = 1.96$, and $s_a = 0.05$, which results in a total number of 384 observations. Equation 1 neglects the effects of the finite population correction factor. The above equation assumes that the total population is much larger than the sample size. However, on many construction sites the population of craft workers is not large enough to drive the finite population correction factor to zero. Equation 2 accommodates this correction:

$$n = \frac{1}{\frac{1}{n_0} + \frac{1}{N}} \quad (2)$$

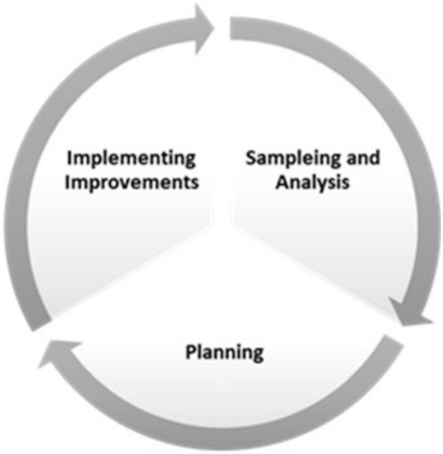
where N is the total population size, and n_0 is the sample size as given by Eq. 1. For example if 1000 port workers were to be studied for an error of ±5 % at a confidence level of 95 %, $n_0 = 384$ and N = 1000. The total number of samples necessary, considering the finite population correction factor, is 278 observations. This is a 27.6 % reduction in the total number of observations necessary had the finite population correction factor not been considered [7].

4.1 Continuous Productivity Improvement

The purpose of activity analysis is first to study and identify safety and productivity challenges, and then to implement improvements to eliminate/reduce these barriers. The authors developed a five-step approach to identifying and reducing productivity barriers. These steps are part of the activity analysis cycle and are illustrated in Fig. 3.

The three steps involve the following: sampling and analysis; planning; implementing improvement. The process starts with activity sampling is done in order to collect a representative data sample. Each discrete data sample or observation is categorized as personal protective equipment (PPE); ergonomic risk factors; direct work; preparatory work; tools and equipment; material handling; waiting, travel; or personal. Once the data have been collected, they are tabulated to determine activity

Fig. 3 Continuous improvement process



percentages. The resulting percentages are analyzed to determine which types of activities are beyond the acceptable range. After the potential causes for unacceptable variances are identified, several potential solutions to improve productivity are considered. These improvements are based on a set of factors that include feasibility, logistics, and costs. Finally, the improvements selected in the planning stage are implemented to increase the direct work rate.

5 Conclusion

Maritime and port activity analysis measures how time is utilized by the labor force along with the safety requirements. In the work sampling technique, observations of what each worker is doing at a particular instant are recorded. The activities of craft workers are typically divided into predetermined categories. The two categories, including safety and productivity, and nine elements utilized in this study include personal protective equipment (PPE); ergonomic risk factors; direct work, prep work, tools/equipment, materials handling, travel, waiting, and personal. The definition of these categories is provided in this paper. Implementing the above improvements requires overcoming barriers that may resist change. A plan of action can be taken once the following activities are followed through: obtain commitment from managers at every level, study each potential course of action, define a schedule and timeline, and investigate the costs that may be involved, consider human resources issues such as training and support, and change relevant forms and documents. It is predicted that the costs of implementing a basic form of port activity analysis would not be costly. It is also estimated that the financial benefits will greatly outweigh the costs. The potential benefits of the proposed Port Activity Analysis Tool include: (1) provides detailed information similar to continuous observation studies, but in less time and at a smaller cost; (2) ability to canvas entire port; (3) no disruption of the work activities of craft or foreman; (4) craft more likely to accept activity analysis compared to continuous observation; and (5) desired level of accuracy possible through statistical techniques. It should be noted that there are no standards for Activity Analysis Processes, another difficult and vital component of the implementing of one. It is recommended that an activity analysis process should be tailored with considerations of a variety of factors affecting productivity and safety in the particular process being evaluated and the circumstances surrounding it.

Acknowledgments The authors would like to acknowledge the support of the Center for Advances in Port Management at Lamar University and the port of Beaumont.

References

1. Bellamy, L.J., Aneziris, O.N., Papazoglou, I.A., Damen, M., Manuel, H.J., Mud, M. and Oh, J. I.: Safety management of safety performance-what factors to measure? Nutritional Care of the Patient with Gastrointestinal Disease, p. 445 (2015)
2. Fabiano, B., Currò, F., Reverberi, A.P., Pastorino, R.: Port safety and the container revolution: a statistical study on human factor and occupational accidents over the long period. Saf. Sci. **48** (8), 980–990 (2010)
3. Beresford, A.K.C., Gardner, B.M., Pettit, S.J., Naniopoulos, A., Wooldridge, C.F.: The UNCTAD and WORKPORT models of port development: evolution or revolution? Marit. Policy Manag. **31**(2), 93–107 (2004)
4. Thomas, H.R., Mathews, C.T., Ward, J.G.: Learning curve models of construction productivity. J. Constr. Eng. Manag. **112**(2), 245 (1986)
5. Aft, L.S.: Work measurement & methods improvement. John Wiley & Sons, Inc., United States of America (2000)
6. Picard, H.E.: Direct labor productivity measurement - as applied in construction and major maintenance projects. AACE International Recommended Proactive No. 22R-01. United States of America: AACE International (2004)
7. Construction Industry Institute (CII): Guide to activity analysis, Univ. of Texas, Austin, TX (2010)

Boat Camping Sailing Yacht: A Study Case of Conscious Yacht Design

Massimo Di Nicolantonio, Jessica Lagatta, Antonio Marano and Andrea Vallicelli

Abstract Camping is a pleasurable yachting activity that enables people to visit beautiful and pristine lagoons, rivers, canals, lakes and marine reserves. The water in these places is often very shallow and there are no landing stages. This calls for a type of vessel suitable for coastal yachting and navigating inland waters. A small yacht can be considered ideal for this purpose. The concept proposed is a small trailerable sailboat for camping, with special performance, from both the naval and ergonomic point of view. This concept introduces the ideas of easy management of the sailing performance, housing activity, beaching and transportation, and an insight into the tactile pleasure of the deck surfaces. This first phase of the research reports a series of critical reflections related to the conscious design of small yachts, taking into account of various aspects: the environmental context, on-board activities, out-board activities and general management of the yacht.

Keywords Boat camping · Easy sailing · Hybrid propulsion · Tactile pleasure · Supple · Trailerability

1 Introduction

Our territory has an immense and widespread environmental heritage, historical monuments and archaeological sites. Also the delicious agricultural products of fishing, hunting and collecting, contribute, according to new

M. Di Nicolantonio (✉) · J. Lagatta · A. Marano · A. Vallicelli
Architectural Department, G. d'Annunzio University, Viale Pindaro 42, 65127 Pescara, Italy
e-mail: m.dinicolantonio@unich.it

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_88

theories of sustainable development,[1] to define the natural capital[2] that today the man has.

The discussion about cultural economy,[3] gives new meanings to the general notion of natural and artistic heritage, replacing it with the new definition of environmental and artistic heritage. It is a way, the environmental heritage looks like a cultural resource, to be enjoyed, to be exploited and, at the same time, to be protected and, if possible, to be played [1]. These first considerations highlight the need to invest in natural capital as compared to the capital produced by man, because even today poorly used. Phenomenon of tourism, among the thematic areas of the contemporary world, can be considered the most strategical and full of attractiveness, both on the socio-cultural, and at the economic point of view; it is so intrinsically connected with the sociological category of leisure.[4] Touristic motivations are most felt by individuals; they may include education, instruction, body care, solidarity, fun and games. The same networks of companies organized like districts, are engaged in an investment on the environmental situation, in order to develop and exploit the material resources, human and cultural factors that characterize the place, because of the needs to support the global competition.

In this scenario of changes related to economic orientations, socio-cultural and production, design can play a center role to generate a new process of conversion of the cultural and environmental heritage in exploitable resources for the future. But to understand the sociological changing sociology contained by the interest of the tourist, one should bear in mind a number of factors [2], as described by the British sociologist John Urry.[5]

2 Boat Camping and Sustainable Tourism

This research introduce an interesting proposal of design solution focused on marine ecotourism,[6] taking account of the new demand of sustainable products and services.

---

[1]With the expression *sustainable development*, introduced in the international discussion from the document of the World Commission for Environment and Development *Our Common Future*, 1987, it refers to all the relationships between human activities and biosphere, such as to afford to human life continue, individuals to meet their needs and different cultures to develop without destroying the global biophysical context.

[2]*Natural capital* is the set of non-renewable resources and systemic capacity of the environment to reproduce renewable resources. The importance of investing in *natural capital* has been highlighted in an article by many hands, entitled *The value of the world's ecosystem services and natural capital*, appeared in May 1997 on Nature.

[3]The field of application of the cultural economy, limited for a long time only to art, according to the Anglo-Saxon tradition, has recently extended its field of action to the show until the museums' economy and heritage, see Benhamou [1].

[4]De Masi (2000, pp. 243–264).

[5]Urry [2].

[6]Marano [4].

Today, the tourist offer must aim to broaden the user groups who would more liking enjoy the environmental benefits generated by thousand relationships between the marine waters with the typical characteristics of coastal landscapes and the characteristics of its seabed. The design vision could point to the promotion of eco-friendly activities closely related with pleasure boating. This would represent an extension of actionable opportunities for all the shipyards to seek new competitive advantages by introducing new eco-friendly products [3]. The establishment of marine protected areas has generated the desire for knowledge of the underwater world; this activity needs new and ecologically correct naval units.

Nautical tourism is one of the forms of marine resource exploitation, spread out in this century in almost all Western countries; yacht are the first image of this attempt of sea colonization by man. Yachting needs of port infrastructure and services; at the same time allows to visit those coasts which are not easily accessible from the inland, thus reducing the overall weight of the tourist sites' merger. The protection of coasts and more in general marine environments, is encouraged by a mature safeguard awareness of the rich touristic heritage; however, it requires appropriate means. This can be a powerful incentive for new products regulated by international norms that consider the environmental heritage a common good to be protected for optimum enjoyment. The nautical filed can therefore have an important part in the growth of a proper eco-marine tourism [4].

Pleasure craft, especially sailing yachts, open interesting perspectives; this particular tipology of yacht can inspire a conscious design focused to translate environmental protection into a wealth's source for the individual; at the same time it could generate new income for local economies. The relationship between nature, society, design and enterprise may have in this coincidence of interests a new opportunity for development and mutual growth. Pleasure crafts can be considered as an extension of man-made water territory tracing by their passage, an ideal network of mending between the earth and the liquid element. A small trailerable sailing yacht can be considered a theme of particular interest for this purpose, related with boat camping activities in the Adriatic sea.

It perform the correct dimension for a pleasure boat and involves many problems to gustily the opening of a research and the discussions based on different topics such as: the performance; processing technologies of composite materials; relation between craft and industry production systems; ergonomics related to living spaces in motion; the formal values of the object; consumption patterns.

3 Concept Description

The proposed typological yacht is considered suitable for coastal and inland's waters navigation; It plays a key role in the research's field of sustainability, strongly linked to that of ergonomics. The market's survey has revealed that the specifically boat for this purpose can be considered a small pleasure boats, trailerable, 7.5 m long, no larger than 2.5 m, with a displacement no more than 1000 kg.

This kind of yacht is part of a developing medium-range commercial market, located on the border between two families of very popular products: the small sailing dinghies, and small sailing cruisers. The research on this tipology of yachts, has showed that the reasons for their low spread depends not only on low demand, otherwise by the low quality of the offering. Most of the yachts in production are not satisfactory, and unable to meet users expectations; they show on the contrary strong deficiencies, both from performance, functional, ergonomics and aesthetic; such as difficult maneuverability, difficult trailerability, high displacement, lack formal identity, no livable comfort.

The yacht on the market can be considered only an aggregation of different morphological elements: traditional monohulls, equipped with aerodynamic and fluid-dynamic appendages. Sometimes these small crafts are equipped as cabin cruisers; the cabin are generally arranged in the bow of the yacht, integrated into the hull and accessible by little hatches.

Mostly there is possibility to have small portholes, but thy are not sufficient to ensure a correct air recirculation; the minimum equipments on board are bad arranged; housing minimum conditions insufficient to meet basic needs; the different internal and external environment modalities of fruition does not allow a correct relationship between users involved in different activities; the volumes are closed, very dark and narrow, characterized by a dis-comfort that affect a good market response. The initial survey of a number of users of this type of boats has shown that they use their small yachts only for short day trips; all the data collected show another attitude of these sailors: all of them only consider this micro habitable volume like a storage for all the tool kit useful in different situation on board and out board.

The analysis of these products, enabled to focus the research intent to formulate a response to some primary needs highlighted by the analytical work: low-displacement, trailerability, semi-craft construction processes, overall rethinking of the deckhouse living system. The material's survey identified the advanced plastic composites as the best compromise to ensure a good relationship between performance, serial production, and manufacturing costs (Fig. 1).

4 Concept Guidelines Development

Boat camping theme was investigated with great attention paid to the external and internal livability of the yacht, preferring continuous surfaces, flowing lines, a large cockpit, and a solution that would integrate the small deckhouse with the whole of the indoor system. The cat solution for the rigging, positioned at the extreme bow, ensure easy management with a single rope. This solution has allowed to free up the entire space of the deck and to take advantage of the forward part by using the entire volume, necessary to meet the minimum habitability requirements (Fig. 2).

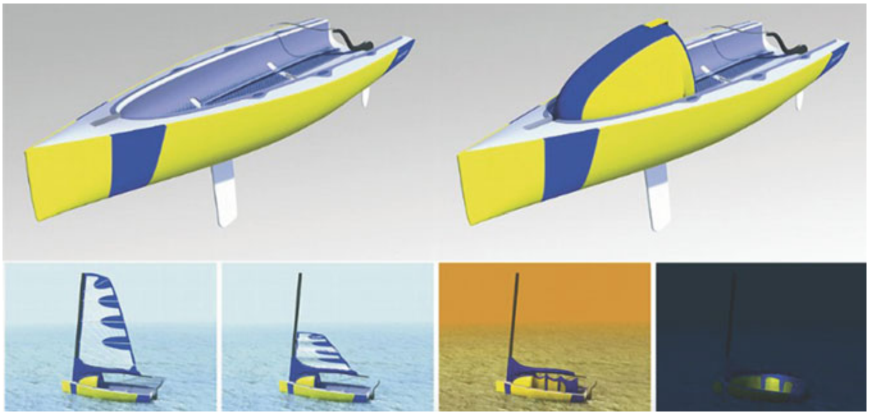


Fig. 1 Isisi photorealistic rendering

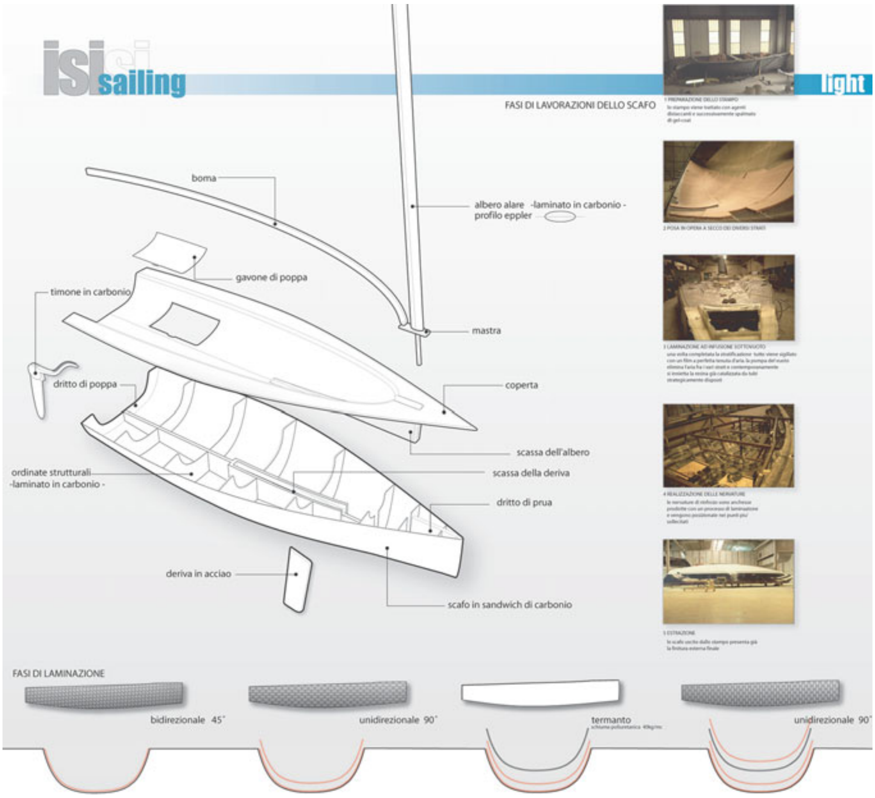


Fig. 2 Construction sketch

4.1 Deck Layout Organization

The main layout has been rational arranged, relating all the expected functions; maneuvering the mainsail and rudder, the outboard motor storage, keel management, and finally, the accessibility to the supple situated at the extreme bow.

The surfaces have been developed to allow the crew to take different postures because of the different activities on board, related to the balance conditions created by sailing, and at the same time to ensure an adequate level of accessibility on board by the adoption of the open stern solution.

The choice of the coating material for the deck was evaluated by the investigation on the needs expressed by a number of potential users of sailing dinghies, who were involved in the design process; the study was conducted through the testing of old drifts used as a testing ground for the application on their decks surfaces of different grip design solutions (Figs. 3 and 4).

Marfran was the final material choiced. Its a biomedical material derived elastomer produced by the italian group Franceschetti Elastomers; this material is produced in different hardnesses and already used in the production of diving equipment (Fig. 5).

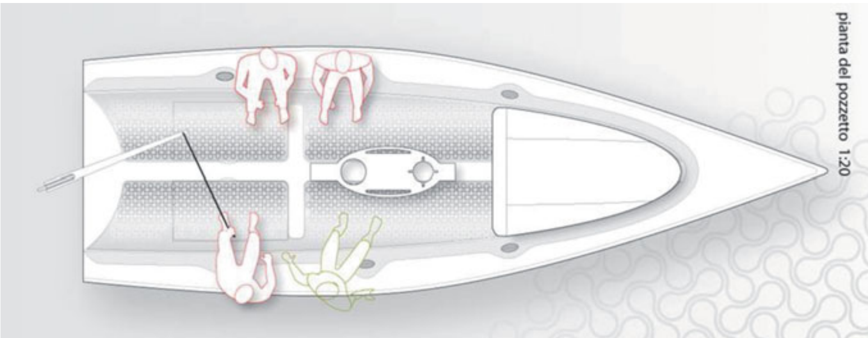


Fig. 3 Deck layout organization

Fig. 4 Analysis of different crew postures

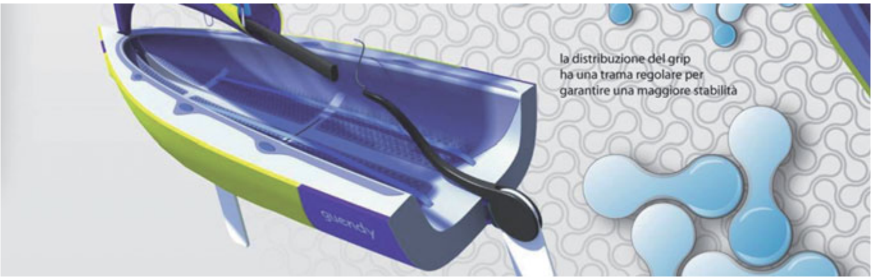


Fig. 5 Anti-skid grip morphological concept

This is an ideal material for covering the interior and exterior livable surface areas, making it non-slip through the application of a grip whose design has been appropriately tested through tactile pleasantness tests conducted using scientific methods (SEQUAM Sensorial Quality Assessment Method) [5], which led the definition of soft spots capable of making the outer surfaces pleasing to the touch and comfortable to use.

The closed material, elastic, pigmented, soft touch, deformable, water-resistant an repellent, is designed as a combination of repeated and cross curved elements to form a regular pattern.

4.2 Supple Design Solution

The supple arranged in the bow of the yacht has been designed as a removable and transportable unit, from the cockpit height, to the shore (Fig. 6).

Its light, compact and ovoid shape with ellipsoidal cross and longitudinal sections; this solution guarantees it's perfect housing in the bow of the cockpit, by the joint into the box of the little tire, positioned on the base of the supple. This solution allows to easily move the supple out of board. A memory foam of silicon gel allows and guarantee the accessibility to its interior; this solution contribute to to keep in safe from the external atmospheric agents and to stow all the essential tools, needed for the survival of the users, in the various and different operating conditions of the boat. A retractable roof is located into the edge of the supple; when the roof its open, it cover the entire deck of the small boat, up to the stern; the roof ensures it's form and stability through the adoption of three ellipsoidal pneumatic rings. The flexible cover is designed to ensure an appropriate level of visibility to the outside, and its equipped with membranes that can be opened to ensure microclimate comfort. If necessary, the membranes can be opened and closed through the adoption of normal silicone hinges, resistant to sea salt and to corrosion.

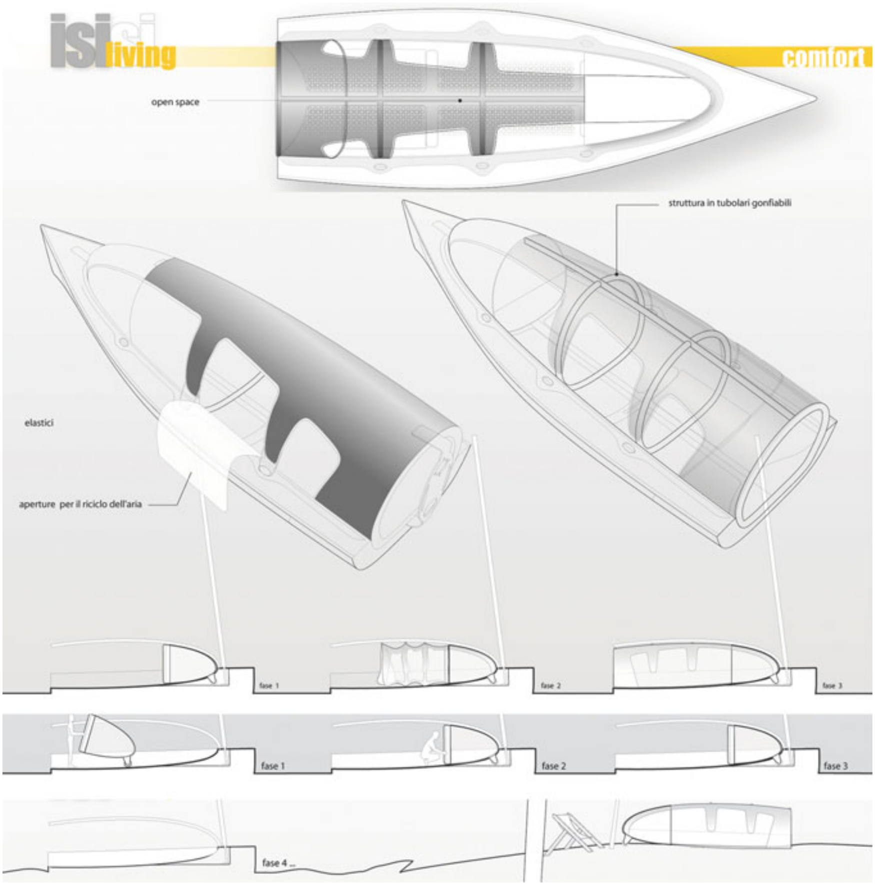


Fig. 6 Supple concept sketch

4.3 Supple Interior Layout

The supple is arranged in a very rational way, simple and functional (Fig. 7).

Inside were designed different solutions related with the principal need of the users; one a removable work surface, which contains the typical camp kitchen and a small sink; a portable sit integrated with a chemical toilet; a small sink; bags made of elastic net are provided, located inside of the shell and integrated in the lining material. The material chosen concur to generate a feeling of tactile pleasure, and allows an adequate level of adaptability to the soft, curvilinear characterizing linguistic and morphological level of the proposed nautical concept. The possibility of pigment enables to choose different color levels to identify different functions; this solution contributes to improving the general affordances.

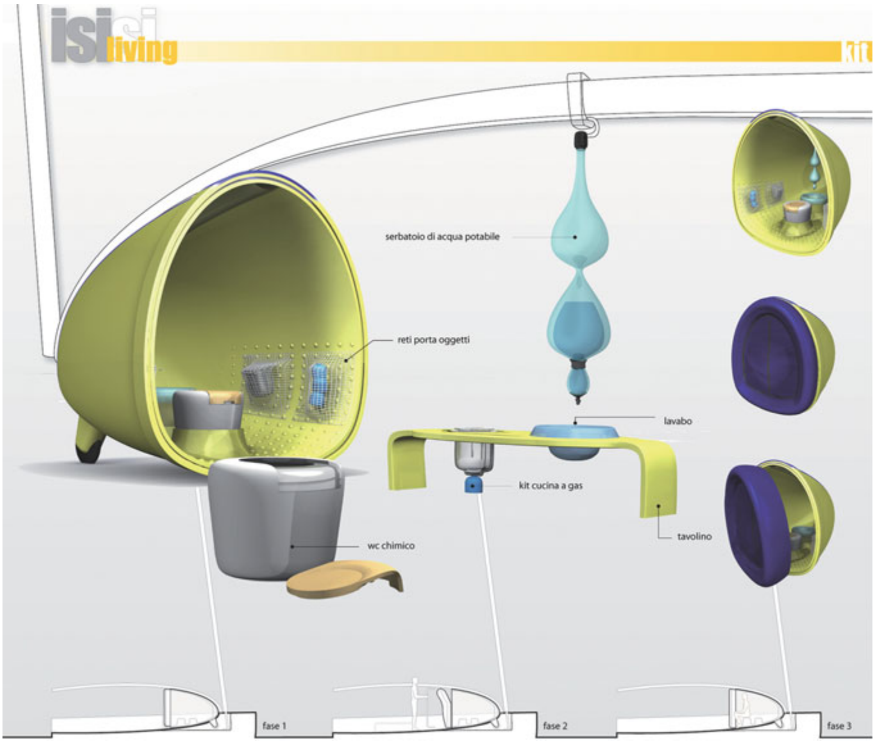


Fig. 7 Supple kit visualization

4.4 *Sail Plan*

Thanks to the adoption of the cat solution for the mast, the final users needs only to conduct a big mainsail only manouvering one mainsheet [6] (Fig. 8).

This option simplifies the maneuvers on board, and contribute to free the entire boat's deck plan, consequently to arrange all the functions in an optimal way, and at the same time to maintain a good sail area. The airfoil mast is made of composite material; light, rigid, easily removable, its the best solution, from the technical and functional point of view. The battened mainsail is hooded on the mast and integrate the boom to guarantee a good aerodynamic shape at sea; its made of monofilm and mylar materials.

The mainsheet allows the helmsman to easily conduct the boat; this simple trick allows the rest of the crew to be able to navigate in comfort and without having to intervene in the maneuvers.

