

Advances in Accounting Behavioral Research
Volume 14

Advances in Accounting Behavioral Research

Vicky Arnold
Editor



**ADVANCES IN ACCOUNTING
BEHAVIORAL RESEARCH**

ADVANCES IN ACCOUNTING BEHAVIORAL RESEARCH

Series Editor: Vicky Arnold

Recent Volumes:

Volumes 1–4: Edited by James E. Hunton

Volumes 5–13: Edited by Vicky Arnold

ADVANCES IN ACCOUNTING BEHAVIORAL RESEARCH
VOLUME 14

ADVANCES IN ACCOUNTING BEHAVIORAL RESEARCH

EDITED BY

VICKY ARNOLD

*Ernst & Young Professor of Accounting,
Kenneth G. Dixon School of Accounting,
University of Central Florida, USA*

SENIOR ASSOCIATE EDITOR

DONNA BOBEK

University of Central Florida, USA

ASSOCIATE EDITORS

B. DOUGLAS CLINTON

Northern Illinois University, USA

ANNE LILLIS

University of Melbourne, Australia

ROBIN ROBERTS

University of Central Florida, USA

CHRIS WOLFE

Texas A&M University, USA

SALLY WRIGHT

University of Massachusetts Boston, USA



United Kingdom – North America – Japan
India – Malaysia – China

Emerald Group Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2011

Copyright © 2011 Emerald Group Publishing Limited

Reprints and permission service

Contact: booksandseries@emeraldinsight.com

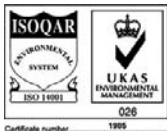
No part of this book may be reproduced, stored in a retrieval system, transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without either the prior written permission of the publisher or a licence permitting restricted copying issued in the UK by The Copyright Licensing Agency and in the USA by The Copyright Clearance Center. No responsibility is accepted for the accuracy of information contained in the text, illustrations or advertisements. The opinions expressed in these chapters are not necessarily those of the Editor or the publisher.

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-1-78052-086-5

ISSN: 1475-1488 (Series)



Emerald Group Publishing Limited, Howard House, Environmental Management System has been certified by ISOQAR to ISO 14001:2004 standards



Awarded in recognition of Emerald's production department's adherence to quality systems and processes when preparing scholarly journals for print



INVESTOR IN PEOPLE

CONTENTS

LIST OF CONTRIBUTORS	vii
REVIEWER ACKNOWLEDGMENTS	ix
EDITORIAL POLICY AND SUBMISSION GUIDELINES	xi
MULTIDIMENSIONAL LOCUS OF CONTROL APPLIED TO THE TAXPAYER/