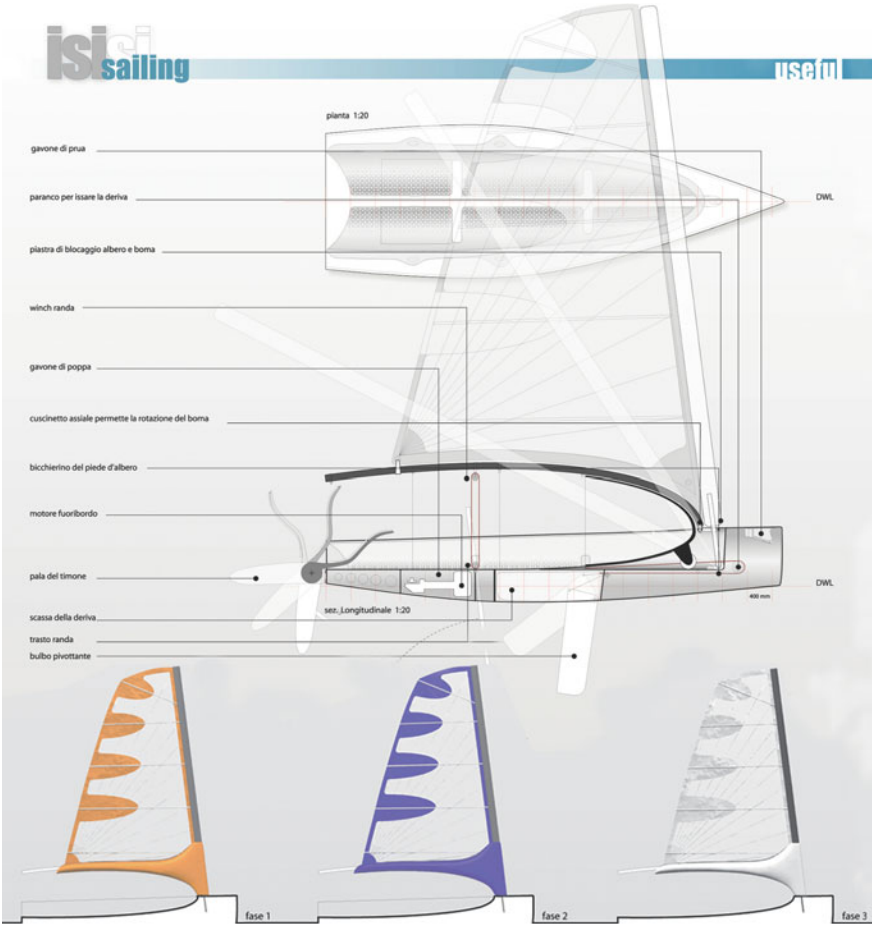


Fig. 8 Sail plan functional scheme

4.5 Trailerability

The final design solution adopted concern the possibility to transport the small boat by road, in compliance with the current legislation described by the road's code.

All design solutions proposed, folding mast, transportable supply, movable appendages, minimum dimensions, together with the fundamental requirement of lightness, allow the user to reconfigure the boat and to be able to hoist with a few movements of a road's carriage, and to be able to carry it in other places (Fig. 9).

This last design solution guarantees the end user to be able to benefit of the small craft, equipped for the accommodation of this particular category of tourists, in many different marine contexts.

Fig. 9 Trailerability concept sketch

5 Conclusion

The research already carried out, (described in this paper), has had its's main objective to bring out as a series of correct design solutions, relating to this particular type of boat, aimed by encouraging an idea of nautical tourism accessible and careful to the environmental resource.

The first purpose was to show how a series of correct design solutions could represent ideas of strong morphological and functional originality, and may constitute certainly interesting material for industry professionals; in terms of spatial organization, aesthetic and morphological solutions, and correct materials' choice.

The second purpose was to highlight how conscious design solutions could help to define new possible scenarios and usage forms of the marine environments.

Finally, try to stimulate a reflection on the idea (of sociological interest) that yachting (its practice), nowadays, is an activity potentially for a wide section of the population, more than its usually believed.

Acknowledgments This paper refers to the results achieved within a Master's degree Thesis in yacht design entitled "*ISISI*" (advisor: Prof. A. Vallicelli,; candidate: D. Maddaluno), edited in the "*Interior Design of Sustainable Living*" Degree Laboratory, at the Department of Architecture of the University of Chieti-Pescara (Italy). All the images reported in this paper are taken from the above mentioned M.Sc. Thesis. The development of the research referred to the technological solutions for a new concept of small camping sailing yacht, conducted directly from the M.Sc. candidate D. Maddaluno under the guide of Prof. A. Vallicelli. The various paragraphs of the present paper can be considered the consequence of a common discussion and a collective review among authors. In particular, the writing of the various paragraphs can be attributed to: Massimo Di Nicolantonio (Abstract, Paragraph 4 and 5), Jessica Lagatta (Paragraph 3), Antonio Marano (Paragraph 1) Andrea Vallicelli (Paragraph 2).

References

1. Benhamou, F.: L'economia della cultura, tr. it., Il Mulino, Bologna (2001)
2. Urry, John: Lo sguardo del turista. Il tempo libero e il viaggio nelle società contemporanee, Seam, Roma (1995)

3. Vezzoli, C.: System design for sustainability, Maggioli editore (2007)
4. Marano, A.: Design e ambiente. La valorizzazione del territorio tra storia umana e natura, POLI. design, Milano (2004)
5. 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. Boca Raton (FL): CRC Press, Taylor & Francis Group (2011)
6. Rossi, E., Di Bucchianico, G., Di Nicolantonio, M.: 10 Meters Daysailer "for All". Sustainable Technological Solutions for Easy Navigation. In: Peter V. (ed.) Advances in Social and Organizational Factors. Boca Raton (FL, USA), CRC Press (Taylor & Francis Group) (2012)

Design for Inclusion in the Field of Sailing Yachts. Design for All Approach for Small Sailing Boats

Jessica Lagatta, Andrea Vallicelli, Massimo Di Nicolantonio and Antonio Marano

Abstract In recent years, a strong attention to the issues of human diversity, social inclusion and equality has developed in the field of yachting. This awareness has led to the creation of sailing boats much more closely-related to Design for Inclusion theories, Universal Design (UD) and Design for All (DfA) in particular. In this paper, we present a final result of a Ph.D. research. The purpose of this study was to define the design guidelines for small sailing boats, based on the principles of DfA. To obtain this result, a comprehensive study of the state of the art was conducted, through the analysis of fifty small sailing boats. This study was carried out in two distinct phases. During the first one, all small sailing boats were filed and the characteristics of all case studies were observed. During the second one, comparative analyses between different boats were conducted. Starting from this analysis, we have drawn up design guidelines based on the principles of DfA.

Keywords Design for inclusion · Sailing yacht · Small sailing boats · Design for all

1 Introduction

Contemporary society is witnessing a strong social change, both in terms of population's ageing and of multiculturalism. Nowadays, sure enough, «human diversity in age, culture and ability is greater than ever. We now survive illness and injury and live with disability as never before» [1]. In recent years, the field of design has tried to meet the needs and aspirations of an increasingly complex population through the creation of new methods, such as *Design for Inclusion*.

J. Lagatta (✉) · A. Vallicelli · M. Di Nicolantonio · A. Marano
Department of Architecture, University of Chieti-Pescara, Viale Pindaro 42, 65127 Pescara, Italy
e-mail: jessicalagatta79@gmail.com

M. Di Nicolantonio
e-mail: m.dinicolantonio@unich.it

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_89

It assumes that the creation of products, services and environments for *standard person* leads to the penalization or the exclusion of most of the population from their use. The moment has come to realize that badly-designed environments, products or services can cause disability, not the opposite. In other words, «good design enables, bad design disables» [1]. This consideration has led to a radical change in design approach towards users. The specificities of each person are no longer conceived as ability or disability but as differences, and diversity begins to be seen as a resource rather than as a problem. Everyone has equal dignity, rights and opportunities to actively participate in the daily life, without any barrier or obstacle. Therefore, the purpose of *Design for Inclusion*, as the name itself says, is not to exclude any individual from the use of a product, an environment or a service. Starting from the awareness that human diversity is a resource for contemporary social and economic development, different approaches and design methods based on inclusion have been developed during the last thirty years. Among them, the most important are: *UD_Universal Design*, *ID_Inclusive Design* and *DfA_Design for All.* Although these theories are newly formed, they have already been applied in many field, such as the Sailing Yacht Design.

The strong expansion of the nautical sector that occurred in the last century—and the consequent increase of users—has been one of the factors that led to the application of Design for Inclusion in this area too. In fact, «today, only in Europe, 48 million people practice one or more nautical activities, and 36 million of these people go by boat (motor or sail). […] Nautical activities, despite being associated by media—often unjustly—exclusively to luxury, are not uniquely addressed to a social élite. The expression *mass recreational boating* can be rightly used» [2].

The increasing awareness about the themes of social inclusion in the different sports disciplines has been another important factor in the development of such methods in the nautical sector. This has also allowed people with disabilities to access to activities they had never practiced before. Sailing sports have not been excluded from this phenomenon. The nautical sector more closely involved in this development is that of small sailing boat (with a length less than ten meters). In the late 70s, in fact, the early boats designed to meet the specific needs of disabled sailors begin to be made. Moreover, the realization of a wide range of special aids to navigation (ergonomic seats, electric winches, joystick, sip and puff systems, etc.), allowed us to adapt existing boats to the different needs of users. From the 70s to today, however, considerable progress has been made. It has been realized that designing sailing boats not only for people with disabilities but also for users with different age, tastes and cultural habits, could represent a competitive solution for the nautical market both economically and socially. During the last years, such an awareness, has led to the creation of sailing boats that are more and more close to the theories of *Design for Inclusion.*

2 Purpose of Research

There is a market made up of nonstandard consumers, partially different from the ones whom the design of small sailing boats was addressed to only a few decades ago. This awareness leads to the need to propose new design solutions that have an active confrontation with the complexity and the added value they entail. DfA, defined by the Stockholm Declaration as the «design for human diversity, social inclusion and equality» [1], might be the most appropriate response to the demand the market is posing right now, which is the creation of more and more inclusive and not ghettoizing sailing boats that would also be able to satisfy the needs of every type of sailor. In financial terms, this approach could offer companies a number of new competitive advantages (increase of potential users, customer loyalty, etc.).

Moreover, as far as it concerns the social level and the collective well-being, basing the market development on the concept of diversity favors the inclusion as well as everyone's active and autonomous participation, thus increasing the wealth and the overall level of well-being.

All of this having been said, the research here described aims at understanding how the DfA approach can be applied to the small sailing boat design. However, in order to address this study, a complete view of the status of this issue is firstly required. In the following paragraphs we will explain how this theme has been dealt with in a Ph.D. Research developed in the Department of Architecture of Pescara (Italy).

3 Method

The study has been developed following three different phases. In the first phase, different cases study have been selected and cataloged. In the second phase, an analysis of all the boats taken into account has been conducted. In the third phase, design guidelines—based on DfA methodology—have been defined.

3.1 Case Studies' Selection and Cataloguing

The small sailing boats considered as case studies are 51. They had been selected according to target users and design methodologies used for their realization. We can summarize the selection criteria in three points:

1. Small sailing boats designed and created exclusively for people with physical disabilities.
2. Small sailing boats designed and created for able-bodied sailors but that, thanks to their particular features (technical ones, related to stability, etc.) and to the

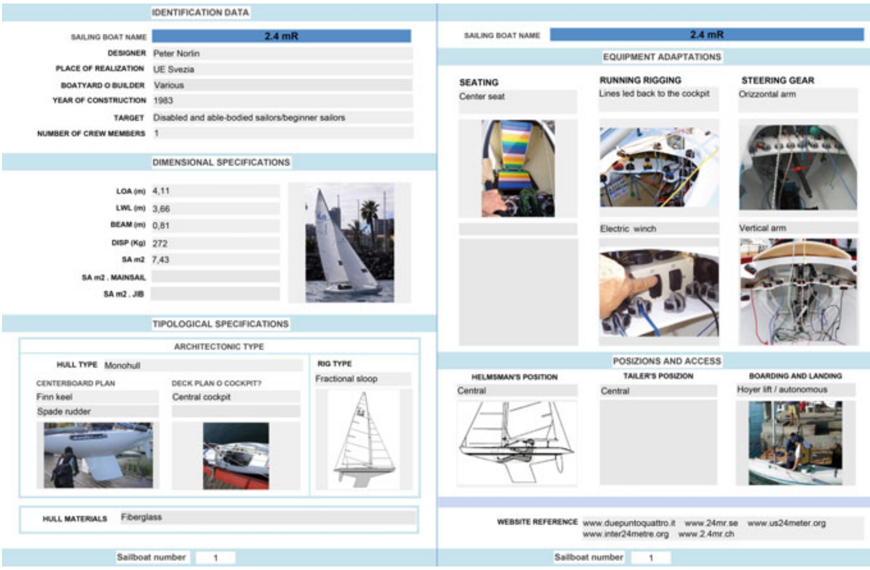


Fig. 1 Example of records cataloguing, created using *FileMaker Pro 12*®. The sheet has been generated by entering the data of a case study: the 2.4 mR one

addition of “special aids”, may be steered by people with physical disabilities, too.

3. Small sailing boats overtly designed and created according to some methodologies of *Design for Inclusion*.

The following data have been collected for each of them: identification data, dimensional specification, typological specification, equipment adaptation, access and positions.

In order to organize the whole material, a specific program has been selected and used: *FileMaker Pro 12*®. Indeed, this software has made it possible to manage a substantial amount of information. Moreover, it has flexible tools which allow the records’ format customization according to the wanted type of data collection. Two records have been created for each small sailing boat.

As Fig. 1 shows, in the first one there are the identifying, dimensional and architectural data. In the second one, there is a cataloguing of the different adaptive changes (additional seats, electrical aids, specific types of steering gear, etc.), the access and positions for sailors who are sailing.

3.2 Case Studies Analysis

FileMaker Pro 12® is a relational database, thus capable of linking multiple records together. Thanks to this special feature, studying the data entered through comparative mode has been possible, as well as obtaining a report of the main technical, typological, dimensional and morphological data of all the boats taken into account. The extensive work of collection, systematization and case studies analysis has shown that many small sailing boats have common peculiarities and technical solutions.

The most used type of hull is the monohull, followed by the trimaran. The most common appendages are the ballast keel (which allows a greater stability) and the underhung rudder, steered by a tiller or a manual joystick. The most common length overall (LOA) is four or five meters. Therefore, the average number of crew members varies from one to three. Helmsman and tailer are often in a central position, this one being fixed in relation with the cockpit thanks to the addition of a rigid seat which is the most common adaptive element. Fractional sloop is the type of sail plan adopted in almost all case studies. In many small sailing boats, running rigging led back to the cockpit. To board is often autonomous and direct. In case of physical disability, instead, it is carried out with the aid of a mechanical or electric hoist.

These data have been obtained by considering the highest percentages of entries in records. Therefore, they represent the main features of a "standard case study".

4 Design for All Approach for Small Sailing Boat

Starting from the observation of the case studies, it has been possible to identify the technological, morphological and ergonomics features of the analyzed boats according to the target audience (able-bodied or disabled sailors, beginners, experts, young people, old people, etc.). Thanks to this analysis, knowledge has been acquired of the different design solutions adopted to meet the needs of various types of users. Starting from this knowledge, some design guidelines based on the principles of DfA (for example on the methodology) have been drawn. Before describing them, however, a clarification is needed: the small sailing boats may have very different dimensional and typological characteristics. According to them, the relationship between the sailor and the elements that compose the boat can considerably vary. For this reason, the small sailing boats have been classified according to their typology (monohull or multihull) and to their size: 1- LOA from 3 to 4 m (typically steered by a single sailor); 2- LOA from 4 to 6 m (usually steered by two or three sailors); 3- LOA from 6 to 10 m (normally steered by two or more sailors). Even if in the Ph.D. research has described the features of all the above-mentioned categories, the decision has been made to report the results concerning the monohull small sailing boats with length from 4 to 6 m. Such a choice allows to synthetically—thus more clearly—explain the application of DfA guidelines for a “standard case study”.

In order to obtain this result, we have particularly based on the instructions given by DfA Italy that in 2010 established two brands: *DfA Start* and *DfA Quality*. Their purpose is «to certify the Design for All’s quality of products, environments and systems, that is, of all the project realizations of material and virtual reality, such as industrial products, architectures, services, graphics and multimedia» [3].

More generally it can be said that, in order to be defined as DfA, a product—in this case a small sailing boat—has to respect some principles, which are shown below: to give value to human diversity; to promote social inclusion and equality; to be easy and pleasant to use for all users; not to ghettoize, either physically or psychologically; to be beautiful; to be socially, environmentally and economically sustainable; to have as its final goal the quality of life’s improvement.

Furthermore, it has to meet very specific requirements. It is appropriate that they are tested in regard to target users, that is in regard to all «people who wish to enjoy or have a reasonable likelihood of independently enjoying the product, environment, or system» [4]. The requirements’ four main categories are: physicality, feeling and sense, comprehension, process. Each of them includes a number of other requirements that have to be tested for each element of the small sailing boat and that will be subsequently presented. A small sailing boat «can be considered an *organized work system* defined by an interrelated set of elements (activities, riggings and people), working in special environmental conditions, pursuing the main object of *sailing* with efficiency, effectiveness and with satisfaction for everybody» [5].

Given that the DfA methodology places target users at the centre of its attention, the application of the guidelines was carried out taking into account some parts of the small sailing boat, which are closely related to the activities and operation carried out by the sailors. Therefore, the requirements will be described in this order: seats, steering system, running rigging and sail plan.

4.1 Seats

The design of a seat is a very complex one. Several factors have to be considered. First of all, the position the sailor takes when sailing. It varies depending on the role he plays (helmsman, tailer, bowman) and on the boat’s size and morphological features. It can generally be said that sailors at sea must frequently change side to counterbalance the heeling moment generated by the wind force on the sails. Therefore, the seats are often located in a lateral position with respect to the cockpit. This recurring movement from one side to the other could be tiring for all those people who, for example, have limited mobility or have recently started practicing sailing sports. In some case studies, “mobile” seats (which rotate on a central pin or move along the tracks) or even central or lateral auxiliary seats have been used to solve this problem. However, they constitute a design response to the specific needs of "limit users", who can be defined as: «that part of All (target users) that has the most critical specificity to independently enjoy the system» [4]. In DfA methodology, instead, these problems «are solved by taking into account the autonomous users, with the awareness that the other users will satisfy their specific needs with ad hoc aids, which they normally already use» [4]. For example, a person with lower limbs disability should enter and sit inside the boat as autonomously as possible and by using his/her own wheelchair. However, the cockpit of the small sailing boats here analyzed has not the space needed to contain it. Therefore, the adopted solutions will have to consider this factor, too. Starting from these premises, the DfA requirements for seat will be described.

4.1.1 Physicality

Minimum Required Efforts and Abilities. During a tacking or a gybing, the sailor may need to shift from one small sailing boat's side to the other. This movement is performed while the boat is moving and requires a particular strength and motor skills, as well as a good sense of balance. If the small sailing boat has only lateral seats, the sailor needs to make a triple action: to get up from one side, to shift and to sit on the opposite side. As already mentioned, not all users may be able to perform these actions with skill and little effort. A good design solution should instead allow the body fluid and not fatiguing movements. A single central seat which connects the two sides (for example like that used in NEO 495), could be a good design solution for all, not only for those who have mobility difficulties. In fact, it allows the body to move without the need to stand up and sit down. In this way, the sailor can move with minimum efforts and skills.

Enabling Prehensility. The seat should provide for some support elements (handles, openings on the supporting surface or on the backrest, etc.) in order to enable everyone to cling when side is changed or the body contrasts the small sailing boat's heeling. Furthermore, a non-slip material should be used both on the seat and the seatback (if there is one), in order to guarantee a greater body's adherence to the surface.

Handling and Use Alternatives Offer. The seat should be handled and used in alternative ways. In other words, it could be used not only for sitting, but also for other functions. For example, as a container element for ropes, fenders or other objects, or as a mobile element that serves as a seat at sea, but that could be used as an access walkway after docking.

Correspondence to Anthropometric Differences. In order to respond to this requirement, the seat should be adjustable according to the size of the user's body. It is therefore necessary to create a design solution that allows to rise the height and make the size of the seat and of the backrest (if there is one) flexible.

Personalization Systems Offer. The personalization systems may relate to the addition of cushions, the possibility or not to insert a backrest, the presence of safety belts, the possible use of a headrest or armrests, etc. In the analyzed case studies, a seat with only one of the systems listed above has been used, depending on the type of ability or disability of users. On the contrary, the analyzed requirement demands the systems to be included in a single artifact as much as possible, or at least added to it, depending on the sailor's demanding.

Safety Features Compliance. Safety concerns various aspects. For example, the possible presence of water on the seat's surface could cause an unintentional slipping of the body. For this reason, non-slip and quickly dry out materials should always be used. In addition, the seat's corners should always be rounded in order to avoid the risk of bruises.

4.1.2 Perception and Sense

Chromatic Contrasts. The eye better visualizes the seat's dimensions if its color is different from that of the cockpit, thus avoiding feet or body bruises. In addition, highlighting parts that have greater ease of grip (as handles or armrests, if there are) by chromatic contrasts might be helpful during the maneuvers. Always thinking on safety, a color differentiation should be used between the rigid part (structure) and a potential soft part (cushions). The same applies to non-slip areas.

4.1.3 Comprehension

Self-Explanation of Use/Operation. If, as previously described, the seat has modular or additional parts, their use and operation should be immediately understandable to everyone.

Formal/Aesthetic Quality. In the analyzed case studies, seats are often inserted into the small sailing boats regardless of their formal/aesthetic component. This approach contributes to create that "hospital image" that is typical of some products dedicated to people with disabilities. Therefore, the formal/aesthetic quality becomes relevant because it helps to branch off from the idea of an artifact designed exclusively for "limit users". Furthermore, the seat should be designed with the boat's same formal language. It would make it a congruent and integrated element with the rest of the small sailing boat, helping to eliminate that feeling of "additional" and "medical" element perceived in some case studies.

4.2 Running Rigging and Sail Plan

A simultaneous description of the requirements for running rigging and sail plan has been chosen here. Indeed ropes, halyards and sheets are needed to govern the sails. Therefore these items cannot be studied separately, because they are closely connected with each other. In addition, running rigging will be taken into account in relation to the hardware (winches, clam cleats, stoppers, etc.) for a more complete analysis. The analysis of the case studies has shown that the fractional rig sloop is the most widespread rigging. In many small sailing boats there is also the bowsprit and the use of spinnaker pole is expected. Thus the guidelines will include the gennaker and spinnaker, too.

4.2.1 Physicality

Minimum Required Efforts and Abilities. Hoisting the mainsail and the jib while rigging may require efforts that are not suitable for all sailors' motor abilities. Some strength's decrease systems, such as winches and travelers, are typically used to facilitate this operation. While sailing, the tailer must maneuver the sails and also move from one side to the other during a tacking or a gybing. Therefore, the whole body is under stress. The seats (which it has already been spoken of) play an important role in decreasing leg fatigue. Instead, travelers and winches are used to minimize the force required to haul and ease the ropes (fatigue of torso, arms and hands). In order to minimize the tailer's and bowman's efforts, limited sweeping gestures or movements are necessary. Therefore it is desirable that all the maneuvers are put back to the cockpit, in order to have the ropes always close at hand and to prevent the bowman from needing to move towards the bow. In order to obtain this result, some solutions already available on the market can be used for each type of sail (jib, mainsail, spinnaker, spinnaker).

- The jib: it should be equipped with a self-tacking/turning system and furling jib. This would enable the sailor to maneuver and to reduce the sail size directly from the cockpit.
- The mainsail: a furling mainsail system could be applied on the boom or mast. This would prevent from manually reefing the sail, allowing to reduce the sail area by the use of a rope.
- The gennaker: in order to use this sail with minimum effort, it is desirable that it is governed by a cockpit. In Skud 18, for example, a closed-circuit rope allows to hoist the gennaker and simultaneously release the bowsprit. It also allows to drop down the sail. This rope is led back to the cockpit and so the user can steer the gennaker while sitting and with minimum fatigue.
- The spinnaker: the use of this sail requires a particular physical effort and a good motor ability of the bowman. He must: attach the spinnaker pole to the mast, fit out it and hoist the spi. During the gybing he must: unhook the spinnaker pole from the mast and fix it to the other spinnaker sheet; unhook the spinnaker pole from the after guy and fix it to the mast. To accomplish these tasks, the bowman has to stand and balance. Moreover, it is necessary that his arms and hands have strength enough to support and handle spinnaker pole and sail.

The use of a maneuvering system different from the just described one would therefore be desirable to respond to the required *minimum abilities and efforts*. A good solution would be the *twin pole spinnaker system* that allows the bowman to hoist, steer and drop down the spinnaker by the use of lines led back to the cockpit and without the need to handle the spinnaker pole.

Enabling Prehensility. Ropes, halyards and sheets must be grasped by everyone with ease. Therefore, they have not to be slippery. In some cases, sheets with different treatments and sections could be accomplished to improve the prehensility; in others, adding the spherical elements of antislip material might be useful. They would be inserted in the terminal part of the rope in case it reaches the stopper or the cam cleat.

Correspondence to Anthropometric Differences. As previously mentioned, the ropes of the headsails should all be led back to the cockpit. They are normally blocked by a cam cleat located on the deck. Cam cleats should be positioned on a console that can scroll towards the bow or stern on rails in order to meet the user's anthropometric differences in a better way. This system (already used in Skud 18) allows to regulate the distance from the console to the tailer, better responding to the body's different dimensions.

Safety Features Compliance. Injuries that may result from the use of ropes, sheets and halyards are varied: the flow can cause abrasions and burns to the hands; the rigging can be an obstacle for the feet and cause danger of falling, etc. Regarding abrasion, it can be avoided only by using protective gloves. To prevent the rigging from being an obstacle, instead, they may be placed under deck (protected by a carter), where it is possible. Regarding sail plan, instead, the boom height should be adjusted according to that of the sailors to avoid head's bruises while tacking or gybing (thus respecting also the requirement: *correspondence to anthropometric differences*). It could be also lined with a soft material that, in case of impact, would help to soften the blow.

4.2.2 Perception and Sense

Chromatic Contrasts. In headsails, chromatic contrasts may be useful to distinguish the different parts of the sail. A mainsail's colors differentiation may be helpful to indicate to what height it should be hoisted, depending on the wind. The sails chromatic contrasts may generally facilitate the boat recognition from a distance. The color differentiation of the ropes, instead, can help to understand which sail they refer to.

4.2.3 Comprehension

Formal/Aesthetic Quality. Sails are an important part in the visual impact of a small sailing boat. Taking care of their aesthetics through an image study coordinated with the hull would help to make the small sailing boat more attractive.

Multisensory Communication. When using the ropes, the mainly involved senses are: sight, touch and hearing. However, not all users are able to use them simultaneously. For example, blind sailors may have difficulty in identifying the position of the ropes and in understanding which sail they refer to. In this case the previously mention color diversity is not useful. This lack may be filled by differentiating diameters and weft. These solutions may be useful for all sailors and not just for a "limit user".

4.3 The Steering System

In the analyzed case studies, various types of steering gear have been used: tiller, steering wheel, manual or servo assist joystick, foot pedal steering, sip&puff system. Each of them is applied to the small sailing boat, according to the user's skill level. The tiller and the manual joystick are the most common ones, probably because they adequately respond to the needs of the largest number of sailors. The second type, in particular, proved to be the most appropriate to respond to the requirements of DfA. The design guidelines are shown starting from this basis.

4.3.1 Physicality

Minimum Required Efforts and Abilities. The functioning of the two types of the rudder above is very different. In fact, the small sailing boat sails in the same direction the manual joystick is conducted in. This principle is instead reversed in the use of the tiller: if it moves to starboard, the boat will go to the left and vice versa. It is evident how the use of the manual joystick, especially for inexperienced sailors, is much simpler and intuitive, thus requiring a minimum ability. Furthermore, given that this it is always positioned at the bow of the helmsman, it is less difficult to maintain the course when a change of side is necessary.

Enabling Prehensility. The helmsman should always monitor the small sailing boat and to do so, he has to grab the governing body in the most firm way. The prehensility, in this case, is a very important requirement. The manual joystick, besides including personalization systems (see. next entry), should be coated with a non-slip material in order to be easily grasped.

Correspondence to Anthropometric Differences and Personalization System Offer. The manual joystick should be adjustable through a telescopic system according to the users' anthropometric differences. Personalized additional elements (that should be placed on the head of joystick) might make the grip more comfortable even for those with motor difficulties.

Safety Features Compliance. The rudder control is essential to ensure safety. Operating errors, in fact, could lead the sailing boat to capsize, to collide with other boats or possible obstacles in the sea, etc. A good prehensility and ease of use, given from the previously described solutions, help maintaining the boat's safety.

4.3.2 Perception and Sense

Chromatic and Tactile Contrasts. They may be useful to highlight the parts of the manual joystick which must be gripped by the user.

4.3.3 Comprehension

Self-Explanation of Use-Operation. As previously mentioned, the fact that the direction of the small sailing boat coincides with that of the manual joystick ensures that its use is expressed by itself, making it easier to learn navigation techniques.

4.4 The Design Process

The "process" entry has deliberately been omitted from the above mentioned requirements. In fact, it has been decided to treat it separately, since it is common to all the elements of analyzed small sailing boats. «In DfA the process is holistic and inclusive, and it allows (by creating, managing and implementing complex systems) everyone's comfortable and pleasant fruition» [4]. The main features that must be observed in the design process are: the participation of users, the multidisciplinary, the involvement of decision makers. The users, or their representatives, should be taken into account not only in all stages of the process (proposal, meta-design, design), but also in the product's ergonomic testing. This allows an assessment of human diversity and an effective comparison with it. The resulting complexity implies a multidisciplinary approach, i.e. a synergistic participation of specialists of other areas (e.g. Psychologists, ergonomists, engineers, etc.). The DfA process also includes the involvement of decision-makers, which is of those who «decide to make a product in one way rather than another, with partners and collaborators rather than others» [4]. When designing a small sailing boat there can be various kinds of decision makers. In the analyzed case studies there are: single individuals (ship owners), public entities (e.g. The Navy League), international associations that promote *sailing for all* (e.g. Sailability World), small local non-profit associations (for e.g. EOS), etc.

The present research has come to an initial proposal (design brief) based on the observation of examined users and small sailing boats. It just represents the first step of the design process required by DfA. Therefore, the requirements described in the previous paragraphs constitute a specific aid for the development of design, which will have to be developed by involving decision makers and users. Moreover, the creation of prototypes or mock-up will have to be an aid to perform tests on the user, before arriving at the final product.

In addition to all has been said above, design guidelines specified in the rules of DfA Italia brands stipulate that the process does not end when the small sailing boat has been realized. It is rather necessary that there is «an extension of DfA principles to the entire value chain» and that «publi-promotional messages promote the DfA philosophy together with the product/system/environment, avoiding contrary viewpoint, especially if they ghettoize» [3].

In other words, all the aspects related to maintenance and disposal of the product, to logistics, to marketing, to advertising promotion, etc. should be analyzed in the process, too. It is clear that such a process is a very long and complex one, but the deriving benefits represent advantages for all the involved ones. First of all, the product's easy and pleasant fruition leads to an improvement in users' quality of life. The single-investment response to the different needs different projects should satisfy leads to an increase in profits and in the number of customers for businesses. The creation of new opportunities for contact with potential users and of a distinctive image in innovation and ethical-social commitment areas improves the work of designers. Finally, the inclusion of excluded groups of users provokes an image relapse in project successes funded or commissioned for public commissions [6].

5 Results and Conclusion

Different results have been achieved through this research. Thanks to the cataloguing of the case studies, a first report on the current state of the art has been possible. Moreover, small sailing boats that meet the characteristics described in accordance with the selection criteria of Sect. 3.1 are increasing over time. The database which has been used and the structure of the records will allow to continue the cataloguing of the material in the future. Thanks to the flexibility of the program, in fact, records can be added and changed, if necessary. This will allow a continuous update not only of the small sailing boats' number, but also of their technological, dimensional and architectural characteristics. Moreover, the analysis of the collected data could also be implemented and it will also help to understand how the characteristics of boats belonging to the analyzed typology can change over time. Finally, the indications explained by DfA requirements may form a basis for the design and the realization of more and more "inclusive" small sailing boats.

Acknowledgments This paper refers to the results achieved in the last phase of a Ph.D. research: "Design for Inclusion in the field of sailing yacht. Design for All approach in small sailing boat". This research is going to be developed at the Department of Architecture of G.d'Annunzio University of Chieti-Pescara (Italy) (Ph.D. student Jessica Lagatta; Tutor Prof. A. Vallicelli; co-tutors Prof. G. Di Bucchianico and Prof. M. Di Nicolantonio). The various paragraphs of the present paper can be considered the outcome of a common discussion and a collective review among authors.

References

1. The EIDD Stockholm Declaration: http://dfaeurope.eu/what-is-dfa/dfa-documents/the-eidd-stockholm-declaration-2004/
2. Pollicardo, L.: CNA Produzione-Nautica. Dinamiche e prospettive di mercato della filiera nautica del diporto. Second edition (2013)
3. Design for All Italia.: I marchi DfA, http://www.dfaitalia.it/marchi
4. Accolla, A.: Design for All: il progetto per l'individuo reale. Franco Angeli, Milano (2009)
5. Vallicelli, A., Di Bucchianico, G.: Evaluation of tasks and postures of a sailing yacht tailer. In: Ergonomics is a lifestyle. Proceeding of: 40th International Annual Congress of the Nordic Ergonomics Society (2008)
6. Vantaggi - Marchio di qualità: http://quality.dfaitalia.it/design-for-all-italia/vantaggi/

Part XV
Road and Rail—Safety, Driver Psychophysiology and Eye Tracking

Measuring Trade-Offs Between Risk and Travel Time Based on Experimental Speed Data

Paolo Intini, Pasquale Colonna, Nicola Berloco and Vittorio Ranieri

Abstract Users make continuous trade-offs between risk and mobility (travel time) while driving, in order to maximize the utility of their travel, which depends on different variables and it varies according to different individuals. One of the measurable outputs of this process is the speed. An on-road experiment was carried out in order to measure the speed behavior of a sample of 19 drivers on a two-lane low-volume highway. Measurements were repeated over six days in order to test if the familiarity acquired with the road can be responsible for speed changes over time. By making some hypotheses based on previous theories, the reduction of travel times due to the acquired road familiarity can be related to a conscious/subconscious shift to a more significant risk-taking behavior. These trade-offs will be analyzed throughout the paper by considering the differences between drivers and the geometric differences between different road sections.

Keywords Risk · Travel time · Road safety · Driving behavior · Speed · On-Road experiment

P. Intini (✉) · P. Colonna · N. Berloco · V. Ranieri
Department of Civil, Environmental, Building Engineering and Chemistry,
Polytechnic University of Bari, Via Orabona, 4, 70125 Bari, Italy
e-mail: paolo.intini@poliba.it

P. Colonna
e-mail: pasquale.colonna@poliba.it

N. Berloco
e-mail: nicola.berloco@poliba.it

V. Ranieri
e-mail: vittorio.ranieri@poliba.it

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_90

1 Introduction

Every time a driver starts its travel on a given road, he accepts, more or less consciously, a certain risk associated to that travel. Generally speaking, the risk related to each activity can be defined as:

$$R = p \times I. \quad (1)$$

Where:

P Probability of the unwanted event;

I Intensity of the consequences of the event itself.

Therefore, the application of the risk definition to the driving process implies that it is obtained by multiplying the probability of occurrence of the traffic accident, by its consequences. Obviously, the traffic accident is an unwanted event. Apart from the rare and extreme case of a person who is deliberately searching for the occurrence of an accident, it happens because an unexpected event occurs while driving and the user is not able to react promptly or accordingly.

Nevertheless, there are several factors (related to drivers, vehicles, road, environment and traffic) which can affect this process by influencing both the probability of the accident and its consequences. Among all the possible factors, as confirmed by early and recent studies [1, 2] the most important is the driver factor. This means that the driver behavior plays a huge role in the occurrence of the accidents and it depends on turn on several variables (e.g. age, experience, exposure, health conditions, attitude to risky driving, driving under influence, use of protection devices, trip purpose).

Furthermore, in normal conditions, the driver tries to maximize the utility of its travel by considering these two antithetical concepts:

- Time is money, so the driver wants to save as much money as possible;
- Life is holy, so the driver wants to preserve his safety as much as possible.

The antithesis is given by the fact that, for the aim of maximizing the money saving, the driver tends to decrease the travel time and so, for example, he can select faster speeds or he can tend to modify lateral position by cutting curves; but these behaviors are evidently in contrast with the safety purposes. Therefore, each driver, while traveling, rules this dynamic equilibrium by modifying his behavior accordingly: each modification can be defined as a “trade-off” between mobility and safety [3].

Anyway, the matter is complex: a part of this process can be conscious [4], but there is evidence that drivers can misperceive risk and travel time [5]. Furthermore, the dynamic equilibrium conducting to trade-offs is absolutely personal: each driver-related variable defined above can influence this process.

Considering a middle-aged driver, with several years of driving license, experiencing a remarkable annual mileage, in good health conditions, not driving under influence of drugs, wearing the seat belts and on considering also that he is driving a new and well-maintained car in good weather conditions and free flow traffic conditions, the trade-offs between mobility and safety will probably depend on his attitude to risky driving, the road geometry and the purpose of the travel.

In this paper, trade-offs between mobility (in terms of travel time) and safety are measured on the basis of experimental speed data, considering that some variables influencing the driver behavior can be fixed according to the experimental approach used. The influence of road geometry and the attitude to risky driving will be considered as the variables influencing the occurrence of trade-offs.

2 Methods

The methods employed are based on a theoretical framework (Sect. 2.1), on an on-road experiment (Sect. 2.2), on the application of the framework to the experimental data (Sect. 2.3) and on the calculation of risk and travel time (Sect. 2.4).

2.1 Theoretical Framework

This study is based on the framework proposed by Noland [3] (as an extension of the model proposed by O'Neill [6]) in which the utility of the travel is defined as follows:

$$\mathrm{U} = \mathrm{f}(\mathrm{P}, \mathrm{T}, \mathrm{C}, \mathrm{A}, \mathrm{R}). \quad (2)$$

Where:

U Utility of the travel;

P Price;

T Travel Time;

C Capability;

A In-Vehicle Activities (such as those that lead to distraction);

R Risk.

According to this equation, the utility of the travel depends on the mobility-related variables (price of the travel, travel time) and safety related-variables (driver capabilities, distraction or other activities leading to distraction while driving, risk associated to the travel). Drivers tend to maximize the utility of the travel (that is minimizing its cost) by acting on the independent variables of the Eq. 2. An example of the way in which this process takes place is the speed selection. If the driver increases his speed he can perceive a benefit for reducing travel time, but he increases the risk related to the travel (even if this increase in risk could not be perceived by everyone). Anyway, it should be considered that an increase in speed results in a greater price due to more fuel consumption, that greater capabilities are needed, and that probably the distraction tendency is reduced while going at higher speeds. All these trade-offs are made by the drivers in order to maximize the perceived utility of the travel.

2.2 On-Road Experiment

In this study, speed data belonging to a previously realized on-road experiment will be used. The experiment was planned for the aim of analyzing speed changes over time due to the acquired familiarity of the drivers with the test road [7]. In fact, a sample of 19 drivers drove on a very low-volume two-lane rural road, on which they had never traveled before, for six times in six different days.

The sample of drivers was chosen according to the following criteria:

- They had never traveled before on the chosen test route;
- Age-homogeneity (mean age of the final sample was 24.45 ± 1.10 years old);
- Minimum three years of driving license (mean: 5.75 ± 1.25 years);
- Availability of their own car;
- Minimum mileage of 10 km/week on rural roads.

In order to meet the age-homogeneity requirement, drivers were recruited with advertisements among the students of the Polytechnic University of Bari.

Two stretches of two-lane two-way rural roads (SP31 and SP18, situated in the district of Bari, Italy, speed limits namely 50 km/h and 70 km/h) were chosen as driving test routes. The complete driving test (see Fig. 1) consisted on traveling

Fig. 1 Scheme of the test road layout (on the *left*) with identification of the cross-sections in which speed was measured (on the *right*)

along a route composed of the above mentioned two stretches of road in the following order:

- Stretch 1, from the Start to the intersection with stretch 2—way there;
- Stretch 2, from the intersection with stretch 1 to the End—way there;
- Stretch 2, from the end point to the intersection with stretch 1—way back;
- Stretch 1, from the intersection with stretch 2 to the Start again—way back.

Drivers were asked to drive freely on this route without any other instruction, while their speed was continuously measured by using a differential GPS technology (fixed antenna, on-board vehicle rover antenna, data recorder). A cost compensation for fuel consumption was considered for all drivers.

For the purpose of the study, users were asked to repeat the same driving test described above six times in six different days. The chronological schedule of the test repetitions over time were fixed for each driver. The first four tests were scheduled in four consecutive days (1st, 2nd, 3rd and 4th days of testing). The other two tests were fixed in the ninth day after the first test (5th day of testing) and in the twenty-sixth day after the first test (6th day of testing), as also the possible presence of a long term memory effect after some interruptions in administering stimuli was investigated. Anyway, since the matter of trade-offs between safety and mobility could be perplexed by the consideration of these other two days, more distant in time than the others, only the tests related to the first four consecutive days were taken into account.

Speed data were assigned to the 137 road cross-sections (61 in the stretch 1 and 76 in the stretch 2) highlighted in Fig. 1 by connecting the value of distance corresponding to each cross-section to the respective value of speed in the speed profile. Speed data corresponding to situations of adverse weather or adverse traffic conditions, based on experience reported by the drivers, were discharged from the dataset.

In order to take into account the different road geometry along the test route, the 274 cross-sections (in both outward and return journeys) were clustered into four classes in respect to their computed value of sight distance. The latter was chosen as a synthetic variable representing road geometry, as long as it takes into account both horizontal and vertical alignments. Four visibility classes were so defined:

- class 1: 53 cross-sections with low sight distance (0–100 m);
- class 2: 110 cross-sections with medium-low sight distance (100–200 m);
- class 3: 38 cross-sections with medium sight distance (200–400 m);
- class 4: 73 cross-sections with high sight distance (400–600 m).

Furthermore, in order to allow the interpretation of speed measurements in terms of different road user behaviors, drivers were divided into three clusters according to their speeding inclination based on measured data. The three drivers' cluster are:

- cluster A: 6 drivers traveling at a speed lower than the sample average;
- cluster B: 8 drivers traveling at a speed near the sample average;
- cluster C: 5 drivers traveling at a speed higher than the sample average.

Users' clustering was conducted by using the K-means algorithm, a nonhierarchical cluster analysis.

2.3 Application of the Theoretical Framework to the Experiment Conducted

The analysis of speed data belonging to the on-road experiment conducted is modeled on considering the theoretical framework presented in the Sect. 2.1.

In particular, given the conditions of the experiment, the following assumptions can be made for the factors in the Eq. 2:

- The price P of the travel was not considered by the drivers as long as fuel compensation was expected;
- The capability of the drivers C can be roughly considered as a constant since the characteristics of homogeneity explained in the Sect. 2.2 were considered for selecting the sample of drivers;
- The in-vehicle activities A could be considered as not present, as long as the drivers were most likely attentive while driving since they know it was an on-road experiment;
- The utility U of the travel can be set to zero, since for all the explanations given above, drivers had no need to minimize the costs related to the travel as long they were volunteers undergoing to an on-road experiment.

All these hypotheses are clearly approximations of the reality, but they were felt as reasonable by the authors for the aims of this study.

Therefore, in this particular case, the Eq. 2 can be rewritten as follows:

$$f(T, R) = 0. \tag{3}$$

The equation can be plotted in a Cartesian plane, where the travel time (a variable representing mobility) is on the x-axis and the risk (a variable representing safety) is on the y-axis. This depiction, represented in Fig. 2, is an adaptation of the mobility-safety plan proposed by Noland [3] and Dulisse [8].

In particular, a negative power function was chosen to relate risk to travel time. Therefore, the Eq. 3 can also be written as follows:

$$R = k \times T^{-b} \tag{4}$$

The Eq. 4 was chosen since this function satisfies the following conditions. On a given road section characterized by its length L, with all other conditions being equal, a travel time close to zero (which can be related to a speed approaching infinity and to more aberrant behaviors such as cutting curves or running a red light) can be related to a very high risk in terms of accident and vice versa. Furthermore,

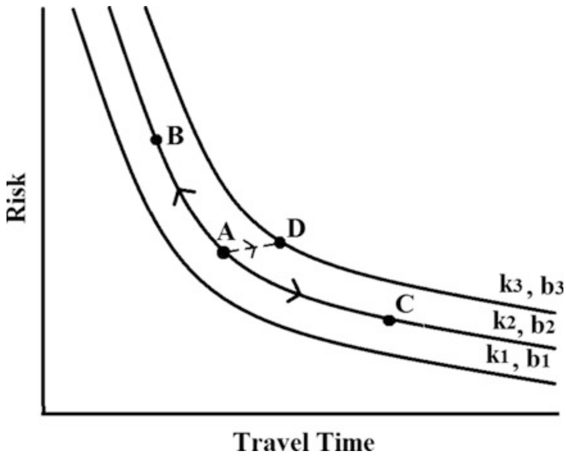


Fig. 2 Representation of the trade-offs between travel time and risk

this formulation allows a wide variability due to the elasticity permitted by the two parameters k and b.

The values of the parameters k and b define the curve representing a particular road system. The curve represents a "drivers' preference curve" and each point of the curve can represent the risk and the travel time for a particular driver, or the mean of a similar group of drivers, traveling on a given road section. Therefore, on considering the point A, it could represent the choice of the driver in terms of level of risk and travel time for the road section characterized by the parameters k_2 and b_2. A modification in the road system (e.g.: the introduction of traffic speed control measures) or the shift of the travel on another road section characterized by a different geometry could result in a different level of risk and travel time. Therefore, in the Fig. 2, the new state of the driver could be represented by the point D on a different branch characterized by the parameters k_3 and b_3.

Instead, let us consider the case of an unfamiliar user driving for the first time on a given road section (point A). After some repetitions of the same travel he will probably modify his behavior in terms of speed and lateral position (see [7, 9, 10]), and in terms of reduced attention capacity due the minor mental workload [11]. Thus, even staying on the same branch (k_2, b_2), behavioral changes will result in a movement to a different state characterized by a different levels of risk and travel time (point B or point C). This shift is clearly depending on personal variables as long as different individuals can show different tendencies in modifying their behavior.

Therefore, the speed measures acquired from the on-road experiment will be used in order to verify the hypothesis that the trade-offs between risk and travel time happen following the patterns described in this section. In the next section, the methods for converting measured speed into measures of risk and travel time are explained.

2.4 Measures of Risk and Travel Time

The process of speed choice was defined by Elvik [5] as not objectively rational and, in the same way, the concepts of risk and travel time can be not exactly perceived by the drivers. Therefore, Noland [3] argues that the identification of preference curves (as the one shown in Fig. 2) and trade-offs is difficult because they are based on the subjective interpretation by the users of various factors involved in the determination of the utility of the travel. However, since this study is not really focused on studying the subjective optimization of the utility related to a travel, but on the measurement of trade-offs between risk and travel time based on measured data, objective measures of both travel time and risk will be taken into account. Furthermore, both measures were converted into monetary costs in order to be comparable.

Risk. The risk taken by the users while driving on the test route was assessed by considering the relationships between speed and the accident risk found in literature. A review of previous literature studies concerning this topic (see e.g. [12]) revealed that it is very difficult to quantify the absolute accident risk related to a given speed of an individual driver. There are some studies relating measured individual speeds to the accidents happened in the previous 3–5 years (e.g. [13, 14]). Instead, in other studies, the relationships between average traffic speed and relative accident risk (referred to a given level of risk) is taken into account (e.g. [15, 16]). For the aims of this study, focused on the trade-offs between travel time and risk over four days of testing, it is not possible to consider a measure of risk based on accidents happened in the past. Therefore, it is only possible to measure risk in terms of number of accidents related to a given level of reference. The equations used are based on the study made by Elvik [17] in his work about a re-parameterisation of the earlier power model [15].

$$\text{Relative number of accidents} = \alpha \times \exp(\beta \times \text{initial speed}). \tag{5}$$

Where:

Relative number of accidents	number of accidents referred to the value of 100 for the maximum speed of 115 km/h considered in the study;
Initial speed	average traffic speed;
α	0.072 (fatal accidents), 1.983 (injury accidents), 2.928 (PDO accidents);
β	0.069 (fatal accidents), 0.034 (injury accidents), 0.032 (PDO accidents)

However, it should be noted that the initial speed is referred to the average traffic speed and not to individual free flow speeds. Anyway, since the equation gives a relative risk and not an absolute estimate of the accident frequency, then it can be considered only as a qualitative measure of how speed can be related to accident risk, coherently with the aims of this study.

Hence, the relative number of accidents (referred to the value of reference of 100 accidents for a speed of 115 km/h) was converted into a relative general cost for the accidents (referred to the value of reference of the cost of 100 accidents). Since fatal, injury and PDO (property damage only) accidents were estimated separately, then their costs were summed in order to obtain a unique value of accident cost for each speed, and after referred to the value of reference corresponding to 100 accidents, by using the following equation.

$$\text{Index of relative cost of accident risk} = \frac{\sum \text{costs for accidents at a given speed.}}{\sum \text{costs for 100 accidents.}} \quad (6)$$

In which the sums refer to the sum of the cost of the accidents related to three levels of severity: fatal, injury and PDO. The average general costs of each accident (considering also societal and health costs for fatal and injury accidents) was set to: 1,642,236 € for fatal accidents, 309,863 € for injury accidents and 10,986 € for PDO accidents [18]. The computed index varies between 0 and 1.

Travel Time. Travel time was computed by easily converting speed measures obtained from the on-road experiment into time measures, because the length of the investigated road section was known (about 7 km on considering both the outward and the return journeys). In particular, since four visibility classes were considered in order to take into account road geometry variability (see Sect. 2.2), then the overall road section was dived into four sub-sections: 1.3 km characterized by low visibility, 2.7 km characterized by medium-low visibility, 1 km characterized by medium visibility and 1.8 km characterized by high visibility. Therefore, travel times were computed for each driver and for each road sub-section.

Travel times were converted into monetary costs by considering a constant conversion factor for all drivers, since they were all young students earning no wages. A value of 3 €/h was used as a conversion factor based on a recent study [19], considering this category of drivers included in the sample.

Since the measure of risk was related to a value of reference corresponding to 100 accidents for a speed of 115 km/h, then also costs related to travel time were related to a value of reference of travel time corresponding to a speed of 115 km/h, in order to obtain two comparable measures. Then, the following index was defined.

$$\text{Index of relative cost of travel time} = 1 - [\text{TT(actual speed)}/\ \text{TT}(115\ \text{km/h})]. \quad (7)$$

The introduction of the unity as a minuend in the subtraction aims to obtain an index variable between 0 and 1 as the one employed for measuring risk. This is possible as long as no average speed is greater than 115 km/h in each road section inquired.

3 Results and Discussion

The main results showing the evolution over time of travel time and relative accident risk in the four days of testing based on measured speed data are shown in Figs. 3 and 4 by using risk—travel time diagrams (see Fig. 2). The travel time on the x-axis is represented by the computed index of relative cost of travel time (see Sect. 2.4) and the risk on the y-axis is represented by the computed index of relative cost of accident risk. Both the axes are variable between 0 and 1 as for definition of the indexes.

In Fig. 3, the diagram showing the estimated relationship between risk and travel time for the overall length of the experimental road section is drawn. It is as a regression curve based on the points (index of risk, index of travel time) obtained for each day of testing (1, 2, 3, 4) and each drivers' cluster (prudent, average and risky drivers) from experimental data of average speed. Since the overall test road was considered, a unique curve was defined, based on hypotheses made in the Sect. 2.3.

In Fig. 4, the diagrams showing the estimated relationships between risk and travel time for each sub-section based on visibility categorization are drawn. They are regressions based on the points (index of risk, index of travel time) obtained for each day of testing (1, 2, 3, 4) and each visibility class (high, medium, medium-low, low) from mean data of speed of all drivers. Since four sub-sections were

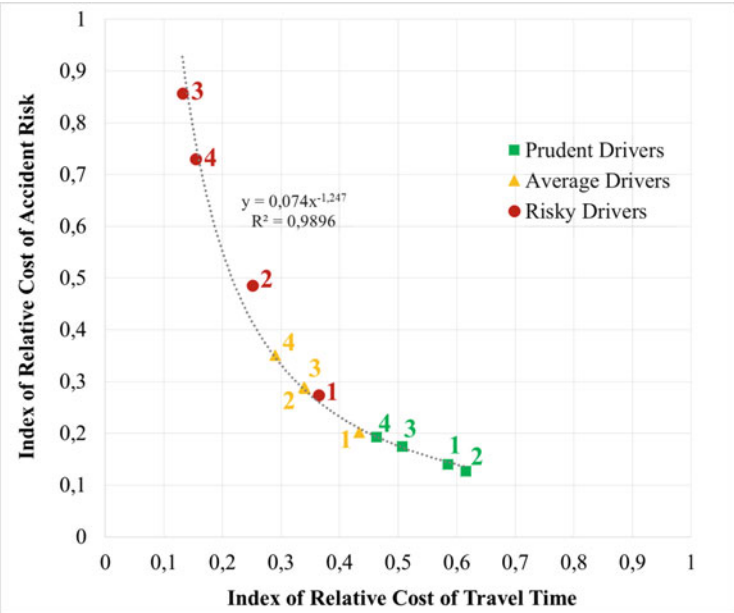


Fig. 3 Estimated risk-travel time curve for the overall road section, showing trade-offs over the days of testing (1, 2, 3, 4) for each drivers' cluster

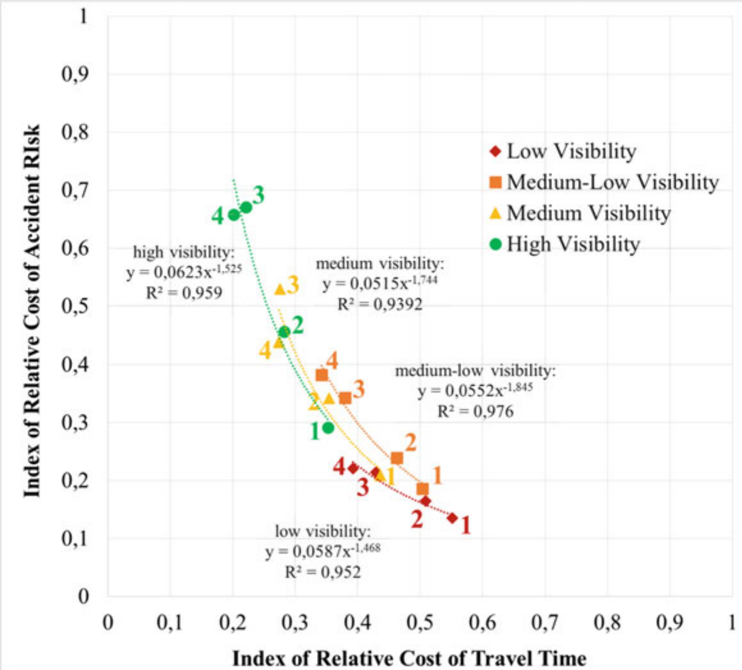


Fig. 4 Estimated risk-travel time curves for the road sub-sections according to sight distance, showing trade-offs over the days of testing (1, 2, 3, 4) for all drivers

considered, four different curves were defined, based on hypotheses made in the Sect. 2.3.

From the analysis of the two figures (Figs. 3 and 4), we can notice that the hypothesis of using a negative power function to link changes in travel time to changes in accident risk is verified. In particular, the values of the estimated parameters based on the regressions vary between 0.05 and 0.08 for the parameter k, and between 1 and 2 for the parameter b. Therefore, the empirical relationship between accident risk and travel time based on the definitions given in this study, is the following:

$$\text{Index of Relative Cost of Risk} = (0.05 \div 0.08) \times \frac{1}{\text{Index of Relative Cost of Time}^{(1 \div 2)}} \quad (8)$$

The definition of these curves (valid only for the particular section of two-lane rural road), which can be considered as the drivers' preference curves in terms of risk and travel time for the considered road sections, allows us to quantify trade-offs between risk and safety. In this study, trade-offs are assessed based on the behavioral changes due to acquiring familiarity with the test road over the days of testing.

It is evident from Fig. 3 that drivers make trade-offs between risk and travel time due to the increase in speed, while going from the first days (1, 2) to the third and fourth days of testing, and that this phenomenon is highly correlated to the risk inclination of the drivers themselves. In fact, a general tendency to move from the bottom-right to the up-left of the diagram by following the preference curve over days can be noted for all the drivers' clusters. However, while for prudent and average drivers, the decrement of travel time is associated to a modest increment of the relative risk, for the risky drivers, a similar decrement of the travel time is associated to a very high increment of the relative risk. Moreover, this discrepancy is more evident considering that the level of relative risk associated to the first day of testing is very similar between average and risky drivers, while the level of risk associated to the last days of testing is much higher for the risky drivers than for the average drivers.

Other considerations can be made by looking at Fig. 4. In this analysis, the focus is on all drivers, but in different visibility conditions. Firstly, it should be considered that even if sight distances are different, all sub-sections belong to the same road category: a very-low volume two-way two-lane rural road. Thus, as expected, there are no huge differences between curves belonging to different sub-sections. Anyway, different tendencies for trade-offs between risk and travel time can be identified for different visibility conditions. The decrement of travel time in respect to the first days of testing resulting in a shift to the left of the index of cost of travel time, is roughly very similar for each category of visibility. Instead, the modification of the index of accident risk goes from a very modest increase in the low visibility conditions (from about 0.15 to about 0.2), to a very robust increase in the high visibility conditions (from about 0.3 to about 0.65). However, in this case too, the state representing the first day of travel on the test road in terms of risk and travel time is not so different in the diverse visibility conditions (especially for lower visibility), while it is on the last day.

Trying to catch the key factors related to these two different analyses, the following concepts arise. The differences between drivers' clusters and visibility conditions are minimum in the first day and maximum in the last days of testing with respect to risk and travel time. Independently from the initial level of the index of travel time (that is independent from the initial speed, more or less high), its decrease over time (that is: increase of speed/decrease of travel time) is quite constant: it does not depend so much on the drivers' cluster or the visibility. Instead, the corresponding variation of the index of accident risk is very different considering different drivers' clusters and visibility conditions, because it follows a power function and the process starts from points representing different speeds. This could happen because drivers with different risk tendencies could have different ways of internalization of the external risk into their own perception, during the process of familiarization with the road [20].

This means that, independently from the more or less demanding condition given by the road, and independently from the attitude to speeding (reflected in the initial speed), there is a great harmony between drivers in considering the relative quantity of travel time which would decrease while becoming more familiar.

However, this circumstance allows this direct consequence: the risk taken is proportionally much higher for less demanding road conditions (high visibility) and for more risky drivers, since in those cases the initial speed is higher.

4 Conclusions

Trade-offs between travel time and risk were analyzed with respect to repeated speed measurements over days belonging to an on-road experiment. Acquiring familiarity with a road has been related to both an increase in risk and a decrease in travel time based on the objective measures defined in this study. More evident trade-offs were highlighted for risky drivers and in less demanding road conditions such as the high visibility road sub-section. A function that correlates relative risk to relative travel time with regard to their cost was defined for the particular condition of an on-road experiment; unless drivers' perception of those quantities could be totally different from what objectively determined. Moreover, the importance of considering different drivers' levels of risk while evaluating trade-offs due to familiarity was highlighted.

Findings from this study can be inserted in the debate about trade-offs between safety and mobility. As long as they can be measured in given conditions, the occurrence of trade-offs could be considered in the future in different situations:

- In the safety countermeasure design; since behavior of drivers can adapt to a modification in the road system by modifying e.g. the speed. This can be considered while assessing costs and benefits of the intervention itself [21].
- In the evaluation of a change in the road system; as long as the evaluation should start from the moment in which drivers already (eventually) completed the adaptation phase to the modification in the infrastructure.
- In the road design stage, since the road should be ideally designed in order to avoid behavioral changes towards riskier behaviors for familiar drivers.
- In on-road experimental design, since familiarity could have an impact on risk and travel time and so, some repeated measures should always be planned in order to avoid what can be called an "unfamiliarity bias".

References

1. Treat, J.R., Tumbas, N.S., McDonald, S.T., Shinar, D., Hume, R.D., Mayer, R.E., Stansifer, R.L., Castellan, N.J.: Tri-level study of the causes of traffic accidents: final report. Executive summary. (Report No. DOT HS-034-3-535). Washington, DC: National Highway Traffic Safety Administration (1979)
2. Singh, S.: Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey. (Traffic Safety Facts Crash, Stats. Report No. DOT HS 812 115). Washington, DC: National Highway Traffic Safety Administration (2015)

3. Noland, R.B.: From theory to practice in road safety policy: understanding risk versus mobility. Res. Transp. Econ. **43**(1), 71–84 (2013)
4. Yannis, G., Kanellopoulou, A., Aggeloussi, K., Tsamboulas, D.: Modelling driver choices towards accident risk reduction. Saf. Sci. **43**(3), 173–186 (2005)
5. Elvik, R.: A restatement of the case for speed limits. Transp. Policy **17**(3), 196–204 (2010)
6. O'Neill, B.: A decision-theory model of danger compensation. Accid. Anal. Prev. **9**(3), 157–165 (1977)
7. Colonna, P., Intini, P., Berloco, N., Ranieri, V.: The influence of memory on driving behavior: how route familiarity is related to speed choice. An on-road study. Saf. Sci. **82**, 456–468 (2016)
8. Dulisse, B.: Methodological issues in testing the hypothesis of risk compensation. Accid. Anal. Prev. **29**(3), 285–292 (1997)
9. Colonna, P., Intini, P., Berloco, N., Perruccio, A., Ranieri, V.: Repeated measurements of lateral position and speed at horizontal curves on very low-volume rural road. In: Transportation Research Board 95th Annual Meeting (No. 16-3426) (2016)
10. Martens, M.H., Fox, M.R.: Do familiarity and expectations change perception? Drivers' glances and response to changes. Transp. Res. Part F: Traffic Psychol. Behav. **10**(6), 476–492 (2007)
11. Young, M., Stanton, N.: Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. Hum. Factors **44**(3), 365–375 (2002)
12. Aarts, L., Van Schagen, I.: Driving speed and the risk of road crashes: a review. Accid. Anal. Prev. **38**(2), 215–224 (2006)
13. Fildes, B.N., Rumbold, G., Leening, A.: Speed behaviour and drivers' attitude to speeding. Monash University Accident Research Centre, Report, 16 (1991)
14. Maycock, G., Palmer, C., Buttress, S.: The factors that influence a driver's choice of speed: a questionnaire study. Transp. Res. Lab., Crowthorne, UK (1999)
15. Nilsson, G.: Traffic safety dimensions and the power model to describe the effect of speed on safety. Bulletin-Lunds Tekniska Högskola, Inst för Teknik och Samhälle, Lunds Universitet, 221 (2004)
16. Baruya, A.: Speed-accident relationships on European roads. In: 9th International Conference on Road Safety in Europe (1998)
17. Elvik, R.: A re-parameterisation of the Power Model of the relationship between the speed of traffic and the number of accidents and accident victims. Accid. Anal. Prev. **50**, 854–860 (2013)
18. Ministero delle Infrastrutture e dei Trasporti. Studio di valutazione dei costi sociali dell'incidentalità stradale. Italia (2010)
19. Fiorello, D., Pasti, G.: Il valore del tempo di viaggio. Guida teorica e applicativa. TRT Trasporti e Territorio Srl, Milano, Italia (2003)
20. Colonna, P., Berloco, N.: External and internal risk of the user in road safety and the necessity for a control process. In: 24th PIARC World Road Congress, Mexico City (2011)
21. Rudin-Brown, C., Jamson, S. (eds.).: Behavioural adaptation and road safety: theory, evidence and action. CRC Press (2013)

Crash Testing High-Power Portable Traction System Device

Núria Parera, Daniel Sánchez, Alba Fornells, Eloi Boix and René Molina

Abstract The Applus IDIADA Passive Safety laboratory has a high accuracy and reliability traction system for performing crash tests. However, the system cannot be used outside the laboratory due to its type of construction. The objective of this project is to design and implement a new, portable, high-power traction system, which enables crash tests to be performed at any location and with new configurations, such as heavy-vehicle crashes, street shows and improved car2car tests. The first phase of this project, which is explained in this paper, consist of the calculations to find a motor with necessary power, the choosing of an automatic transmission system, the design of the pulley installation, anchorages and tensor mechanism. Also, includes the choosing of a suitable steel wire for the traction. The second phase that is still in progress includes the motor's software development, all the modifications of the mechanical parts, the hardware implementation and the complete system installation. The last phase will consist of a validation test for the whole system and single components.

Keywords Crash testing · Traction system · Portability · New configuration impacts

1 Introduction

Nowadays, to ensure a high safety level to the vehicles, thousands of crash tests are performed to assess their crashworthiness. They are performed in crash tests laboratories where environment conditions are controlled and the results are analyzed.

IDIADA is an automotive engineering partner, expert in every stage of the vehicle development; from the first design of the vehicle to the last homologation procedures before the marketing of the product. One of the critical phases in a vehicle development is the definition of the passive safety systems. Passive Safety

N. Parera · D. Sánchez · A. Fornells · E. Boix (✉) · R. Molina
Applus IDIADA, L'Albornar PO Box 20, E-43710 Santa Oliva, Tarragona, Spain
e-mail: eboix@idiada.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_91

department is in charge to develop and improve vehicles equipment to minimize the consequences of traffic collisions. To assess the different restraint systems installed in the vehicle, several crash tests are carried out at IDIADA Passive Safety laboratory.

The vehicle is accelerated and guided through a straight track and it finally impacts against a barrier (rigid or deformable) or another vehicle. The traction system consists on a rope closed loop, in which the vehicle is attached, moved at the required speed by an electric motor. The system is completely accurate, reliable and solid making possible to reproduce several tests with exactly the same conditions, crucial factor in order to compare results.

However, the current traction system shows some limitations. Traction capacity only makes possible to test vehicles up to 7000 kg and not heavier vehicles as trucks. Moreover, it is fixed and testing in other different places is not possible. Some kind of passive safety tests are not able to be performed with the current traction system:

- *Misuse*: The Passive Safety system not only has to be deployed when the accident occurs, it must also not to be activated when it is not required. These scenarios are called misuses and several tests are carried out to assess if the vehicle identifies that different kinds of situations very close to the limit but not enough to active the restraint systems. Dynamic misuse test should be performed on flat or irregular (with bumps, holds, etc.) surfaces outside the laboratory.
- *Rollover*: It is also interesting to know vehicles behavior in scenarios such as rollover. This test is performed on a defined embankment ground with determined slope and length. Track for this test is placed outside the laboratory.
- *Heavy vehicles*: To perform trucks and buses crash tests a longer run-up trucks and a more power traction system is required. The current traction system is not able to move so big weighs neither to achieve the required speed in a reduced track length.
- *Road restraint systems*: For road restraint system tests, cars and heavy vehicles are crashed at very high speeds as described in the European regulation EN 1317-1,2. These tests should be performed outside the laboratory in order to achieve the heavy trucks speed

To solve this issue, a new portable traction system has been developed. It consists of a closed loop metallic rope guided by a pulley system and driven by a movable engine. The new system should be able to be installed anywhere, and heavy vehicles as buses or light trucks should be crashed at high impact speeds.

2 Objectives

The objective of this project is to design and implement a new, portable, high-power traction system which will offer a reliable mechanism to perform crash tests. This new device will allow performing crash tests at any location and new configurations that now, with the current system is not possible to be done, such as heavy-vehicle crashes, street shows and improved car2car tests.

In order to accomplish this main objective, smaller objectives had to be achieved such as follows:

Study of the truck requirements: in order to be able to perform crash tests with heavy vehicles, the portable system will need to fulfill requirements such as:

- Maximum weight of the test vehicle
- Minimum test vehicle speed
- Engine power
- Longitude of the test track
- It should be installed everywhere and easy to transport
- It has to be a robust, solid and reliable system

To accomplish those requirements, some calculations and actions are needed to be achieved:

- *Calculations* of the power of the traction engine, maximum traction force, maximum lineal velocity, design and calculations of the pulley system and calculations for the wire thickness and materials.
- *Design of the structure of the traction engine*. This part will include the study of the structural modifications and replacements that are needed to be done in order to reduce the current mass and convert the truck in a portable traction system. Although the structural modifications are aimed to make the portable system lighter, the system has to reproduce a crash test with heavy vehicles up to 40,000 kg of mass, so it should weigh enough to not to show any displacement during the process of an impact test.
- *Electronic improvements* are needed to be done like the development of the portable system software and all the new cable connections with the new structural relocations and improvements.

Finally, a validation test has to be done in order to validate the single components and the whole system.

3 Requirements

The new system should be installed anywhere and be able to perform tests with heavy vehicles outside a laboratory. Finally, a minimum speed should be ensured. Different elements of the system have been studied and dimensioned as it can be observed in the following lines:

3.1 General Testing Requirements of Mass and Velocity

The system should drive a maximum mass of 38,000 kg at a maximum speed of 65 km/h (required conditions to perform a test of road restraint systems under the EN 1317-1,2) or a maximum mass of 3000 kg at a maximum speed of 120 km/h (limit requirements of passenger car tests). The required engine power to perform these kind of tests has been calculated taking into account the point as follows:

3.2 Option 1: Truck of 38.000 kg at 65 km/h

To overcome the standstill static inertia forces caused by the rolling resistance forces in order to make the vehicle start its movement. The rolling resistance coefficient of a truck tires in an asphalt surface is $\mu = 0.03$ and the force is:

$$\text{Rolling Resistance} = F_r = m \cdot g \cdot \mu = 38{,}000 \cdot 9.81 \cdot 0.03 = 11.183\ \text{N} \tag{1}$$

To achieve the acceleration of 38,000 kg at 65 km/h in a maximum displacement of 1000 m.

$$\begin{aligned} a &= \frac{v^2}{2\Delta r} = \frac{\left(65 \times \frac{1000}{3600}\right)^2}{2 \times 100} = 1.63\,\frac{\text{m}}{\text{s}^2} \\ F_a &= m \cdot a = 38,000\,\text{kg} \times 1.63\,\frac{\text{m}}{\text{s}^2} = 61,940\,\text{N} \end{aligned} \tag{2}$$

That means a total force of 73.124 kN.

3.3 Option 2: Car of 3000 kg at a Maximum Speed of 120 km/h

To overcome the standstill static inertia forces caused by the rolling resistance forces in order to make the vehicle start its movement. The rolling resistance coefficient of a passenger car tires in an asphalt surface is $\mu = 0.3$ and the force is:

$$\mathrm{F_r} = \mathrm{N}\,\mu = \mathrm{m\,g}\,\mu = 3000 \cdot 9.81 \cdot 0.3 = 8829\,\mathrm{N} \tag{3}$$

To achieve the acceleration of 3000 kg until 120 km/h in a maximum displacement of 1000 m.

$$a = \frac{v^2}{2\Delta r} = \frac{\left(120 \times \frac{1000}{3600}\right)^2}{2 \times 100} = 5.55\,\mathrm{m/s^2}$$
$$F_a = m \cdot a = 3000 \cdot 5.55 = 16{,}650\,\mathrm{N} \tag{4}$$

That means a total force of 25.479 kN.

Option 1 was defined as the worst case. Therefore, the engine necessities have been calculated in the basis of these parameters.

3.4 *Traction Force and Engine Solution*

To fulfill the objectives described and the traction necessities it has been decided to adapt a tractor unit from a truck.

The advantages are as follows:

- Using a thermic engine makes possible to install the system anywhere even if there is no electrical connection.
- The rims of the traction axle could be used as the traction pulleys of the system.
- The traction unit structure could be used as the base of the entire Portable Traction System, packaging all the other required devices—such as pulleys or ropes—together during the system transfer.
- The differential of the traction-steering axle can be used to change the relation of the power-speed to one wheel. The traction-steering axle is com-posed by the two wheels and the differential. It allows each of the driving wheels to rotate at different speeds and if one wheel is not used the differential gives higher power to the other wheel with a relation 4.1.

However, some specific requirements regarding the tractor unit and thermic engine should be taken into account:

- The thermic engine should have automatic gear as to achieve a constant acceleration.
- The truck should be 4 × 2 in order to focus the traction force only in one axle. It will make the pulley system easier.
- The engine should provide the required 73.124 kN.

Finally, the chosen traction unit was an Iveco Cursor 4 × 2 450E33T. Its thermic engine characteristics can be observed as follows.

To assure that the selected engine is powered enough for this porpoise, some calculations such as the traction force and the maximum speed that can be achieved have been carried out.

Maximum Speed. When the engine gives 238.6 kW power, it rotates at a maximum of 2400 rpm (revolutions per minute). If the engine is connected with the 16th gear, which have a gear ratio of 0.85 and only one wheel is used, the differentials gives a relation on 4.1, the speed should be:

$$\begin{aligned} \text{Traction axel revolutions } &= \text{engine rpm/gear and differential ratio} \\ &= 2400/0.85 \times 4.1 = 688.67\,\text{rpm} = 72.12\ \text{rad/s} \end{aligned} \quad (5)$$

Then, if the wheel is used as a pulley, the speed traction transmitted to the rope should be:

$$\begin{aligned} \textit{Rope maximum speed} &= \textit{traction axcle rpm} \cdot \textit{rim radius} \\ &= 72.12\,\frac{\text{rad}}{\text{s}} \times 22.5'' \times 0.0254\,\text{m} = 41.21\,\text{m/s} \\ &= 148.4\,\text{km/h} \end{aligned} \quad (6)$$

As a consequence, the maximum speed that can be achieved by the engine is higher than the required for the tests, so this engine is enough powered.

Maximum Traction Force. To calculate the maximum traction force the engine rotates at 1300 rpm and the engine torque is 1200 Nm (according to the Fig. 1). The gear ratio range of the 1st gear is 15.39 and the differential ratio is 4.1 so the driving axle should transmit the calculated torque as follows:

$$\textit{Traction axle torque} = 1200 \times 15.39 \times 4.1 \cdot 0.85 = 64,361\,\text{Nm} \quad (7)$$

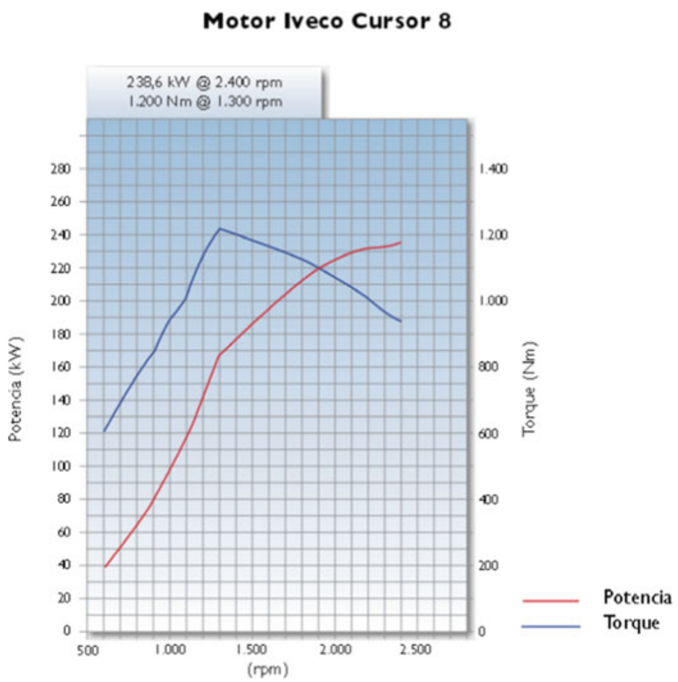


Fig. 1 Thermic engine characteristics

As the wheel (22.5 in) is used as a traction pulley, the lineal traction achieved is as follows:

$$\begin{aligned}\text{Lineal traction} &= \text{Traction axle torque/rim radius} = 64{,}361\,\text{Nm}/22.5 \times 0.0254\,\text{m} \\ &= 12.6\,\text{kN}\end{aligned} \tag{8}$$

The obtained value is higher than the required power calculated of 73.124 kN.

3.5 Rope Requirements

The requirements that the rope should fulfill to drive a heavy vehicle have been calculated below:

It has to support a maximum tensile stress of 73.124 kN plus a safety coefficient of the 50 %. Therefore, the total tensile stress to be withstood is 109.7 kN.

It has to ensure a constant tensile stress so the rope should be rigid and non-deformable. However, it should be flexible enough to be adaptable to the pulleys shape and to avoid the sliding. The most common cables to install are helicoidally shaped made of steel.

The rope selected is the AISI 316 7 × 19 + 0. The minimum diameter to fulfill the tensile stress requirements is 14 mm. This rope shows a breaking load of 112.77 N, higher than the maximum tensile stress. The final safety coefficient has been calculated by dividing the breaking load of the chosen rope by the calculated real maximum tensile stress.

$$safety\,coefficent = \frac{breaking\,load}{maximum\,tension} = \frac{112.77}{73.124} = 1.542 \tag{9}$$

3.6 Pulley System

Finally, the system to drive the vehicle to be tested has been designed. It consists of a closed loop rope system where the tested vehicle is pulled by a rope fitted in a pulley system that is driven by the engine of the truck. As it has been explained above, the tensile force and torque are transmitted to the rope by the tractor axle by means of the rims of the wheels. The scheme of the portable traction system is showed in the Fig. 2.

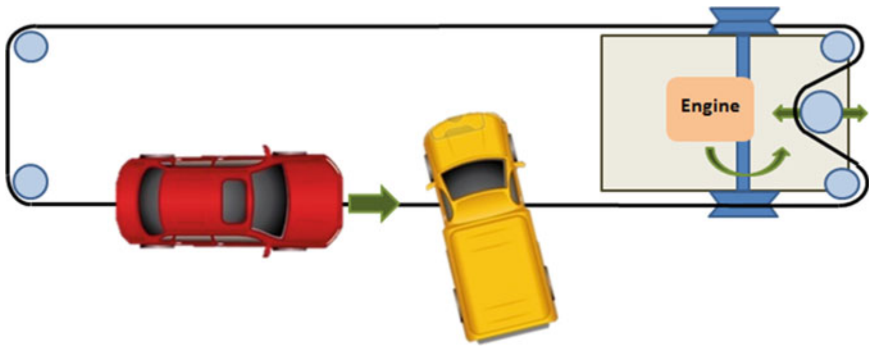


Fig. 2 Scheme of the pulley system and traction unit

4 Design of the Portable Traction System Device

The portable traction system consists of two main parts; the powertrain and the closed loop rope and pulley system. It fulfills the requirements specified above to carry out specific and heavy vehicles crash tests.

The truck's engine and its powertrain are the core of the portable traction system, but the huge volume of the chosen truck are too big. To make the system portable, the powertrain and the pulley had to be relocated on the same structure and the size of several elements had to be reduced. The chassis structure was modified as it can be seen in Fig. 3.

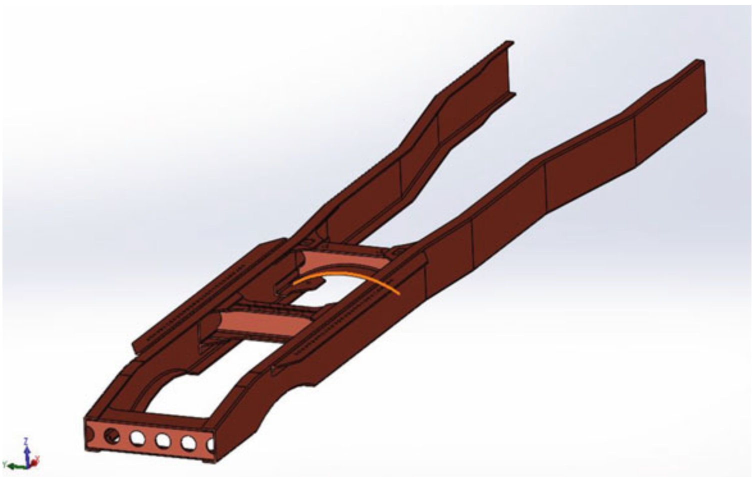


Fig. 3 New chassis design

Figure 4 shows the new chassis of the portable traction system with the current truck's gas tanks, gear box and cardan.

The new dimension of the portable traction system with the current truck elements are shown in Fig. 5.

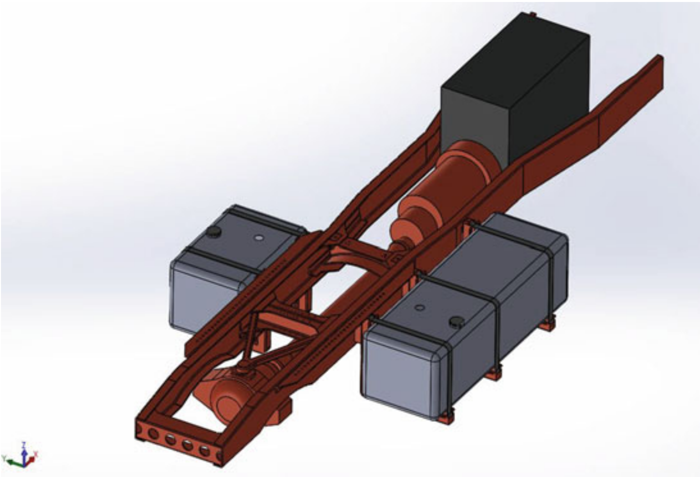


Fig. 4 New chasis desing with the actual truck elements

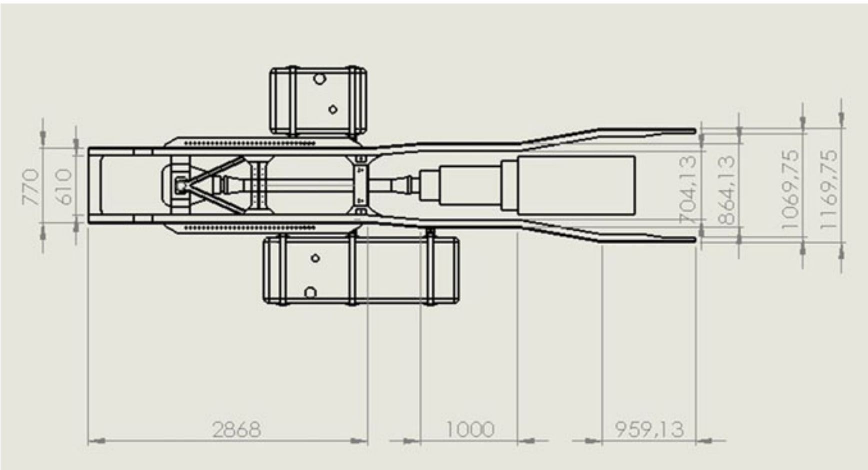


Fig. 5 Portable traction system dimensions

5 Structural Truck Modifications and Improvements

The base of the portable traction system is a truck of which structure has been modified. To achieve those new dimensions, the bodywork has been removed as much as possible as well as vehicle components such as seats, cabin, the two gas tanks were removed and replaced by one smaller gas tank. However, power train elements such as the engine, automatic gear, wheels and differential have been kept as well as electronic elements and vehicle ECU in order to control the engine functionalities.

Figure 6 shows the truck status once it has been dismantled.

The hydraulic mechanism cables needed to be spliced in order to relocate all the electronic cables at the right side of the truck frame. As a result, the hydraulic and pneumatic systems were isolated on the left side of the truck frame. An IP67 box was installed to keep all the electronic cable system protected. Moreover, the acceleration pedal was dismantled and several modifications were done to adequate it for the portable traction system requirements. A new one was brought as a replacement in case of mal function. The filter of the engine and the support of the batteries were dismantled in order to relocate them on another part of the frame. After these actions, the width of the system was reduced.

Figure 7 shows the structure designed to cover the engine. The box is made of a metallic sheet. The box seals perfectly the engine and is easy to open. Moreover, the frontal part of the box has small ventilation holes in order to help the engine cooling.

The ECU analysis was carried out to assess the status of the software system of the truck. Some errors were shown, thus in the second stage of this project software modifications will be performed to delete all of them. Despite these errors, the gears were properly working.

Fig. 6 Dismantling of the truck

Fig. 7 Protection box for the motor

6 Conclusions

Testing heavy vehicles and not having a robust guiding structure worsens the test accuracy and small impact point deviations are the consequence. This portable traction system is the only one that offers the possibility of performing tests outside the laboratory and testing vehicles up to 40,000 kg of mass. It also makes possible to perform misuse tests without a human driver.

This system also enables new services to be offered, such as street shows, new configurations, testing of heavy trucks and buses and serves as a support for the laboratory traction system, increasing the performance of rollover and misuse tests and reducing the cost of complex car2car tests.

The next steps of the project are the improvement of the electronic system and developing the ECU software in order to fulfill the requirements of the portable traction system. Moreover, the development of the pulley system and its attachment at the chassis needs to be defined A system of support legs needs to be developed in order to fix the portable system to the floor.

Finally, a validation test must be performed.

Traffic Light Assistant Simulation: Foggy Weather

Jacob Poirier, Fabian Ludwig and Sanaz Motamedi

Abstract In recent years, a variety of new assisted driving systems have been introduced to the automobile market. Advanced Driving Assistance System (ADAS) is becoming a more popular feature in car systems today. It is worth noting that ADAS process best in clear weathers where cameras and sensors are able to detect objects in order to warn the driver in time. However, the weather is not always ADAS friendly. Therefore, this study tested intersection assistance influence on driving behavior in foggy weather. A driving simulator experiment was conducted by using an intersection assistance to send warnings to the driver when approaching an intersection. Various type of warning systems such as audio warning, video displays, and hybrid warning (combination of both audio and visual warning) were examined and compared to no warning system. As a result, we were able to conclude if there was a significant difference in the driver's behavior due to a certain warning type. Compared to no warnings at all, the experiment concluded that a hybrid warning system is the most effective in terms of awareness and speed approaching the intersection followed by audio warnings and then by video displays. For further calculations of our experiment we were able to calculate how much more distance each driver had to stop due to each warning system.

Keywords Advanced driving assistance system (ADAS) · Driving simulator · Warning systems · Bad weather conditions

1 Introduction

Bad weather conditions are a crucial threat in terms of safety to all drivers in any environment and situation. Whether it is raining, snowing, foggy, or even clear at night safety is considered the most important factor to most drivers. More than

J. Poirier (✉) · F. Ludwig · S. Motamedi
Industrial and Systems Engineering, University of Rhode Island, Kingston, Kingston, RI, USA
e-mail: Poirierj@my.uri.edu

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*, Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_92

40,000 people die and over one million people are injured in road collisions just in the European Union alone [1]. ADAS's have been trying to improve these results globally. Whether it is an out-of-the-vehicle or inside-the-vehicle advanced warning system the goal is to decrease the number of accidents in all different environments and situations. There have been many studies that have already looked into the impact on various types of these systems. For instance, Osman et al. [2] experiment was conducted in order to understand the inattentiveness and distraction factor of drivers and try to alter the outcome by implementing advanced warning information inside of the vehicle. The result proved that the system held promise for improving the drivers' safety and decreasing the risk for crashes or collision. However, the detection systems such as cameras and sensors of ADASs functions best in clear weathers in order to warn the driver in time. In the real world, the weather is not always ADAS friendly. Therefore, the objective of the study is to examine the hypothesis if intersection assistance has a significant influence on driving behavior in foggy weather.

In addition to detection system of ADAS's, there are types of warning systems that are affecting drivers' performance. There are three types of warning systems that have been implemented in cars. These three systems are video, audio, and hybrid. Video ADAS systems are usually located between the passenger seat and drivers' seat and tend to be a little more distractive to the driver by taking the attention away from the road in order to see what the warning is. Audio warning systems can occur through common car systems such as Bluetooth or built-in systems in the car that allow the driver to keep their eyes on the road and allow them to hear what is approaching without being distracted. Hybrid systems, for example a GPS, allow the driver to visualize as well as listen in order to fully understand the appropriate task or warning that is coming up.

In this research, a driving simulator experiment was conducted by using an intersection assistance to send various warnings to the driver when approaching an intersection. These warnings include audio, visual displays and a combination of both. Understanding the effects of ADAS systems in such bad weather conditions can give us better understanding of what is the best way of interaction with drivers to keep them safe while they use ADAS system in potential dangerous situations.

1.1 Background

The first contribution regarding traffic light assistants can be found in the late eighties from Australian researchers [3]. However, these first procedures of informing a driver as he or she approaches an intersection were first conducted with the main goal to save fuel—not for safety reasons [4]. Also in recent studies, the detection of intersections and traffic lights have been developed in many research centers but for safety aspects. For instance the Mercedes Benz Research Center along with a European research project developed Prometheus, a vision system for autonomous vehicles including traffic light recognition [5].

There are different ways to inform the driver: First it should be distinguished between inside and outside e.g. on the road side driver notification. Apparently the road side information has the disadvantage of accession and maintenance costs as well as it is depending on the weather conditions. In addition to lead a driver towards an intersection, the signals must be repeated along the road [6]. The other option is to use inside warning signals which can be audio or video, transmitted by a screen or the vehicle speakers.

Within the past couple of years there has been numerous studies regarding the behavior of drivers within certain circumstances. For instance, Liu et al. [7] created a driving simulator study in order to test out which in vehicle ADAS system would allow participants to have a faster response and decision performance in different tasks. In his experiment he examined an audio system, visual system, and a hybrid system in order to see which performed the best. Based on his conclusions it was stated that the hybrid system performed better than both the audio and visual systems in both spatial, response, and divided attention tasks.

Another study that dealt with drivers' performance and behavior was done by Wang et al. [8] who aimed to analyze the effects of environment, vehicle and driver characteristics on risky driving behaviors. In order to see how participants driving behavior performed he designed a decision tree in order to receive results. According to his findings, risky driving behavior was associated with bad weather, poor road and light conditions, and single lane roads. These studies are intended to help provide safer decisions in terms of driving behavior and how much of a difference these aspects can make.

This paper is proposed to test the influence of the in-vehicle driver assistance during bad weather conditions (foggy weather). The main objective is to analyze the different driver behavior with different warning signals and also to derive recommendations for safer driving.

2 Design of the Experiment

To gain insight into drivers' performance associated with the different type warning system in bad weather condition (foggy), a driving simulation experiment were conducted. The factors investigated in the developed driving simulation experiment were categorized into two types: main factors and blocking factors (see Table 1). The two main factors are the Operation (warning systems) and Scenario. In the experiment, four different *operations* including: no assistance, audio assistance,

Table 1 Driving simulation experiment factors and levels

	Factors	Levels
Main factors	*Operation*	None, audio, visual, hybrid
	Scenario	Route 1, Route 2, Route 3, Route 4
Blocking factors	Gender	Male, female

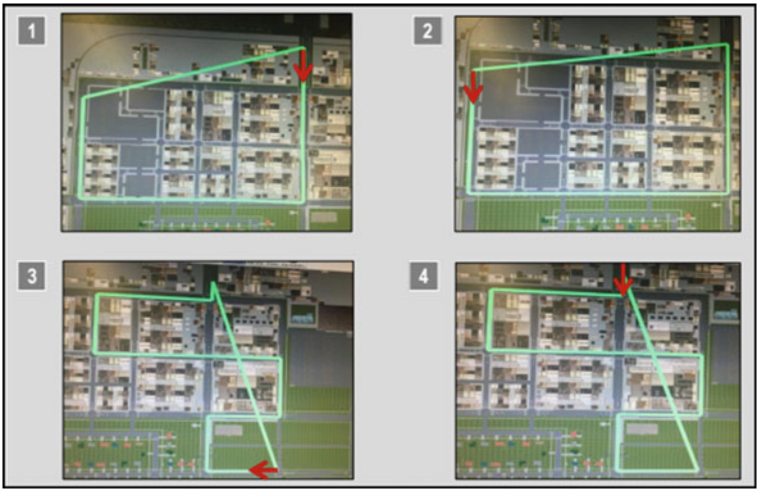


Fig. 1 All four scenario routes (arrow determining direction car travels in)

video assistance, and a hybrid of both audio and video assistance were examined. Moreover, four different types of scenarios were designed. Each scenario varied in turns in direction to keep the experiment randomized and from people able to get familiar with the course. In each scenario we created six different traffic lights that participants had to stop at. All four scenarios are not all exactly the same but are similar. As you can see in Fig. 1, there were only two routes but each route has a different number of right and left turns.

In order to calculate and get a response time for each participant the time it took the participant to go between a predetermined object marked on the side of the road to the traffic light was recorded. The other measurement was the speed that each participant was going every time they passed a predetermined object on the side of the road.

In order to investigate how traffic light assistance influences the driving behavior, it is necessary to conduct the experiment with randomly chosen participants. A total of 10 drivers balanced in gender took part in the experiment whose ages ranged from 19 to 27. The participants were recruited from the University of Rhode Island.

The participants were given a 10-min warm-up run, followed by the experiment. In total, a participant went through 4 scenarios in random sequence. In addition to the random sequence of 4 scenarios, all main factors were randomly chosen. Once the participant felt familiar with the simulator we started the experiment.

In order to input the visual display and audio display of *operations* during the experiment, we had to create two functions. The first function was for the visual *operation.* In order to make a visual *operation* we designed a PowerPoint presentation with alternating slides of a black screen and a sign of a traffic light as shown below (Fig. 2). Since the driving simulator did not have a built in screen we

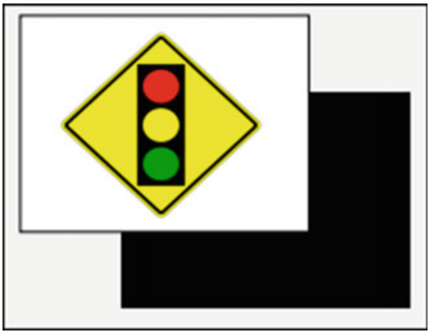


Fig. 2 Alternating slides in powerpoint for video display

were able to place an Ipad where a normal screen would go in a car and use a wireless clicker in order to switch slides when the driver passed a predetermined object. The second function that we needed to provide was an audio voice of a warning signal. We created this using a computer voice recorder and would play the sound at the appropriate time in the experiment. For the hybrid component we were able to add the recorded audio along with the PowerPoint slide with the picture of a stop light. In each experiment one of the group members would be in charge of timing how long it took for the driver to stop from the predetermined object until the stop light, while the other member was in charge of implementing the warning system at the correct time as well as recording the drivers speed at the moment it passes the object.

All experiments were conducted in the Driving Simulation Lab at the University of Rhode Island. A virtual-reality driving simulator in the lab was employed in the experiment. The simulator provides high-fidelity real-world driving environments that can be customized for various applications [9–11]. The TranSim VS IV driving simulator produced by the L3 Corporation was used in this study (Fig. 3). It is a fixed-base simulator consisting of a regular driving module and three channel plasma monitors in an immersive driving environment that combines the look and feel of a real vehicle Participants interact with the simulator using a sedan steering wheel and pedals that provide real-time feedback.

Fig. 3 Driving simulator used in experiment

3 The Analysis

As stated before the aim of this experiment was to see which factors had a significant effect on a participants driving. Through the data analysis, the first objective was test the normality of the distribution. By looking at our probability plot (Fig. 4) we can see the normality of our responses using a 95 % confidence interval. The 95 % confidence level plot has p-value = 0.941 which means the data was normal.

Along with checking the normality of our data, we were then able to construct an Analysis of Variance. The Table 2 illustrated the significance in main factors (scenarios, *operation*) as well as the blocking factors (gender). Since gender was a blocking factor, we already know that it may make a significant impact to our results. Interestingly, two factors of scenario, *operation*, and the interaction between them had a significant influence on drivers' performance. As you can see all of these factors and the interaction have a p-value less than our alpha 0.05, as well as gender. This means that each *operation* and scenario have a significant impact on the response.

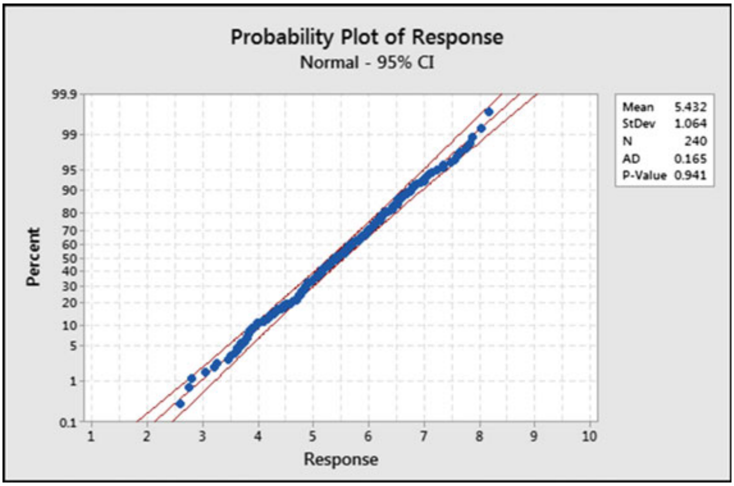


Fig. 4 Probability plot of response

Table 2 ANOVA table

Source	DF	Adj SS	Adj MS	F-value	p-value
Scenario	3	13.840	4.6133	5.88	0.001
Operation	3	38.085	12.6950	16.18	0.000
Gender	1	14.767	14.7670	18.83	0.000
Scenario*operation	9	30.220	3.3578	4.28	0.000
Error	223	174.916	0.7766		
Lack-of-fit	10	9.876	0.9876	1.27	0.246
Pure error	213	165.039	0.7748		
Total	239	270.555			

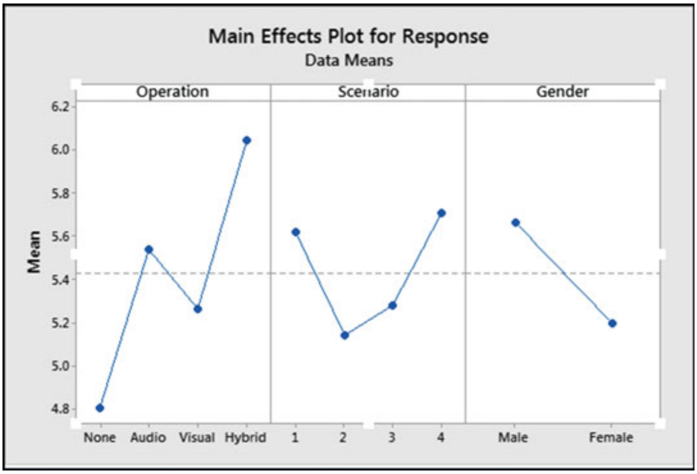


Fig. 5 Main effects plot for response for each level of each factor

In the main effect plot below each factor is being compared to the average mean response time (see Fig. 5). In the *operation* section, it is easily visible that *operation* four, hybrid, has the greatest output with an average of 6.02 s. *Operation* one, no assistance to the driver, is clearly the worst *operation* compared to the rest having an average response of 4.81 s. Another interesting fact about this section is that according to the data the audio assistance had a faster response than the visual assistance by about 0.3 s but is still not as productive as the hybrid.

Since the scenarios were a main factor as well, we can see which scenarios allowed drivers to produce a higher average mean. According to Fig. 5, the scenario four was statistically the best scenario. The second best scenario that produced the highest mean was scenario one followed by the scenario three and lastly scenario two. Even though that the time between each scenario is fairly small there is still a significance of which scenario has a greater effect on the driver. We assume that there is significance in these scenarios in terms of the uniqueness of each route. Although each route is similar in terms of environment and streets, different scenarios have stops at four-way intersections as well as dead ends where drivers are less distracted on seeing other roads and streets approaching in foggy weather.

Although gender was a blocking factor in the experiment its p-value was less than 0.05 we can determine whether each level of each factor performed better than another. According to ANOVA table, gender has a p-value of 0.000 which means that its levels do have significance on the outcome of the response. According to the data, male participants scored a higher mean than females did.

Figure 6 shows the interaction between each scenario with each *operation* to evaluate which combination of the two will produce the most efficient response. As you can see, *operation* four in scenario four will produce the highest response mean for the experiment. According to the plot, the worst combination in our experiment

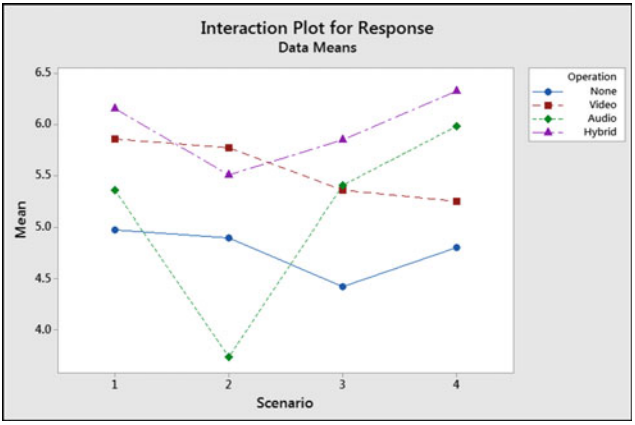


Fig. 6 Interaction plot for responses

would be *operation* three in scenario two. It is worth noting that *operation* four and scenario four would produce the best output because they were both ranked the highest in their levels of factors.

3.1 Differences in Operations

After we conducted our experiment and the average speed for each *operation* was found. We were able to look into further implications to show drivers how much of a bigger difference each *operation* is from one another. As mentioned before in the paper, while calculating the time it took for each participant to go from a predetermined object to the intersection, the speed that the vehicle was going at the predetermined object was recorded. This process employed for each participant on each trial at each object. From this data, the average speed that a vehicle was driving at the specific object was calculated. Therefore, the average mean time that each *operation* took from the object to get to the intersection was measured. The average speed at the predetermined object was about 32.86 m/h = 48.18 ft/s. As you can see in the main effects plot (Fig. 5), the average time each *operation* took for example, no assistance took 4.81 s, audio assistance took 5.54 s, visual took 5.36 s, and hybrid took 6.02 s. With both of these data values, we could conclude how much more distance is available for each *operation* to use while braking to a complete stop. By calculating the difference in time between no assistance versus each *operation* and multiplying it by the average speed, a value in distance was calculated. As a result the operation that allowed the most distance for braking was the hybrid *operation* allowing approximately 58.43 more feet in order to come to a stop versus no assistance at all. Moreover, each *operation* gave the driver about 35.21 ft more than no assistance, and video gave about 26.86 ft more than no assistance.

4 Conclusions and Discussion

After performing the experiment and coming out with our results it is clear to see that different operations certainly do have an effect on a driver's behavior approaching an intersection. We can say that the hybrid operation is the most efficient for this type of experiment. As determined in other studies hybrid warning systems tend to get the attention of drivers more effectively than video or audio systems The study proposed by Liu and Jhuang [7], supports our conclusion that hybrid warning systems in vehicles produced the fasted response time in all response times compared to any other system. Also according to a study formed by Weng et al. [8] concluded that risky driving behavior is associated with bad light or weather conditions. With having hybrid assistance in a car it allows the driver to know what's coming up even if the weather is foggy and as mentioned before gives the driver more time to stop and avoid a collision. However these elements do have an off-putting behavior to the driver. In the real world having a traffic assistance operation in a car could be distracting if you are living in a city and every couple of minutes an audio or a hybrid warning will be announced.

During the experiment we understood that there were some limitations that prevented us from the exact calculations. The two main problems were the car simulator and the stopwatch that we used to calculate the time aspect as well as the visual displays and audio playing of our warning systems. The car simulator obviously is not as realistic as the real world but is the closest representation possible. Since it has all the same functions as a car and most people are at the same skill level, we accepted the fact that there would be some error. The stop watch error was caused by human error in trying to time the moment where the participant passed the predetermined object on the side of the road and the moment at which time the driver came to a complete stop at the intersection. In order to input our visual display of our warning system we had to again physically push a button as soon as the participant crossed the predetermined object on the side of the road which allowed room for human error.

References

1. Plainis, S., Murray, I.J., Pallikaris, IG.: Road traffic casualties: understanding the night-time death toll. Inj. Prev. 2016
2. Osman, O.A., Codjoe, J., Ishak, S.: Impact of time-to-collision information on driving behavior in connected vehicle environments using a driving simulator test bed. JTLE J. Traffic Logist. Eng. **3**(1) (2015)
3. Trayford, R., Crowle, T.: The ADVISE traffic information display system. In: Proceedings of the Conference on Vehicle Navigation and Information Systems. pp. 105–11 (1999) (Conference Record)
4. Trayford, R., Doughty, B.W., van der Touw, J.W.: Fuel economy investigation of dynamic advisory speeds from an experiment in arterial traffic. Transp. Res. Part A Gen. **18**(5–6), 415–419 (1984)

5. Liu, H.X., Ran, B.: Vision-based stop sign detection and recognition system for intelligent vehicle. Transp. Res. Board TRB Annu. Meet. **1748(**20) (2001)
6. Chen, X., Zhou, H., Hsu, P.: What do we know about signal countdown timers. ITE J. web. pp. 72–76 (2009)
7. Liu, Y., Jhuang, J.: Effects of in-vehicle warning information displays with or without spatial compatibility on driving behaviors and response performance. Appl. Ergonom. **43**(2012), 679–686 (2011)
8. Weng, J., Meng, Q.: Effects of environment, vehicle and driver characteristics on risky driving behavior at work zones. Saf. Sci. **50**(2012), 1034–1042 (2011)
9. Wang, J. H., Song, M.: Assessing drivers' tailgating behavior and the effect of advisory signs in mitigating tailgating. In: Proceedings of the 6th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 27–30. Lake Tahoe, California, USA (2011)
10. Motamedi, S., Hasheminejad, M., Choe, P.: Driving safety considered user interface of a smartphone: an experimental comparison. In: Proceedings of the Cross-Cultural Design Applications in Mobile Interaction, Education, Health, Transport and Cultural Heritage (pp. 150–160). Springer International Publishing (2015)
11. Motamedi, S., Wang, J.H.: Study of the impact of hand-held and hands-free cell phone texting on driving performance. Transportation Research Board 95th Annual Meeting (No. 16-3617) (2016)

Skill Metrics for Mobile Crane Operators Based on Gaze Fixation Pattern

Jouh Yeong Chew, Koichi Ohtomi and Hiromasa Suzuki

Abstract This paper proposes skill metrics that analyzes the gaze fixation pattern of mobile crane operators. This study focuses on the fixation data because they commonly represent visual information processing. First, scenes of crane operation are divided into content-based Area of Interests (AOIs) and the Markov Chains is used to model the gaze transitions between these AOIs. Four metrics were introduced to interpret the model at different expertise levels. Results suggest that experienced operators exhibit lower metrics compared to their novice counterparts, and adapted entropy measures exhibit similar patterns as the original ones in the previous study. Most importantly, the proposed metrics are able to address the following issues, i.e. large and sparse model, as highlighted in the previous study.

Keywords Gaze fixation · Markov chains · Skill metrics · Crane operation

1 Introduction

Gaze behavior analyses have attracted huge interest because visual cues consist of rich information that aid decision-making. Some common applications were in Kansei Engineering for product design and development [1], computer interface design [2], and medical equipment design to improve the detection of breast cancer [3]. In sports engineering, it was used to analyze gaze patterns that help to save soccer penalties [4], and to study visual and motion perception in tennis [5]. These studies suggested the relevance of understanding gaze behaviors in attempts to achieve better performance.

J.Y. Chew (✉) · K. Ohtomi · H. Suzuki
Department of Precision Engineering, University of Tokyo, Tokyo, Japan
e-mail: jychew@delight.t.u-tokyo.ac.jp

K. Ohtomi
e-mail: Koichi.ohtomi@delight.t.u-tokyo.ac.jp

H. Suzuki
e-mail: suzuki@den.t.u-tokyo.ac.jp

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_93

In the case of a highly specialized task such as crane operation, gaze patterns may conceal knacks of efficient operation, and this is similar to what one encounters in sports engineering [4–6]. This task is highly dependent on the swing control and load oscillation, both which result from the crane movement inertia and the environmental context such as the wind speed. Thus, an appropriate line of vision may suggest a better operator-machine interaction to resolve these issues. This paper proposes skill metrics that interpret and compare the gaze patterns of crane operators at different expertise level in order to validate the assumption above.

Fixation and saccade are the two primary types of gaze patterns. Fixation refers to fixed gazes on a specific point-of-interest and a saccade refers to the transitions between two fixations. Analyses were primarily done by observing the scan paths [7] and the frequency of fixation [8]. However, this is inadequate to provide a quantitative interpretation of gaze behaviors, especially when the number of fixation is large. Thus, [4, 9, 10] proposed to model the gaze behaviors using Markov Chains. This method is acclaimed because it provides transition probabilities between different AOIs, and thus is regarded as a good summary of gaze behaviors. However, two issues remain as challenges—(1) the interpretation of the model in the form of transition matrix, and (2) the implementation of the model in solving real problems.

A study [10] proposed to use Shannon entropy to interpret the transition matrix model. It evaluated the correlation between aesthetic experiences and gaze behaviors in which scenes were divided into small Area of Interests (AOIs) using grids; and the transitions between these AOIs were calculated using Shannon entropy. Results suggested that these measures were suitable for quantitative interpretations of gaze behaviors. However, the model used in the study was small and non-reducible. In other words, the number of AOIs was small and there was no AOI with zero fixations.

This study proposes (1) the skill metrics that evaluate the experience of crane operators using their gaze behaviors, and (2) additional and adapted measures for the quantitative evaluation of gaze patterns. To our best knowledge, there has been no prior study that uses gaze fixation as the experience measure. The following contributions are proposed—(1) a novel method that measures experience by modeling gaze fixation after Markov Chains and (2) the analysis of large and sparse gaze transition model is made possible using new measures, after which the results exhibit a similar pattern as that of in the previous study.

2 Experiment Setup

This section discusses the background of crane operation and its corresponding visual scenes. The basic control parameters of a crane and the operation used for the experiment is briefly explained. In addition, the setting of the visual scenes into the content-based AOIs is elaborated.

2.1 Crane Operation

There are many challenges in achieving effective crane operation. One of the challenges is to overcome swings and the vibration of the load. It is also difficult to move the load in radial direction, where accurate depth perception is desired. Despite this, crane control consists of only four primary inputs (i.e. Hoist, Boom Angle, Boom Length, and Slew, as illustrated in Fig. 1), each of which manipulates one parameter. However, the operation used for the experiment was performed by manipulating three parameters. Boom Length is excluded because of its small operating radius. This operation is further illustrated and explained in Fig. 2.

Experiments were conducted by measuring the gaze fixations of mobile crane operators at different expertise level, i.e. Beginner, Intermediate, and Expert. One of the challenges in this study is the difficulty in getting participants to conduct the experiment. This is primarily because crane operation is regarded as a highly

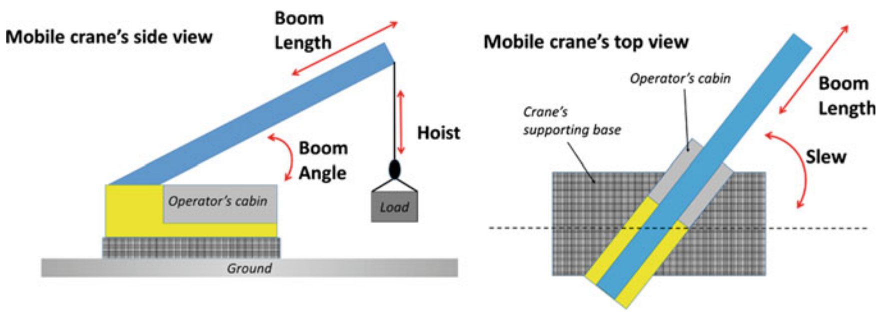


Fig. 1 Basic control parameters of a crane on the side view (*left*) and *top* view (*right*). There are four basic controls, i.e. (*boom length*), (*boom angle*), (*hoist*), and (*slew*). (*red arrows*) illustrate the corresponding motion for each parameter

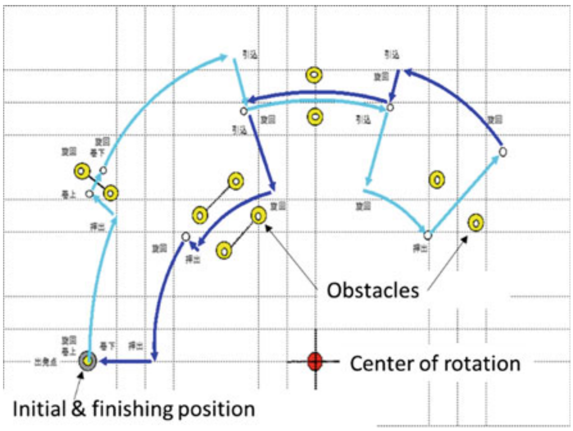


Fig. 2 *Top* view of the path where the load is being moved with respect to the center of rotation (*red circle*) during operation. The load starts from the initial position (*grey donut*), moves in the direction (*blue and cyan arrows*), passes through obstacles/poles (*yellow donuts*) before finishing at the initial position

specialized task, during which operators need to overcome swings and the oscillation of the load. Thus, there are safety elements to be considered while conducting the experiment. There were seven participants and each of them performed multiple trials, which provided 19 samples for analysis. Despite the relatively small sample, it is adequate to provide a preliminary overview of the relationship between experience and gaze behaviors.

2.2 Gaze Fixation

Previous study [11] suggested that scene perception usually takes between 260 and 330 ms. In other words, fixation is more commonly associated with visual information processing and thus, serves as the main interest in this study compared to saccadic response. One of the challenges in gaze behavior analyses for crane operation is the large operating background, which makes grid-based AOIs analysis difficult. This was addressed by focusing on the area around the moving load because the load was commonly within operator's view. Content-based AOIs analysis was used because the background scene was dynamic and it mostly consisted of irrelevant information. This setting limited the spatial information to area around the load and filtered gaze fixations on the background, which was regarded as noise in this study. Gaze fixation was measured using the Tobii Glasses 2 and classified into different AOIs using the Tobii Analysis software. The original scene and the corresponding AOIs are shown in Fig. 3. Table 1 explains the nine AOIs that were used in this study. The number of AOIs is significantly larger compared to the previous study [10], which used up to three AOIs. This setting facilitates study on large and sparse model, which was cited as one of the challenges.

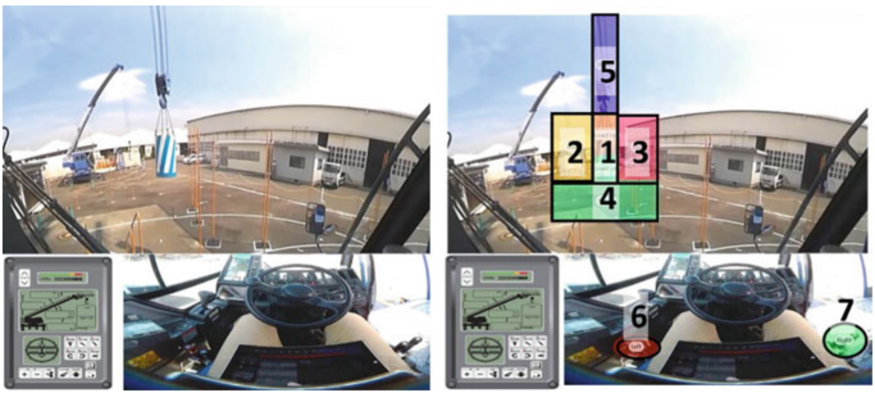


Fig. 3 The original scene (*left*) and content-based AOIs (*right*)

Table 1 Settings of AOIs for gaze fixation analysis

AOIs	Definition
Load (1)	The object that is being moved around by the crane
West (2)	Area on the right of the load
East (3)	Area on the left of the load
South (4)	Area below the load
Winch (5)	The hook/winch above the load
Left (6)	Control levers on the operator's left
Right (7)	Control levers on the operator's right
NoAOI (8)	Virtual AOI which defines gaze fixations on areas other than AOIs listed above
NoFix (9)	Virtual AOI which defines eye movements other than fixations (includes missing data)

3 Results and Analysis

This section explains the representation of gaze behavior using Markov Chains and the corresponding measures to interpret the model. Transition between fixations on these AOIs was modelled as transition matrix, where rows represent current gaze fixation and columns represent the next fixation. This model is preferred compared to scan paths [7] and heat map [8] because it facilitates the estimation of next fixation given the current one using the first order Markov Chains. Qualitative interpretation of Markov Chains model is possible by observing the color scale of probability distribution. However, this method is subjective and difficult for larger models. Thus, additional measures are required to interpret the model for design applications.

3.1 Markov Chains

This section proposes the Markov Chains as a quantitative model for gaze fixation. The AOIs are modelled as state space S in Eq. (1), where M represents the total number of AOIs. Gaze data from the Tobii eye tracker is down-sampled from 50 to 25 Hz before being used to model the Markov Chains. Given X_n is the AOI at nth time step and $i, j \in S$, transition from state i to state j is represented by Eq. (2), where p_{ij} is an element of transition matrix. Transition and stationary probability is calculated by Eqs. (3) and (4), respectively. Variable n_{ij} represents the count of transition from state i to j, and $n_i = \sum_j n_{ij}$. The model also considers additional rules based on [12], where $p_{ij} = 0$ and $p_{ii} = 1$, when $n_i = 0$. This yields $\sum_j^M p_{ij} = 1$.

$$S = \{s_1, s_2, \ldots, s_M\}. \quad (1)$$

$$P(X_{n+1} = j | X_n = i) = p_{ij}. \quad (2)$$

$$p_{ij} = n_{ij} / \sum_j n_{ij}. \quad (3)$$

$$\pi_i = n_i / \sum_i n_i. \quad (4)$$

Markov Chains model was calculated for 19 trials recorded from all participants, and Fig. 4 compares gaze behavior for an Expert and a Beginner. Qualitative observations of these models suggest two significant differences. First, diagonal elements of the Expert's gaze model on the left are mostly higher than the Beginner's. This indicates an Expert is more likely to gaze at the same AOI during the operation. Secondly, the Expert's model consists of more zero elements and is less uniformly distributed compared to the Beginner's. For instance, the Expert exhibits no gaze interaction on the Left and Right AOIs. This is reasonable considering that the Expert is familiar with the position of operation levers. This compares with the Beginner who needs to check the position of these levers during operation. However, this interpretation is subjective and may not be useful when the difference is not significant or when the model size is larger. In other words, visualization of the model with many AOIs is not effective. Thus, it is desirable to use quantitative measures for objective interpretations of the model.

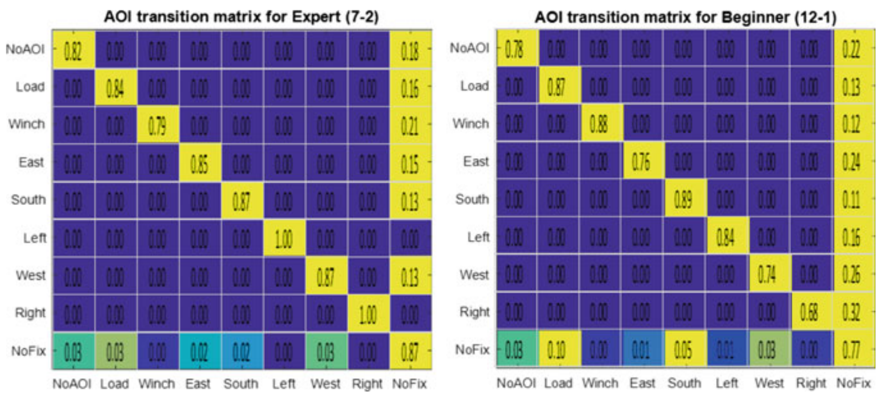


Fig. 4 Comparison of Markov Chains model for an Expert (*left*) and a Beginner (*right*). The *rows* represent current AOI and the *columns* represent the next AOI

3.2 Skill Metrics

Markov Chains model has been used in previous study [4, 9, 10]. However, there are limited works to interpret the model and to put it to practical use. A study [10] proposed entropy measure for non-reducible and small model, which consisted of up to three AOIs. The study cited higher number of AOIs and sparseness of model as challenges, which were addressed in this study. Larger number of AOIs was used for the analysis of mobile crane operator's gaze fixation, which resulted in sparse and reducible model. Four metrics were introduced to interpret the model for each level of expertise.

The first and second metrics were adapted from the previous study [10], i.e. stationary and transition entropy. They are represented by H_s and H_t, respectively. The stationary entropy indicates the distribution of attention, where larger value corresponds to uniform distribution. On the other hand, transition entropy represents exploratory behavior, where larger metric corresponds to transition that is more active or random. In this study, infinitesimal $\alpha = 10^{-6}$ was added to the log term in Eqs. (5) and (6) to enable calculation for sparse and reducible model. This adaptation resulted in patterns that is similar to those of the original ones, and successfully addressed one of the concerns highlighted by [10], i.e. application on sparse model.

$$H_s = -\sum_{i \in X} \pi_i \log(\pi_i + \alpha). \tag{5}$$

$$H_t = -\sum_{i \in X} \pi_i \sum_{j \in X} p_{ij} \log(p_{ij} + \alpha). \tag{6}$$

Matrix density ρ_m serves as the third metric to indicate the sparseness of attention. Larger metric suggests that more AOIs were covered during gaze transitions. In other words, it calculates the ratio of non-zero elements to the total number of matrix elements. The last metric D serves as an indicator of continuous fixation or duration of attention. It was calculated using norm-2 of eigenvalues $\|\lambda_2\|$, where smaller metric suggests continuous fixation or focus on specific AOIs. In this case, variable M refers to the model size or number of AOIs. The third and fourth metrics were introduced to better interpret the enlarged model, which was the second challenge cited by [10]. Calculations of the last two metrics were done using Eqs. (7) and (8).

$$\rho_m = \frac{\#\,\text{Non-zero matrix elements}}{M^2} \tag{7}$$

$$D = \sqrt{M} - \|\lambda\|_2. \tag{8}$$

Metrics for all the 19 transition matrices were calculated using Eqs. (5)–(8) and the results are shown in Fig. 5. The bars represent average metrics which are

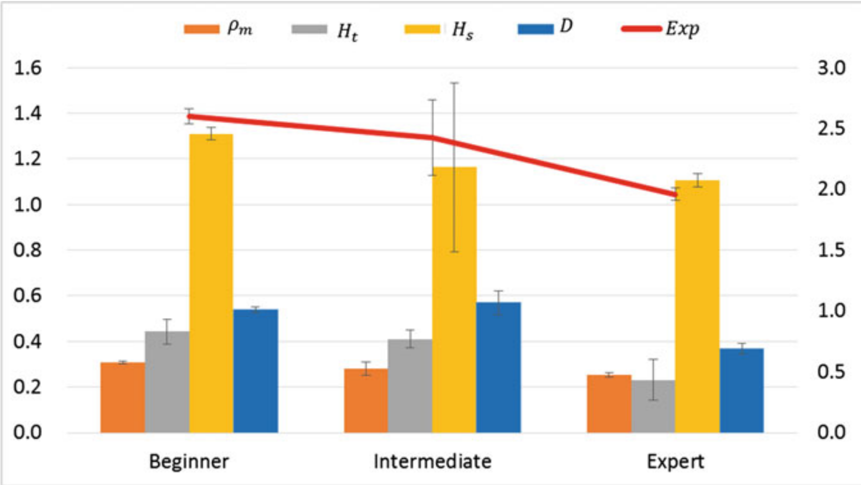


Fig. 5 Skill metrics (*red line*) and average of individual metrics (*bar charts*) for Beginner, Intermediate and Expert. Results suggest Experts tend to exhibit smaller values, which is consistent for all metrics

grouped into three clusters for Beginner, Intermediate and Expert. Comparison between Beginner and Expert suggests that the latter exhibits smaller values, which is consistent for all metrics. The red line represents the skill index, which is the summation of all metrics. In this case, a smaller value indicates higher skills, and the vice versa. In general, Intermediate exhibits metrics which mostly fall between those of Beginner and Expert. It is also worth noting that the standard deviation for this group is higher. This possibly results from the error in classifying the skills of the operator. In short, the results suggest different response at each level of expertise, and adapted entropy measures exhibit similar pattern as original ones in the previous study. In other words, the new measures seem to be useful to define the skills of crane operators, especially when gaze transition matrix model is large and sparse.

4 Discussion

The results of skill metrics suggest that Expert tends to exhibit non-uniform gaze distribution because of small H_s. It is reasonable to consider that Expert knows where to look at during the operation. This behavior is observed from the time plot of gaze fixation on each AOI in Fig. 6, where Expert exhibited less uniform gaze which covered only five AOIs. It is also consistent with small H_t, which suggests that Expert exhibits less random gaze during the operation. In this case, continuous fixation is observed in Expert's gaze fixation in Fig. 6, which is also consistent with

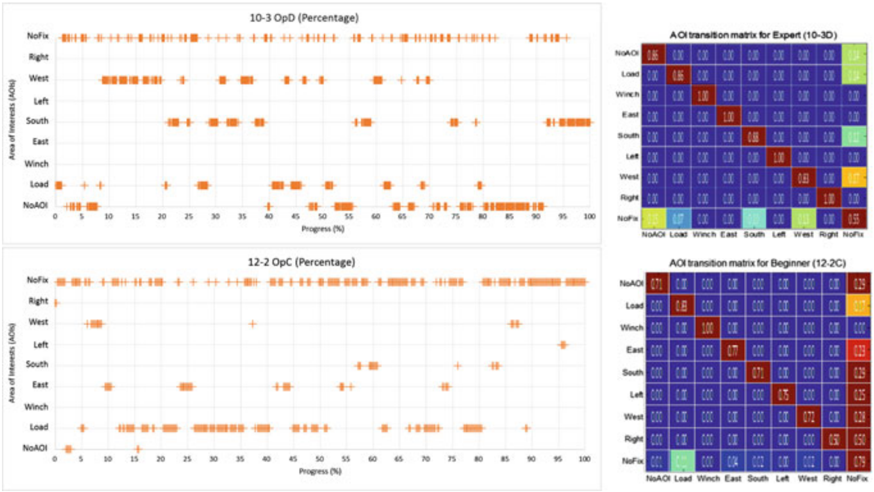
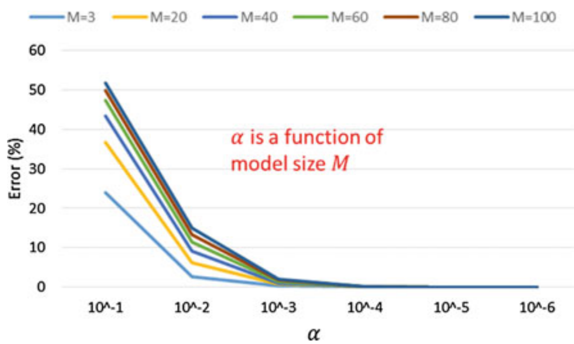


Fig. 6 Time plot of gaze fixations on each AOI for Expert (*top*) and Beginner (*bottom*). Metrics $(H_s, H_t, D, \rho_m, Exp)$ for Expert = (1.23, 0.05, 0.28, 0.21, 1.77) and Beginner = (1.23, 0.58, 0.69, 0.27, 2.77)

smaller D. It suggests that Expert spends more time on a smaller number of AOIs during the operation. In short, Experts exhibit relatively lower metrics, i.e. gaze pattern is non-uniform, and stable with focus. On the other hand, Beginners exhibit relatively higher metrics, i.e. gaze pattern is uniform and random without focus. This behavior is consistent with those of previous studies [2, 6], where expert tennis players exhibit consistent gaze because they know where to focus during the game. This same reason could be applicable to crane operators.

The results also suggest that the skill metrics are suitable to analyze large and sparse model, which was cited as one of the challenges in [10]. In this study, entropy metrics H_s and H_t were adapted by adding an infinitesimal $\alpha = 10^{-6}$ to the log term in Eqs. (5) and (6) to enable calculation for sparse models. Figure 7 suggests that this adaptation has minimal effects on the original calculation. Despite

Fig. 7 Error calculated using adapted H_s and H_t compared to the original equations in [10]. The error decays as M gets larger and as α becomes smaller



α being a function of model size M, it is adequate to use an infinitesimal constant $\alpha = 10^{-6}$ for M = 100 because the error decays as M gets larger and as α becomes smaller. In addition, additional measures are introduced to better interpret the gaze behavior of transition matrix. Referring to Fig. 6 again, it is worth noting that H_s did not exhibit significant differences. This possibly results from the variation of enlarged models. However, it is compensated by other metrics which exhibited smaller metrics for the Expert. Thus, the proposed skill metrics complement each other in interpreting large and sparse gaze transition models.

Although previous studies [4, 9] have proposed Markov Chains to model gaze behaviors, it is difficult to provide an objective interpretation and to use it to solve real problems. These issues are addressed in this study, where skill metrics are used for objective interpretations of Markov Chains models. In addition, the models are used to represent the gaze behavior of crane operators in actual operating environment, which serves as practical implementation.

5 Conclusions

Markov Chains are used to model the gaze fixation behavior of mobile crane operators and skill metrics are used to interpret the models to estimate their experience level. Adaptation of entropy measures and introduction of new measures facilitate the interpretation of large and sparse models. Despite the adaptation, the results are consistent with those of previous study. However, the results serve as preliminary review due to the relatively small sample size. Further validation is necessary but it remains a challenge to increase the number of samples. One of the possible solution is to use simulator for validation. The skill metrics are useful to facilitate future design works, such as to identify optimal user interface to improve user-machine interaction, and to improve ease of use of a system.

Acknowledgments This research was supported by New Energy and Industrial Technology Development Organization (NEDO) of Japan, and we would like to thank them for their assistance.

References

1. Button, C., Dicks, M., Haines, R., Barker, R., Davids, K.: Statistical modelling of gaze behavior as categorical time series: what you should watch to save soccer penalties. Cogn. Process. **12**, 235–244 (2011)
2. Ward, P., Williams, A.M., Bennett, S.J.: Visual search and biological motion perception in tennis. Res. Q. Exerc. Sport. **73**, 107–112 (2002)
3. Kohler, M., Falk, B., Schmitt, R.: Applying eye-tracking in Kansei Engineering methodology for design evaluations in product development. Int. J. Affect. Eng. **14**, 241–251 (2015)
4. Goldberg, J.H., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. **24**, 631–645 (1999)

5. Mello-Thoms, C., Nodine, C.F., Kundel, H.L.: What attracts the eye to the location of missed and reported breast cancers?. In: Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 111–117. ACM, New York (2002)
6. Singer, R.N., Williams, A.M., Frehlich, S.G., Janelle, C.M., Radlo, S.J., Barba, D.A., Bouchard, L.J.: New frontiers in visual research: an exploratory study in live tennis situations. Res. Q. Exerc. Sport. **69**, 290–296 (1998)
7. DeBruin, J.A., Malan, K.M., Eloff, J.H.P.: Saccade deviation indicators for automated eye tracking analysis. In: Proceedings of the 2013 Conference on Eye Tracking South Africa, pp. 47–54. ACM, New York (2013)
8. Grindinger, T., Duchowski, A.T., Sawyer, M.: Group-wise similarity and classification of aggregate scanpaths. In: Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 101–104. ACM, New York (2010)
9. Otsuka, K., Yamato, J., Takemae, Y., Murase, H.: Quantifying interpersonal influence in face-to-face conversations based on visual attention patterns. In: Human Factors in Computing Systems, pp. 1175–1180. ACM, New York (2006)
10. Krejtz, K., Smidt, T., Duchowski, A.T., Krejtz, I.: Entropy-based statistical analysis of eye movement transitions. In: Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 159–166. ACM, New York (2014)
11. Keith, R., Monica, C.: Eye movements. Scholarpedia **2**, 3649 (2007)
12. Teodorescu, J.: Maximum likelihood estimation for Markov Chains. arXiv:0905.4131v1 [stat.CO] (2009)

Towards a Real-Time Driver Workload Estimator: An On-the-Road Study

Peter van Leeuwen, Renske Landman, Lejo Buning, Tobias Heffelaar, Jeroen Hogema, Jasper Michiel van Hemert, Joost de Winter and Riender Happee

Abstract Driver distraction is a leading cause of crashes. The introduction of in-vehicle technology in the last decades has added support to the driving task. However, in-vehicle technologies and handheld electronic devices may also be a threat to driver safety due to information overload and distraction. Adaptive in-vehicle information systems may be a solution to this problem. Adaptive systems

P. van Leeuwen (✉) · J. de Winter · R. Happee
Department of BioMechanical Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands
e-mail: P.vanLeeuwen@TUDelft.nl

J. de Winter
e-mail: J.C.F.DeWinter@TUDelft.nl

R. Happee
e-mail: R.Happee@TUDelft.nl

R. Landman
Ergos Human Factors Engineering, Hengelosestraat 448-a, 7521 AN Enschede, The Netherlands
e-mail: Renske.Landman@Ergos.nl

L. Buning
HAN University of Applied Sciences, Ruitenberglaan 29, 6826 CC Arnhem, The Netherlands
e-mail: Lejo.Buning@HAN.nl

T. Heffelaar
Noldus Information Technology, Nieuwe Kanaal 5, 6709 PA Wageningen, The Netherlands
e-mail: T.Heffelaar@Noldus.nl

J. Hogema
TNO, Perceptual and Cognitive Systems, Kampweg 5, 3769 DE Soesterberg, The Netherlands
e-mail: Jeroen.Hogema@TNO.nl

J.M. van Hemert
TomTom BV, Oosterdokstraat 114, 1011 DK Amsterdam, The Netherlands
e-mail: JasperMichiel.vanHemert@TomTom.com

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_94

could aid the driver in obtaining information from the device (by reducing information density) or prevent distraction by not presenting or delaying information when the driver's workload is high. In this paper, we describe an on-the-road evaluation of a real-time driver workload estimator that makes use of geo-specific information. The results demonstrate the relative validity of our experimental methods and show the potential for using location-based adaptive in-vehicle systems.

Keywords Driver distraction · Adaptive in-vehicle information (systems) · Driver workload estimation

1 Introduction

Driver distraction is a leading contributor to road traffic crashes [1]. A recent naturalistic driving study showed that as much as 78 % of crashes were related to distraction [2]. Because of the increasing prevalence of technological aids, road safety has improved considerably in the last decades. However, certain in-vehicle technologies such as infotainment systems and handheld electronic devices are themselves a source of distraction and crash risk [1, 3–6]. Distracted driving not only reduces lane-keeping accuracy [7, 8] but also increases the brake reaction time to critical environmental events [9]. Furthermore, a complex in-vehicle display may result in an 'information overload' [10].

A potential remedy to these problems may be the use of adaptive information systems [11]. Adaptive information systems aid the driver by warning for upcoming high-workload situations or by adapting the information presentation. For example, when driver workload is high, an adaptive system may (1) switch to auditory presentation instead of visual presentation, (2) reduce the amount of information, or (3) present no information.

An workload-adaptive in-vehicle information system not only requires knowledge of the current driver workload. An estimate of the *future* workload is required as well. The use of the momentary workload only as input to the adaptive in-vehicle system would result in the adaptation being too late for the driver, and therefore drivers would not benefit from such a system [12]. Prediction of driver workload may seem a difficult task [13] due to the dynamics of traffic, interactions between road users, and moment-to-moment driver variability. Verwey [4, 14] found that the traffic situation is a vital determinant of workload that could be used for real-time workload estimation.

Similar to the approach by Verwey [4], we propose to estimate driver workload based on the location of the vehicle in the world. Specifically, using GPS coordinates and an online route map database, the driver's workload was estimated in real time based on road type, distance to junctions, and vehicle speed and acceleration. In our project, we explored the feasibility of using the workload estimate for real-time adaptation of visual information presentation on a navigation device (see also [15]).

Previous research has demonstrated the measurement of driver workload using physiological measures [16–18], measures of driver performance [19], and self-report evaluations [20]. In the present paper, we evaluated our experimental vehicle and our driver workload estimator in an urban, rural, and highway environment. Specifically, we evaluated vehicle speed, driver inputs, heart rate, respiratory rate, eye gaze, pupil diameter, and self-reported effort as a function of travelled distance along the route, a secondary mental arithmetic task, and the estimated workload level.

2 Methods

2.1 Participants

Six participants from the HAN employee community volunteered to participate in this research. Participants filled out an intake questionnaire with general items (age, gender, wearing glasses, driving history, use of navigation systems).

The participants were four males (mean age: 28.5, *SD* = 7.8) and two females (mean age: 29.0, *SD* = 1.4). Participants were in possession of a driver's license for an average of 8.7 years (*SD* = 5.1) with a minimum of 3 years and reported a mean annual mileage of 12,217 km (*SD* = 11,398). Four participants mentioned the use of glasses, one participant wore glasses during the experiment, and one participant reported the use of contacts. All participants indicated the use of navigation devices in their normal driving.

2.2 Apparatus

The experiment was conducted using a manual drive E91 320d BMW test vehicle. The vehicle was equipped with data acquisition connected to the CAN bus, allowing the collection of vehicle state variables (e.g., speed) and driver input variables (e.g., steering angle). Participants' physiological responses were measured using ECG electrodes and a respiration belt from TMSi (PolyBench, software version 1.30.0.3521) placed around the chest. Eye and head movements were recorded using a remote-mounted eye tracker from SmartEye (SmartEye Pro, software version 6.1.4). All data were synchronized and stored using The Observer XT (Noldus, software version 12.0.825 NBD) at sampling rates varying from 5 to 60 Hz. The navigation device was an Android tablet with prototype TomTom navigation software (Samsung Galaxy Tab 2, P3110 with Android 4.0).

2.3 Driver Workload Estimator

TNO, in collaboration with TomTom, developed the real-time workload estimator prototype. The estimator made use of vehicle and driver input data as well as road type estimated from the geographical location, based on GPS coordinates and a route map database.

On a high level, the estimation process had several components: road type, time/distance to junctions ahead, acceleration of the car, driving speed (with respect to the speed limit), and time-on-task (how long the driver has been driving without a break). For each component, decision rules were specified that indicated the relationship between the component and workload. The components were merged into a final output of the driver workload estimator, representing a 6-point workload estimate ranging from *very low* to *very high*.

2.4 Procedures

Before the start of the experimental sessions, participants received oral instructions explaining the experiment and procedures. Furthermore, participants filled out the intake questionnaire, a consent form, and a traveling cost form. Next, participants performed a Landolt C test [21] to determine their visual acuity. If the visual acuity was at least corrected-to-normal, the participants were allowed to participate.

After the visual acuity test, participants received oral instructions about the driving task. Furthermore, the self-report procedures and the secondary task were explained and practiced by the participants. After taking place in the vehicle, participants adjusted their seat to their own preference. The ECG and respiration hardware was connected to the participants, and the eye tracker was calibrated by means of a series of eye and head movements.

Participants drove three sessions: a baseline session and two times the same route of approximately 40 min. Participants drove the baseline session starting from the university campus to the starting point of the first session, a drive that took approximately 5 min. After the first 40 min session, participants had a 10 min break after which they drove back to the starting point. After completing the second 40 min session, participants drove back to the campus. When arrived at the campus, participants filled out a questionnaire regarding their driving behavior and received a gift card.

While driving, participants performed a secondary arithmetic task and rated their effort using the Rating Scale Mental Effort (RSME) [22]. An experimenter sat on the passenger seat, and initiated the secondary task and marked the RSME scores. Furthermore, the experimenter marked sudden events (e.g., pedestrians crossing, unpredictable behavior of other road users).

2.5 Driving Task and Environment

Prior to the baseline session, participants received oral instructions to drive as they would drive their own car, to adhere to Dutch traffic rules including speed limits, and to be aware of other road users. In addition, drivers were asked to perform a secondary task to temporarily add workload to the driving task. Specifically, at several moments during the drive, the experimenter instructed the participants to countback in steps of three from a random number between 60 and 100.

The route was identical for all participants and both sessions, and started and ended at the same locations. Each participant drove the same route twice. A tablet with TomTom navigation concept software provided the participants with driving directions by means of a Dutch voice. After completing the first session, participants drove from the endpoint of Session 1 to the starting point of Session 2.

The route was chosen so that different traffic situations were likely to occur. The route was near the city of Arnhem (NL) and had a length of 21.5 km. The route consisted of intersections (with and without traffic lights), roundabouts, urban areas with a 30 kph speed limit, a small segment of rural area, and a highway.

The countback task and RSME rating were performed at several locations along the route. On average, participants were requested to score their RSME 6 times per session and perform the countback task 5 times per session.

2.6 Dependent Measures

The following dependent measures were computed per session. The measures can be categorized as (1) vehicle performance, (2) driver input, (3) driver physiology, (4) subjective reports, and (5) the driver workload estimate.

1. Mean speed (kph) and absolute vehicle acceleration (m/s^2) were calculated as a measure of task efficiency, driving style, and driving safety.
2. The mean absolute steering speed (°/s) and steer steadiness (%, defined as the percentage of time the absolute steering speed was lower than 1 °/s) were used to represent steering activity [23, 24]. The mean absolute throttle speed (%/s) was used to indicate throttle activity.
3. The mean heart rate (1/min) and the mean respiration rate (1/min) were recorded from the ECG and the respiration belt hardware, respectively. The mean pupil diameter (mm) measured by the eye tracker data was used as a measure of workload [25]. Eye gaze data were classified into four regions of interest: (1) the road center (defined as a cone with 8° radius around the road center), (2) the peripheral area (defined as the region outside the road center, but within the windscreen perimeter), (3) the dials and navigation, (4) and other. For a definition of the dials and navigation, see Fig. 2. The mean percentage gaze at

the road center (GRC, %) represents the amount of attention directed to the road ahead [26]. Eye movement data were low-pass filtered at 5 Hz because the eye tracker data were sensitive to external noise, such as vibrations. Missing data (e.g., eye blinks and camera obstructions) were automatically removed.

4. The mean RSME (0–15) was determined from the rating scale [22] that was administered during driving. To keep interference with the driving task to a minimum, the participants indicated their effort orally on a scale from 0 to 15 (equivalent to the RSME vertical line of 15 cm) where 3 is ‘normal driving’ or ‘a comfortable task load’ and 12 is more than ‘extreme effort’.
5. The driver workload estimate (1–6) was obtained from the online estimator. As mentioned above, workload was estimated based on vehicle location and vehicle state.

3 Results

Due to low quality eye tracker data (defined as less than 20 % reliable data), the gaze data of two participants were removed. Of the remaining four participants, on average 30 % (*SD* = 14 %) of eye tracker data were removed, due to the tracker’s inability to record eye movements. One participant made a navigation mistake and drove an additional segment (approximately 1.06 km) during the first session. The data of this additional segment were removed.

3.1 Descriptive Results

Figure 1 provides an overview of several of the variables during the experimental route. The figure illustrates the diversity in road types (e.g., the first 4.5 km of the route consisted of a highway) and the differences in driving speed and steering activity along the route. The figure also shows the percentage of gaze at the road center, illustrating the gaze activity near corners and intersections. The RSME values seem to reveal an elevated self-reported workload at travelled distances of 5 and 19 km. Furthermore, the driver workload estimator shows that levels 3 and 4 occurred most frequently, whereas level 5 occurred intermittently.

Figure 2 shows the gaze distribution of one selected participant, illustrating the regions of interest and the main areas of visual attention. The gaze samples are centered on the 8° circle that represents the road center (averaged across the two sessions, participants directed their gaze 60 % of the time at the road center). The dials and the navigation device were glanced at for 5 % of the time (for all participants during both sessions). The gaze swirls to the left and right of the road center indicate lateral eye movements, for example while driving in a curve.

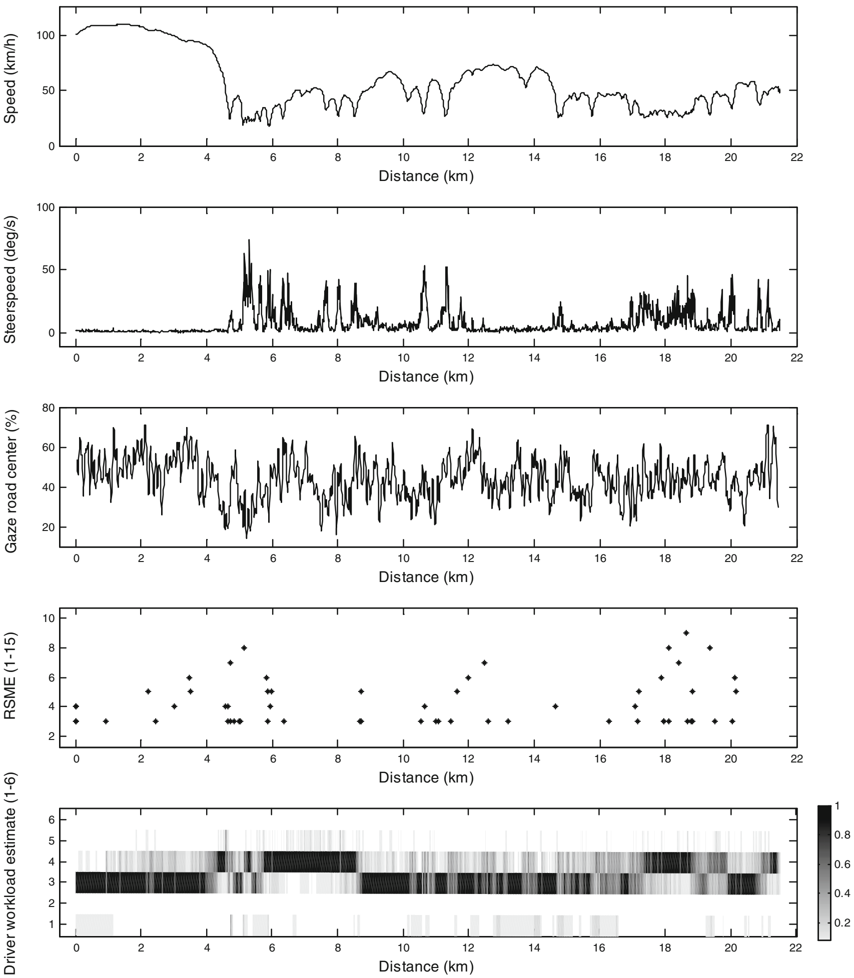


Fig. 1 Driving speed, absolute steering speed, gaze road center, Rating Scale Mental Effort (*RSME*), and workload estimate distribution as a function of travelled distance along the experimental route. The speed, absolute steering speed, and gaze road center were averaged across participants and sessions. All RSME reports for all participants and both sessions are shown. The workload estimate distribution was determined by averaging across the six participants and two sessions, and ranges from 0 out of 12 (*white*) to 12 out of 12 (*black*)

Figure 3 illustrates the association between driving speed and gaze distribution (left) and between driving speed and steer steadiness (right). It suggests that participants were more likely to allocate attention to the road center with increasing driving speed (left). Moreover, steering steadiness increases with increased driving speed (right).

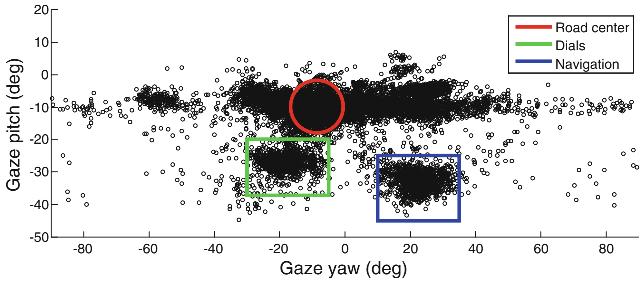


Fig. 2 Raw gaze data of one selected participant, together with regions of interest

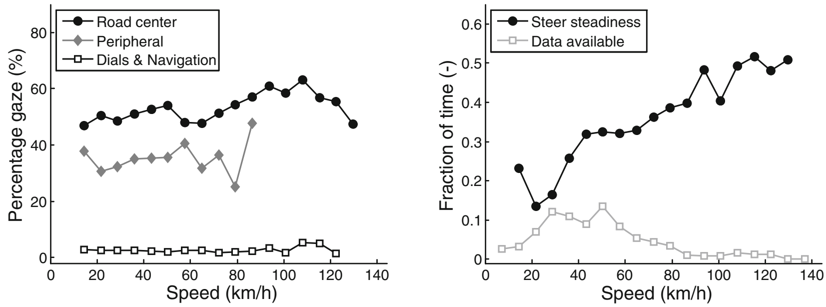


Fig. 3 Gaze distribution (*left*) and steering activity (*right*) as a function of driving speed. Data were extracted per 7.2 kph wide bin, and averaged across participants and both driving sessions. The data from 0 to 10.4 kph were removed from the figure. Note that participants drove faster than 90 kph for less than 5 % of the time, which explains the oscillatory behavior of the distributions for speeds greater than 90 kph. The *gray line with square markers* in the right figure indicates how much data were available at a given driving speed

3.2 Countback Task

Table 1 shows the results of selected measures averaged across the 10 s period before and the 10 s after the start of the countback task. Figure 4 illustrates the effect of the countback task on participants' heart rate and respiration rate. No statistically significant differences and small effect sizes between the two periods were observed for the driving performance measures (mean speed, steer speed, and throttle speed). The heart rate increased slowly from the start of the countback task and peaked at about 10 s after the start of the countback task. Furthermore, the respiration rate decreased after the start of the countback task. No clear differences were observed for the pupil diameter before versus after the start of the countback task.

A scatter plot of the 45 trials (of all participants) of the countback task illustrates the small increase in heart rate (Fig. 5 left) and decrease in respiration rate (Fig. 5 right). Furthermore, large differences between participants can be seen.

Table 1 Means (standard deviations in parentheses) for the 10 s period before and the 10 s period after the start of the countback (CB) task

Dependent measure	10 s before CB	10 s after CB	p value ($\lvert d_z \rvert$)	Correlation (ρ)
Mean speed (kph)	10.95 (3.1)	10.54 (2.2)	0.380 (0.43)	0.943
Acceleration (m/s^2)	0.48 (0.14)	0.40 (0.15)	0.449 (0.56)	−0.543
Steer speed (°/s)	15.93 (8.7)	15.88 (6.1)	0.992 (0.01)	0.143
Throttle speed (%/s)	6.25 (1.1)	5.75 (2.5)	0.620 (0.24)	0.086
Heart rate (1/min)	77.37 (8.2)	78.22 (8.1)	0.053 (1.13)	1.000
Respiration rate (1/min)	18.20 (1.6)	15.25 (3.2)	0.076 (1.00)	0.314
Pupil diameter (mm)	1.99 (0.37)	1.99 (0.50)	0.994 (0.00)	0.900
Gaze navigation (%)	2.91 (1.16)	2.99 (1.40)	0.791 (0.17)	1.000
Gaze road center (%)	53.0 (15.5)	52.5 (17. 3)	0.920 (0.06)	0.800

p values from the dependent t test (effect size in parentheses) and correlations (Spearman ρ, $N = 6$, $N = 4$ for the gaze measures) between the before and after segments are shown. Effect sizes were determined as Cohen's $d_z = t/N^{0.5}$

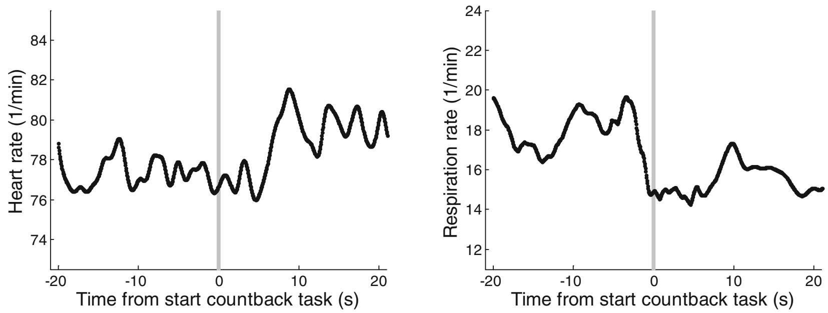


Fig. 4 Mean heart rate (*left*) and respiration rate (*right*) before and after the start of the countback task. Means were computed by averaging across all trials. The start of the countback task is indicated by the *vertical line* at 0 s. The countback task lasted approximately 10 s. Note that due to the manual annotation, the starting time slightly varied across trials

3.3 Evaluation of the Driver Workload Estimator

Table 2 shows the means and standard deviations of the dependent measures per estimated workload level. The driver workload was estimated to be either level 3 or level 4 for over 80 % of the total time. As can be seen by the low mean speed, the first level of the workload estimator was related to low speeds or the vehicle standing still. The fifth workload level occurred less than 2 % of the total time and was related to strong vehicle accelerations, indicated by the throttle speed and the acceleration. The missing values for workload levels 2 and 6 can be explained by

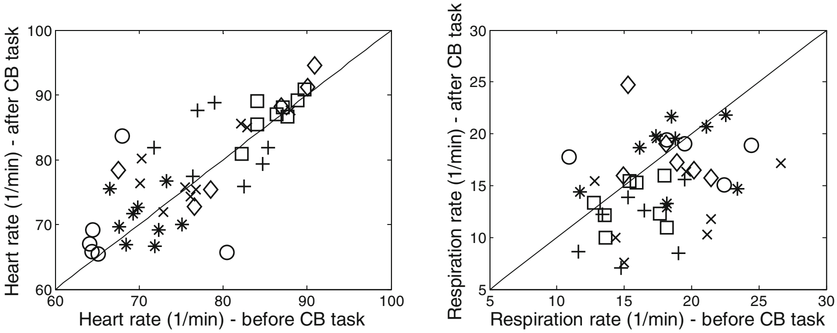


Fig. 5 Scatter plot of the mean heart rate (*left*) and the mean respiration rate (*right*) for the period 10 s before and the period 10 s after the start of the countback (*CB*) task. Each participant is indicated by a different marker

Table 2 Means (standard deviations in parentheses) of the dependent measures for the different levels of the estimated driver workload (N = 6, N = 4 for the gaze measures)

Dependent measure	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Fraction of time (%)	6.5 (3.2)	0 (0)	42.6 (11)	40.8 (9.1)	1.9 (1.5)	0 (0)
Mean speed (kph)	18.8 (23.0)	0 (0)	49.7 (3.1)	34.3 (3.5)	35.6 (7.0)	0 (0)
Acceleration (m/s^2)	0.39 (0.15)	0 (0)	0.45 (0.04)	0.49 (0.05)	0.62 (0.14)	0 (0)
Steer speed (°/s)	5.9 (2.5)	0 (0)	9.4 (1.8)	15.1 (2.4)	13.6 (5.5)	0 (0)
Throttle speed (%/s)	7.0 (3.5)	0 (0)	5.7 (0.9)	6.8 (0.5)	8.3 (4.0)	0 (0)
Heart rate (1/min)	76.8 (7.7)	0 (0)	77.0 (7.6)	77.8 (7.1)	79.3 (8.7)	0 (0)
Respiration rate (1/min)	17.9 (3.6)	0 (0)	17.5 (2.1)	18.2 (2.2)	18.5 (2.8)	0 (0)
Pupil diameter (mm)	2.22 (0.54)	0 (0)	2.22 (0.36)	2.27 (0.44)	2.34 (0.5)	0 (0)
Gaze navigation (%)	2.5 (1.5)	0 (0)	3.0 (1.3)	2.8 (1.4)	5.6 (4)	0 (0)
Gaze road center (%)	45.7 (17.9)	0 (0)	54.6 (14.5)	52.6 (11.8)	49.5 (9.4)	0 (0)
RSME (0–15)	3.8 (1.4)	0 (0)	4.3 (1.4)	4.5 (1.7)	4.4 (1.7)	0 (0)

the absence of criteria for the estimator to estimate these levels within the current experimental scenarios.

Several dependent measures showed an increase from level 3 to level 4. Figure 6 shows the effects between level 3 and level 4 for the heart rate (left), respiration rate (middle), and RSME reports (right). It can be seen that individual differences were large relative to the difference between level 3 and level 4.

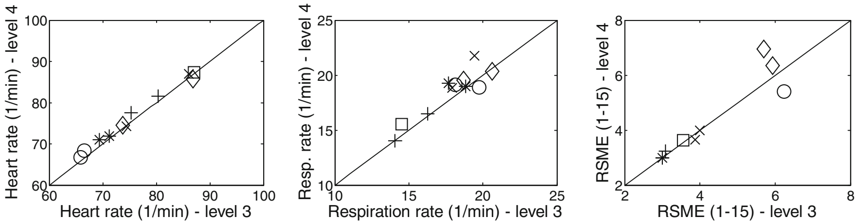


Fig. 6 Scatter plot of heart rate ($N = 11$), respiration rate ($N = 11$), and Rating Scale Mental Effort (RSME; $N = 9$) between level 3 and 4 of the driver workload estimator. Markers are session-averaged values per participant. Each participant is indicated by a different marker

4 Discussion

In this paper, we described the methods and results of an on-road experiment including an online driver workload estimator. Consistent with results from Verwey [4], the results suggest that driver workload is location-dependent. Averaged across participants, the RSME values were high at specific locations in our experimental route. This is further illustrated by the steering activity and gaze behavior along the route, ranging from low steering activity and a higher percentage of gaze directed to the road center on the highway to high steering activity and a lower percentage of gaze directed to the road center in the urban area.

The percentage of gaze directed at the road center tended to increase with increasing driving speeds, whereas the steering activity decreased (i.e., steering steadiness increased) with increasing driving speeds. These results are similar to results found in driving simulator studies [23, 24, 27], and illustrates the relative validity of the measurements obtained with our experimental setup.

Consistent with the literature, the secondary arithmetic task resulted in an elevated physiological response. Specifically, the secondary task resulted in increased heart rate, a finding consistent with Reimer [28] who found similar results when participants performed an n-back arithmetic task (see also [29]). Our results also illustrate that the heart rate response was relatively slow (Fig. 4) [24]. The respiration rate responded quickly to the elevated cognitive load as the participants initiated the countback task. However, this response may be caused by the nature of our secondary task; literature has shown a reduction of respiration rate as participants engage in speech tasks [30]. No substantial effects of the secondary task were found on the control activity of the participants. This finding may be explained by the small cognitive impact of the secondary task as compared to the complex driving task.

Our driver workload estimator estimated the workload to be at intermediate levels (levels 3 or 4 on the 6-point scale) for more than 80 % of the time. Trends were observed between the workload estimate and the RSME results, heart rate, and respiration rate. However, further research into the workload estimator is recommended. Considering the fact that individual differences are large, particular attention is needed to creating person-specific baseline values.

Conducting experiments in a complex naturalistic environment entails several limitations. Because of the exploratory nature of this research, our small sample size does not allow firm conclusions. The naturalistic environment has strong ecological validity, but also introduces side effects (e.g., weather conditions, varying traffic, and vibrations). These effects not only influence experimental control, but also influence the quality of the physiological data. For example, we found no significant effect of the arithmetic task on pupil diameter, which could be explained by the influence of variable lighting conditions [31].

With this study, a first step has been taken to explore the feasibility of estimating workload in a naturalistic driving environment. Our results correspond to previous findings in driving simulators and in the literature, and demonstrate the validity of the instrumented vehicle for assessing driver workload. The implementation of geo-specific data for driver workload estimation shows promise for application in future adaptive in-vehicle information systems.

Acknowledgments We thank Menno Merts (vehicle and test equipment preparation), Arjan Stuiver, Dick Lenior (test set up/development), and Henny Wilke (test leader) for their support in the research project. The research was supported by the Netherlands Organisation for Scientific Research (NWO) of the Ministry of Education, Culture and Science through the RAAK-PRO project: "ADVICE: Advanced Driver Vehicle Interface in a Complex Environment". RAAK-PRO focusses on the enhancement of applied scientific research by Universities of Applied Sciences, in cooperation with the industry.

References

1. 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. Proc. Natl. Acad. Sci. **113**, 2636–2641 (2016)
2. 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. Technical report, Virginia Tech Transportation Institute (2006)
3. Lee, J.D.: Technology and teen drivers. J. Saf. Res. **38**, 203–213 (2007)
4. Verwey, W.B.: On-line driver workload estimation. Effects of road situation and age on secondary task measures. Ergonomics **43**, 187–209 (2000)
5. Kountouriotis, G.K., Merat, N.: Leading to distraction: Driver distraction, lead car, and road environment. Accid. Anal. Prev. **89**, 22–30 (2016)
6. 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. Technical report, Virginia Tech Transportation Institute (2006)

7. Engström, J., Johansson, E., Östlund, J.: Effects of visual and cognitive load in real and simulated motorway driving. Transp. Res. F Traffic Psychol. Behav. **8**, 97–120 (2005)
8. Santos, J., Merat, N., Mouta, S., Brookhuis, K., De Waard, D.: The interaction between driving and in-vehicle information systems: comparison of results from laboratory, simulator and real-world studies. Transp. Res. F Traffic Psychol. Behav. **8**, 135–146 (2005)
9. Hibberd, D.L., Jamsom, S.L., Carsten, O.M.J.: Mitigating the effects of in-vehicle distractions through the use of the psychological refractory period paradigm. Accid. Anal. Prev. **50**, 1096–1103 (2013)
10. Green, P.: Crashes induced by driver information systems and what can be done to reduce them. In: Conference Proceedings Society of Automotive Engineers, Warrendale, PA (2000)
11. Hancock, P.A., Verwey, W.B.: Fatigue, workload and adaptive driver systems. Accid. Anal. Prev. **29**, 495–506 (1997)
12. Piersma, E.H.: Real time modelling of user workload. In: Quéinnee, Y., Daniellou, F. (eds.) Designing for Everyone, pp. 1547–1549. Taylor & Francis, London (1991)
13. Zhang, Y., Owechko, Y., Zhang, J.: Driver cognitive workload estimation: a data-driven perspective. IEEE Intelligent Transportation Systems, Washington D.C., USA (2004)
14. Verwey, W.B.: How can we prevent overload of the driver? In: Parkes A.M., Franzen S. (eds.) Driving Future Vehicles, pp. 235–244. Taylor & Francis, London (1993)
15. Piechulla, W., Mayser, C., Gehrke, H., König, W.: Reducing drivers' mental workload by means of an adaptive man-machine interface. Transp. Res. F **6**, 233–248 (2003)
16. Brookhuis, K.A., De Waard, D.: Monitoring drivers' mental workload in driving simulators using physiological measures. Accid. Anal. Prev. **42**, 898–903 (2010)
17. Recarte, M.A., Perez, E., Conchillo, A., Nunes, L.M.: Mental workload and visual impairment: differences between pupil, blink, and subjective rating. Span. J. Psychol. **11**, 374–385 (2008)
18. 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)
19. He, J., McCarley, J.S.: Effects of cognitive distraction on lane-keeping performance loss or improvement? In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 1894–1898. Sage Publications (2011)
20. Pauzie, A.: A method to assess the driver mental workload: the driving activity load index (DALI). Intell. Transport Syst. **2**, 315–322 (2008)
21. ISO 8596:2009: Ophthalmic optics—visual acuity testing—standard optotype and its presentation. Geneva, Switzerland (2009)
22. Zijlstra, F.R.H., Van Doorn, L.: Construction of a scale to measure perceived effort. Department of Philosophy and Social Sciences, Delft University of Technology, Delft, Netherlands (1985)
23. Van Leeuwen, P.M., Gómez i Subils, C., Jimenez, A.R., Happee, R., De Winter, J.C.F.: Effects of visual fidelity on curve negotiation, gaze behaviour and simulator discomfort. Ergonomics. **58**, 1347–1364 (2015)
24. Rendon-Velez, E., Van Leeuwen, P.M., Happee, R., Horváth, I., Van der Vegte, W.F., De Winter, J.C.F.: The effects of time pressure on driver performance and physiological activity: a driving simulator study. Transp. Res. F: traffic psychology and behaviour **24**, 169–182 (2016)
25. Beatty, J.: Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychol. Bull. **91**, 276–292 (1982)
26. Van Leeuwen, P.M., Happee, R., De Winter, J.C.F.: Changes of driving performance and gaze behavior of novice drivers during a 30-min simulator-based training. Procedia Manufact. **3**, 3325–3332 (2015)
27. Bartmann, A., Spijkers, W., Hess, M.: Street environment, driving speed and field of vision. In: Gale A.G., Brown I.D. (eds.) Vision in Vehicles III, pp. 381–389 (1991)
28. Reimer, B., Mehler, B.: The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation. Ergonomics **54**, 932–942 (2011)

29. Mehler, B., Reimer, B., Coughlin, J., Dusek, J.: Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. Transp. Res. Rec. J. Transp. Res. Board **2138**, 6–12 (2009)
30. Bernardi, L., Wdowczyk-Szulc, J., Valenti, C., Castoldi, S., Passino, C., Spadacini, G., Sleight, P.: Effects of controlled breathing, mental activity and mental stress with or without verbalization on heart rate variability. J. Am. Coll. Cardiol. **35**, 1462–1469 (2000)
31. Watson, A.B., Yellott, J.I.: A unified formula for light-adapted pupil size. J. Vision **12**, 1–16 (2012)

Impact of Motorcyclists' Travel Behavior on Delay and Level-of-Service at Signalized Intersections in Malaysia

Lee Vien Leong and Jong Hui Lee

Abstract High composition of motorcycles in a mixed traffic situation is common in most Asian countries. In Malaysia, there are approximately 11 million registered motorcycles and the fast growing rate of motorcycle ownership has become a critical issue especially in the safety and management of road traffic system. Improper treatment of motorcycles when signalized intersections are designed will result in additional delay and congestion. In this study, the queuing behavior of motorcycles at signalized intersections was segregated into two categories; namely motorcycles outside flow and motorcycles within traffic flow. Motorcycles outside flow are motorcycles that stopped in front of the stop-line of an approach road and motorcycles travelling alongside other vehicles. Motorcycles within flow are the motorcycles that follow behind other vehicles in the traffic flow. In this study, we aimed to examine the behavior of motorcyclists and their influence on delay and level-of-service of signalized intersections in Malaysia.

Keywords Motorcyclists · Travel behavior · Signalized intersections

1 Introduction

Signalized intersections play an important role in the performance of arterial roads, particularly in urban areas where signalized intersections can be the main source of traffic congestion and traffic accidents [1]. The method currently used to estimate delay in Malaysia in Arahan Teknik (Jalan) 13/87 [2], was based on the method developed by Webster and Cobbe in the United Kingdom in the 1950s and 1960s in which the behavior of motorcyclists was not taken into consideration [3]. More

L.V. Leong (✉) · J.H. Lee
School of Civil Engineering, Universiti Sains Malaysia, Engineering Campus, 14300, Nibong Tebal Penang, Malaysia
e-mail: celeong@usm.my

J.H. Lee
e-mail: ljh104534@student.usm.my

N.A. Stanton et al. (eds.), *Advances in Human Aspects of Transportation*,
Advances in Intelligent Systems and Computing 484,
DOI 10.1007/978-3-319-41682-3_95

recently Malaysian authorities have been using the U.S. Highway Capacity Manual, U.S. HCM 2010 [4] to design and analyze signalized intersections. Due to differences in the road system, urban travel behavior, and the mix of vehicle composition in Malaysia as opposed to dual categories of vehicles, i.e., light and heavy vehicles, in the United States, the application of this manual may not be the best choice for analyzing local traffic conditions in Malaysia. In Malaysia, registered vehicles include passenger cars, motorcycles, buses and medium and heavy lorries, but almost 50 % of the registered vehicles are motorcycles. Significant work has been done to understand the effects of mixed traffic composition on the capacity of signalized intersections, and only more recently; researchers have taken into consideration the uniqueness of motorcycles in the analysis of signalized intersections. Nevertheless, in a study conducted by Cuddon and Odgen [5], they have grouped motorcycles travelling within a lane (not between lanes) with passenger cars, whereas Stokes [6] concluded that motorcycles and bicycles have little effect on saturation flow. Branston and Van Zuylen [7] suggested that unless the proportion of motorcycles or bicycles in the traffic stream is more than 20 %, they have very effect on saturation flow and can be ignored for practical purposes.

In Malaysia, the impact of motorcycles should not be disregarded because they are the main mode of transportation apart from passenger cars. Even though Malaysia has a higher percentage of motorcycles than other western countries, the treatment of motorcycles in Malaysia is simplistic and effective estimates of delay could not be obtained. Improper inclusion of motorcycles on the road, especially at signalized intersections, will lower the level of service of the junction and may also lead to more accidents. Hence, our main objective in this study was to investigate motorcyclists travelling behaviour in the estimation of delay and ultimately the level-of-service.

2 Background

The characteristics of heterogeneous traffic flow are quite different with the homogenous traffic flow. Due to the wide variation in operating and performance characteristic of vehicles, heterogeneous traffic systems tend to operate in a different way as compared to the homogenous traffic system. According to Maini and Khan [8], the discharge characteristic of mixed traffic system is quite different with the homogenous traffic system. They have categorized the heterogeneous vehicles that move through the intersection area into four different motorized vehicle types, which are two-wheelers, three wheelers, cars and buses. Then, the speeds of these heterogeneous vehicles were analyzed in detail. From their study, it clearly showed that the variation of clearing speed for different type of vehicles is low, although two-wheelers and cars have acceleration rate that are 33 and 66 % higher than those of auto-rickshaws and buses, respectively. Therefore, it can be concluded that those high performance vehicles will be limited by those low performance vehicles as they are discharged in a single platoon.

Rongviripranikh and Suppattrakul [9] have investigated the effect of motorcycles on traffic operation at signalized intersections. In their study, the relationship between the number of motorcycles discharged in each signal cycle and the start-up lost time of passenger cars were identified using the discharge headway of vehicles. In addition, the effectiveness of motorcycle queue storage is also examined. Through their research, it was proven that the start-up lost time of passenger cars is found to be directly proportionate to the number of motorcycles waiting in the queue. Motorcycle queue storage is found to increase start-up lost time of the traffic. Besides that, they have also developed passenger car equivalent (pce) of motorcycle at different traffic volume and proportion of motorcycles in traffic stream. They have proven that pce of motorcycles consistently decreases with the share of motorcycle in total traffic.

According to Hsu et al. [10], motorcycles will gather at the front of the traffic stream at intersection and will depart together within a very short time when the traffic signal turns green. This is because motorcycles are able to achieve high initial acceleration within a short duration, but less than that of cars when driving above 40 km/h [11]. As a result, a huge motorcycle wave will be generated once the traffic signal turns green. In addition, motorcycle will also have negative starting delay, mainly due to the tendency of the motorcyclists who will always stop ahead of the stop line while waiting for the green phase [12]. Besides that, the saturation flow of motorcycles is dependent on their queuing behavior near the stop line [13].

Due to the small size of motorcycles, motorcyclists tend to maneuver within the traffic stream and get to the front line of traffic queue while waiting for the green phase. According to the study done by Minh et al. [14], a maneuverability model framework had been proposed for motorcycles to simulate when and where motorcycles maneuver in queues at signalized intersections. Since motorcycles are flexible and may not follow lane discipline as other vehicles do, an adapted definition for a motorcycle's lane has been introduced. A threshold distance was estimated to identify when motorcycles need to maneuver. Also, in order to determine where motorcycles maneuver, the lane selection model which was developed with a multinomial logit model was used in the study. Subsequently, it is showed that the 77.88 % of the observed maneuvers, either staying in the current lane or turning left or right could be modeled correctly by using this proposed model.

In Malaysia, the most common type of motorcycles have a small engine that is less than 250 c.c. Motorcyclists can traverse a signalized intersection by three different ways. In the first group, motorcyclist may use the lateral gaps between larger vehicles to weave in and out of the traffic stream to reach the front of the queue while the signal is red. Due to the large number, as a result, motorcycles most stop beyond the stop line. These motorcycles are grouped as one type. Apart from that, as most approach lanes at signalized intersections in Malaysia are wide enough for motorcyclists to travel alongside other vehicles while they traverse the intersection, they are classify as another group. Finally, as some motorcyclists do follow other vehicles in a structured manner, these motorcycles are categorized as another separate group.

Nevertheless, in this study, we divided motorcyclist's behavior at an intersection into those within the flow and those outside the flow. Motorcycles within the flow follow a first-in-first-out rule, implying that they travel either in front of or behind other vehicles in the traffic stream. Motorcycles outside flow are those that do not follow the first-in-first-out rule. This second category consists of motorcycles stopping in front of the stop line during red signals and those that eventually cross the intersection beside other vehicles in a single approach lane.

For the past few years, researchers have conducted studies on the topic of traffic delay. However, the effects of motorcycles are not taken into consideration in these studies. As a result, there will be a small discrepancy exist between the estimated and observed delay. Therefore, this study aims to investigate the effect of motorcyclist travel behavior in the estimation of delay and ultimately on the determination of level-of-service.

3 Study Methodology

In this study, data on motorcyclists travel behaviour were collected using audio cassette recorder at signalised intersections in Central Business District (CBD) areas where the volume of motorcycles is high. In this study, motorcycles were segregated into two categories:

(a) Motorcycles outside flow
    i. Motorcycles that stopped in front of the stop line.
    ii. Motorcycles travelling alongside other vehicles.
(b) Motorcycles following other vehicles in a structured manner or motorcycles within the traffic flow

Saturation flow is the maximum constant departure rate from the queue during the green period. The saturation flow concept assumes that when the signal changes to green, traffic discharges at a constant rate (saturation flow rate) until either the queue is exhausted or the green period ends. The departure rate is lower during the first few seconds as vehicles accelerate to normal running speed and similarly during the period after the end of green interval as the flow of vehicles declines [15, 16]. The rate of discharge (saturation flow) was assumed not to vary from cycle to cycle [17]. In this study, vehicle discharge patterns during green time were collected with the audio cassette recorder to help determine the saturation flow rate. The term "green time" refers to the "green plus amber" period. Data were collected for an average of 30 signal cycles for each straight-through traffic lane of an approach. By observing vehicle discharge pattern, the number of vehicles discharging from a queue in successive six second intervals can be determined. This method is commonly used to compute saturation flow. In the analysis, only saturated intervals were considered. The flow for each saturated interval except the first and the last interval is averaged

Table 1 Conversion factors to passenger car unit

Vehicle class	Vehicle type	Pce values
Class 1	Passenger cars	1.00
Class 2	Light vans	1.75
Class 3	Medium lorries	1.75
Class 4	Heavy lorries	2.25
Class 5	Buses	2.25
Class 6	Motorcycles	0.33

and the saturation flow calculated [18]. Only motorcycles within flow were considered in the calculation of saturation flow.

In a mixed traffic situation, the proportion and type of vehicles in the traffic stream also must be taken into consideration. Saturation flow usually is measured in vehicles per hour, so weighting factors, or passenger car equivalents (pce), are used for other vehicles to enable saturation flows to be quantified as passenger car units per hour (pcu/h). Based on Arahan Teknik (Jalan) 13/87, the pce values used in the design and analysis of signalized intersections in Malaysia are shown in Table 1.

The average delay per vehicle was then estimated based on Arahan Teknik (Jalan) 13/87 (which is adopted from Webster's delay model) as shown in Eq. (1). However, this equation is not suitable to be used at high degree of saturation as it will greatly estimate delay. In the initial analysis, only motorcycles within flow will be included in the calculation of delay.

$$d = 0.9\left[C(1-\lambda)^2/2(1-\lambda x) + x^2/(2q(1-x))\right] \quad (1)$$

where

- d average delay per vehicle (s)
- C cycle time (s)
- λ proportion of the cycle that is effectively green for the phase under consideration (g/C)
- q flow (pcu/h)
- x degree of saturation, which is the ratio of actual flow to the maximum flow that can pass through the approach (q/(λS))
- S saturation flow (pcu/h)

Subsequently, level-of-service for each of the straight-through lane is determined. According to Arahan Teknik (Jalan) 13/87, the level-of-service for signalized intersections is based on stopped delay but U.S. HCM 2010 uses control delay to determine the level-of-service. The U.S. HCM 2010 assigns LOS F when the degree of saturation is more than 1.0 regardless of the delay values calculated. Table 2 shows the level-of-service for signalized intersection based on stopped delay and control delay.

Table 2 Level-of-service

Level-of-service	Arahan Teknik (Jalan) 13/87	U.S. HCM 2010
	Stopped delay (s)	Control delay (s) (v/c ≤ 1.0)
A	d < 5.0	≤ 10
B	5.1 < d < 15.0	10 < d ≤ 20
C	15.1 < d < 25.0	20 < d ≤ 35
D	25.1 < d < 40.0	35 < d ≤ 55
E	40.1 < d < 60.0	55 < d ≤ 80
F	> 60.0	80 < d

4 Data Collection

In this research, traffic flow data were collected in CBD areas under dry-weather conditions for through traffic with a level gradient in several major Malaysian cities. The through-only lanes reflect an ideal situation with no interference due to cars, taxis or buses picking up and setting down passengers, commercial vehicles loading or unloading goods and no parking activity in the adjacent lane. The sites selected have different geometric conditions such as lane width, approach grade and the position of the straight-through traffic lanes (nearside or non-nearside). The signalized intersections being studied is fully saturated or have adequately saturated portions of the green interval of longer than 20 s [19].

Data were collected by observing vehicle discharge pattern using the audio cassette recorder. An audio cassette recorder was used because the time headway between successive vehicles must be measured and observers otherwise might not have enough time to record the actual time headway for each vehicle passing through the stop line. By using a cassette recorder, events in the observed lane such as the beginning of the green interval, the passage of the rear axle of each passing vehicle over the stop line the vehicle type, the end of saturation flow and the beginning of the amber and red intervals can be noted as they occurred. The number of motorcycles stopped in front of the stop line during red period and the number of motorcycles travelling alongside other vehicles during green interval also were recorded. This entire process is repeated for subsequent green intervals. Data were not collected if vehicles travelled through the intersection during the red interval. The vehicle types distinguished in this study are the same as the vehicle classes shown in Table 1.

Data were collected during peak periods (morning peak: 7.30–9.30 am, afternoon peak: 12.00–2.00 pm or evening peak: 4.30–6.30 pm) on weekdays when traffic flows at the intersections are typical, but saturated. Traffic flow data were collected at 14 signalized intersections in various states throughout Malaysia. For each signalized intersection, traffic flow data were collected simultaneously from several approach lanes that satisfied the predetermined conditions described above. From the 14 signalized intersections, traffic flow data for a total of 27 single lane streams were

Table 3 Summary of data collection sites

States	No. of sites	No. of lanes
Penang	2	2
Perak	2	4
Kuala Lumpur/Selangor	10	21
Total	14	27

collected (Table 3). For any single lane, data were collected for average of 30 signal cycles and saturation flow rates were computed based on the observed vehicle discharge. Subsequently, delay for each lane was estimated using Eq. (1).

5 Results and Discussions

Based on the collected data, majority of the vehicles observed at the signalized intersections are cars and motorcycles. On the average, the traffic flow consists of 60 % cars, 30 % motorcycles, 5 % respectively for lorries and buses and 1 % trailers.

In general, the recorded number of motorcycles following other vehicles are the highest in which from the total number of motorcycles recorded, 48 % of them are motorcycles following other vehicles, followed by 29 % of motorcycles in front of stop line and 22 % of motorcycles beside other vehicles. Figure 1 shows the

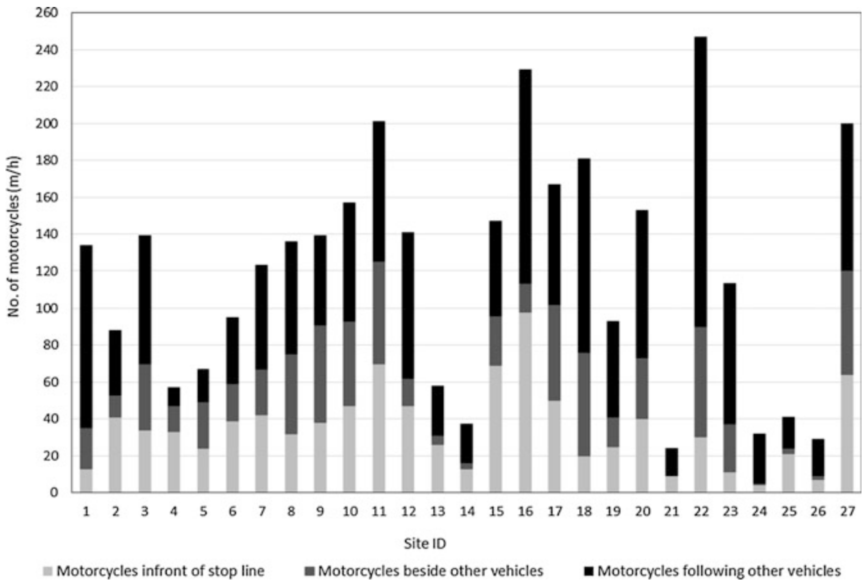


Fig. 1 Observed motorcyclists travel behavior for each of the straight-through lane

observed motorcyclists travel behavior for each of the straight-through lane. Site with ID number 22 recorded the highest total number of motorcyclist and motorcyclist travelling within a traffic lane.

In order to investigate the motorcyclists travelling behavior in the estimation of delay, three different groups of delay were estimated. The first delay group, d1 which is the base group, only takes into consideration motorcycles within the traffic flow while the second group of delay, d2 takes into consideration the total number of motorcycles within a traffic lane, which is the summation of motorcycles inside flow with motorcycles beside other vehicles. Lastly, the third group of delay, d3 will takes into consideration the total number of motorcycles observed for a particular lane regardless of the travel behavior. Table 4 shows the green ratio, degree of saturation flow and delay values estimated for each group. In order to better understand the effect of motorcyclist travel behavior on delay, the percentage of change between the estimated delays were plotted against the corresponding additional number of motorcycles taken into consideration in the delay group. Figures 2, 3 and 4 show the results obtained.

Based on the graphs plotted in Figs. 2, 3 and 4, comparing the two types of behavior when motorcyclists' are travelling outside the flow, the total number of motorcyclist travelling beside other vehicles has more significant impact on the estimation of delay as compared to motorcyclists stopping in front of the stop line. This is based on the strong correlation obtained between the percentage of change in delay estimated in Group 2 (based on the number of motorcycles travelling within a traffic lane) versus Group 1 (delay estimated based on the number of motorcycles following other vehicles), with the number of motorcycles travelling beside other vehicles, as shown in Fig. 2. Based on the regression equation obtained in Fig. 2, for every 17.6 number of motorcyclists travelling beside other vehicles, delay will increase by 1 %.

As the coefficient of correlation is the highest between the percentage of change in delay estimated in Group 2 (based on the number of motorcycles travelling within a traffic lane) versus Group 1 (delay estimated based on the number of motorcycles following other vehicles), with the number of motorcycles travelling beside other vehicles, this indicated that motorcyclists travelling within a traffic lane yield more accurate estimation of delay.

Figure 5 shows the frequency of each level-of-service based on the stopped delay and control delay obtained for each delay group while Table 5 shows the detail of level-of-service determined for each traffic lane based on the stopped delay and control delay of each delay group. Level-of-service is used to classify varying conditions of traffic flow. The highest level, "A" is the flow when drivers are able to travel at their desired speed with freedom to maneuver while the lowest level, "F" represent the worst traffic situation during the congested stop-start conditions. The results obtained indicated that most of the lanes operated at level-of-service "D".

Table 4 Summary of delays, green ratio and degree of saturation flow

Site ID	Name	Green ratio, λ = g/C	Degree of saturation, x = q/(λS)			Delay (s)		
			x1	x2	x3	d1	d2	d3
1	Gama 1	0.3072	0.7484	0.7625	0.7708	35.29	36.17	36.74
2	Jln Ampang 1	0.2992	0.5034	0.5101	0.5330	28.59	28.73	29.22
3	Jln Ampang 2	0.3113	0.5354	0.5580	0.5794	28.97	29.52	30.09
4	Jln Imbi 1	0.2774	0.5287	0.5374	0.5579	35.82	36.03	36.55
5	Jln Imbi 2	0.2880	0.6153	0.6310	0.6462	37.57	38.09	38.63
6	Jln Imbi 3	0.2898	0.5647	0.5775	0.6026	36.02	36.38	37.14
7	Jln Leong Boon Swee 1	0.2789	0.7382	0.7530	0.7778	48.12	49.00	50.66
8	Jln Leong Boon Swee 3	0.2847	0.6020	0.6271	0.6458	42.01	42.82	43.47
9	Jln Pahang 1	0.4802	0.3750	0.3917	0.4037	15.17	15.37	15.53
10	Jln Pahang 2	0.4703	0.5042	0.5204	0.5369	17.65	17.93	18.22
11	Jln Pudu 1	0.1990	0.7419	0.7814	0.8316	49.98	52.82	58.11
12	Jln Pudu 2	0.4376	0.5539	0.5608	0.5824	24.91	25.07	25.62
13	Jln Sultan Iskandar 2-1	0.3719	0.3841	0.3863	0.3978	22.06	22.09	22.26
14	Jln Sultan Iskandar 2-2	0.3225	0.4179	0.4197	0.4275	25.97	26.00	26.13
15	Jln Tuanku Abdul Rahman 1	0.3332	0.4958	0.5107	0.5107	35.35	35.71	35.71
16	Jln Tuanku Abdul Rahman 2	0.3606	0.6312	0.6379	0.6818	36.23	36.44	37.96
17	Jln Tun Perak 1	0.3825	0.6069	0.6300	0.6522	33.83	34.53	35.26
18	Mayang Hyper Store 1-2	0.3821	0.7378	0.7656	0.7756	51.82	53.60	54.30
19	Mayang Hyper Store 1-3	0.2831	0.5422	0.5523	0.5682	55.02	55.35	55.88
20	Medan Tuanku 1-4	0.2910	0.6325	0.6532	0.6783	44.29	45.08	46.13
21	Medan Tuanku 1-8	0.2981	0.2388	0.2388	0.2446	34.56	34.56	34.65
22	Medan Tuanku 2-1	0.3192	0.7788	0.8106	0.8265	48.20	50.70	52.23
23	Medan Tuanku 2-4	0.2930	0.3488	0.3647	0.3714	36.84	37.13	37.26
24	MPSP	0.2646	0.2574	0.2582	0.2600	40.64	40.65	40.68
25	PJ Hilton 2	0.2289	0.5163	0.5190	0.5380	49.50	49.59	50.19
26	PJ Hilton 3	0.2312	0.5900	0.5918	0.5981	51.77	51.84	52.08
27	Wisma Sime Darby 1	0.3768	0.5806	0.6095	0.6425	34.38	35.26	36.38

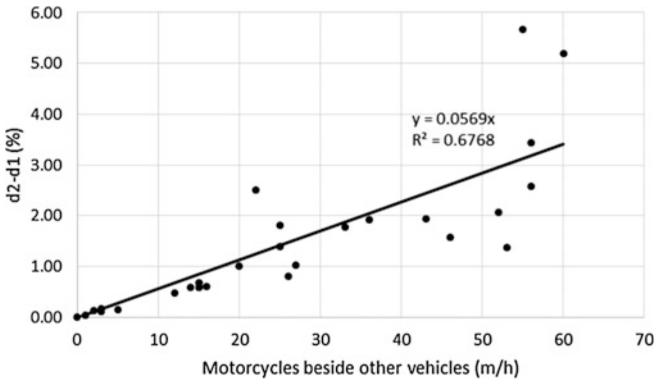


Fig. 2 Percentage of change between the estimated delays due to motorcycles beside other vehicles

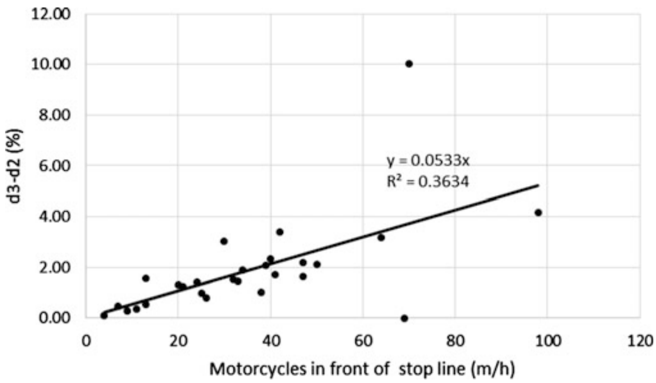


Fig. 3 Percentage of change between the estimated delays due to motorcycles in front of stop line

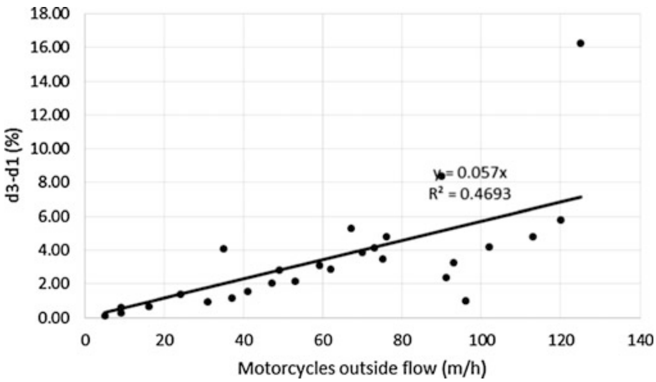


Fig. 4 Percentage of change between the estimated delays due to motorcycles outside flow

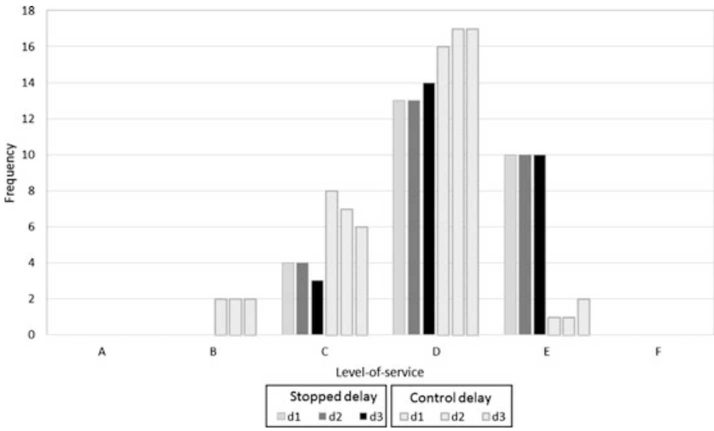


Fig. 5 Frequency for each level-of-service determined based stopped delay and control delay for each delay group

Among the three of groups of delay, almost all of the level-of-service determined based on the delays estimated in Groups 1 and 2 are the same except for one lane which contributes to different level-of-service. Between Group 2 and 3, three lanes contributed to different level-of-service and between Group 1 and 3, four lanes contributed to different level-of-service.

If the level-of-service is determined based on the stopped delay criteria, only level-of-service between "C" to "E" were obtained but if the level-of-service is determined based on the control delay, wider range of level-of-service that is between "B" to "E" were obtained. Additional, higher frequency of level-of-service "D", but lower frequency of level-of-service "E" was obtained based on the control delay criteria as compared to the stopped delay criteria. Hence, level-of-service based on the control delay criteria produces wider range of traffic conditions. Furthermore, the level-of-service based on stopped delay criteria should only be used when the delay is estimated based on the time of a stopped vehicle spends waiting at the intersection in a queue for the signal to turn green. The delay estimated in this study, however is calculated based on the equation given in Arahan Teknik (Jalan) 13/87, which is essentially the Webster's delay model, is the average total delay experienced by vehicles at a signalized intersection controlled by a pre-timed traffic signal. In this case, the control delay criteria are more appropriate to be used to determine level-of-service.

Table 5 Level-of-service determined based stopped delay and control delay for each delay group

Site ID	Name	Delay group 1		Delay group 2		Delay group 3	
		Stopped delay	Control delay	Stopped delay	Control delay	Stopped delay	Control delay
1	Gama 1	D	D	D	D	D	D
2	Jln Ampang 1	D	C	D	C	D	C
3	Jln Ampang 2	D	C	D	C	D	C
4	Jln Imbi 1	D	D	D	D	D	D
5	Jln Imbi 2	D	D	D	D	D	D
6	Jln Imbi 3	D	D	D	D	D	D
7	Jln Leong Boon Swee 1	E	D	E	D	E	D
8	Jln Leong Boon Swee 3	E	D	E	D	E	D
9	Jln Pahang 1	C	B	C	B	C	B
10	Jln Pahang 2	C	B	C	B	C	B
11	Jln Pudu 1	E	D	E	D	E	E
12	Jln Pudu 2	C	C	C	C	D	C
13	Jln Sultan Iskandar 2-1	C	C	C	C	C	C
14	Jln Sultan Iskandar 2-2	D	C	D	C	D	C
15	Jln Tuanku Abdul Rahman 1	D	D	D	D	D	D
16	Jln Tuanku Abdul Rahman 2	D	D	D	D	D	D
17	Jln Tun Perak 1	D	C	D	C	D	D
18	Mayang Hyper Store 1-2	E	D	E	D	E	D
19	Mayang Hyper Store 1-3	E	E	E	E	E	E
20	Medan Tuanku 1-4	E	D	E	D	E	D
21	Medan Tuanku 1-8	D	C	D	C	D	C
22	Medan Tuanku 2-1	E	D	E	D	E	D
23	Medan Tuanku 2-4	D	D	D	D	D	D
24	MPSP	E	D	E	D	E	D
25	PJ Hilton 2	E	D	E	D	E	D
26	PJ Hilton 3	E	D	E	D	E	D
27	Wisma Sime Darby 1	D	C	D	D	D	D

6 Conclusions

Motorcycles travelling through a signalized intersection can be divided into different categories. Different motorcycle travel patterns have different impacts on traffic flow. Generally, when the green ratio is low and the degree of saturation is high; more motorcyclists tend to stop beside and behind other vehicles, as they are blocked by other queuing motorcycles to move to the front of the queue. Once the motorcyclists have filled up the area in front of the queue including the area in front of the stop line, the queue of motorcycles stopping beside other vehicles will build up and more and more motorcyclists will have to stop, either beside or behind other vehicles. Therefore, motorcyclists stopping in front of the stop line generally won't have any significant impact on delay as they tend to move across the intersection even before the drivers can react and accelerate during the onset of green signal. Consequently, inclusion of motorcycles travelling within a traffic lane will yield more realistic estimation of delay and better representation of level-of-service. Clearly, the integration of motorcyclists travel behavior into the estimation of delay is crucial for designing safer, more efficient signalized intersections in Malaysia.

References

1. Minh, C.C., Sano, K.: Analysis of motorcycle effects to saturation flow at signalized intersection in developing countries. J. Eastern Asia Soc. Transp. Stud. **5**, 1211–1222 (2003)
2. Ministry of Works Malaysia: A guide to the design of traffic signals. Arahan Teknik (Jalan) 13/87, Kuala Lumpur (1987)
3. Webster, F.V., Cobbe, B.M.: Traffic signals. Road Research Technical Paper No. 56, Road Research Laboratory, London (1966)
4. Transportation Research Board: Highway capacity manual 2010. Washington D.C. (2010)
5. Cuddon, A.P., Odgen, K.W.: The effect of heavy vehicles on saturation flows at signalised intersections. In: Proceedings 16th ARRB Conference, Part 5, pp. 1–18 (1992)
6. Stokes, R.W.: Some factors affecting signalized intersection capacity. ITE J. pp. 35–40 (1989)
7. Branston, D.M., Van Zuylen, H.J.: The estimation of saturation flow, effective green time and passenger car equivalents at traffic signals by multiple linear regression. Transp. Res. **12**, 47–53 (1978)
8. Maini, P., Khan, S.: Discharge characteristics of heterogeneous traffic at signalized intersections. In: Transportation Research Circular E-C018: 4th International Symposium on Highway Capacity, Maui, Hawaii, pp. 258–270 (2007)
9. Rongviripranikh, T., Suppattrakul, C.: Effects of motorcycles on traffic operations on arterial streets. J. Eastern Asia Soc. Transp. Stud. **6**, 137–146 (2005)
10. Hsu, T.P., Dao, N.X., Ahmad, F.M.S.: A comparative study on motorcycle traffic development of Taiwan, Malaysia and Vietnam. J. Eastern Asia Soc. Transp. Stud. **5**, 179–193 (2003)
11. Wu, J.J.: A study on the mixed traffic flow characteristic and simulation of right-turn PCE. Institute of Civil Engineering, National Taiwan University (1983) (In Chinese)
12. Hsu, T.P.: Capacity analysis and application at signalized intersection. Institute of Civil Engineering, National Taiwan University (1982) (In Chinese)
13. Hsu, T.P.: Integrative analysis of the geometric layout and signal timing for motorcycle traffic at signalized intersection. NSC-85–2211-E002-042, National Research Council (1996) (In Chinese)
14. Minh, C.C., Sano, K., Matsumoto, S.: Maneuvers of motorcycles in queues at signalized intersections. J. Adv. Transp. pp. 39–53 (2010)
15. Teply, S., Jones, A.M.: Saturation flow: do we speak the same language? Transportation Research Record 1320, Transportation Board, Washington, D.C., pp. 144–153 (1991)
16. Akcelik, R.: Traffic signals: capacity and timing analysis. Research Report ARR No. 123, Australian Road Research Board, Melbourne, Australia, pp. 12–22 (1981)

17. Miller, A.J.: The capacity of signalised intersections in Australia. Australian Road Research Bulletin No. 3, Australian Road Research Board, Melbourne, Australia (1968)
18. Road Research Laboratory: A method for measuring saturation flow at traffic signals. Road Note 34/196, Department of Scientific and Industrial Research, Her Majesty's Stationery Office, London, pp. 1–13 (1963)
19. Brown, G.E., Ogden, K.W.: The effect of vehicle category on traffic signal design: a re-examination of through car equivalents. In: Proceedings 14th ARRB Conference, Part 2, pp. 27–37 (1988)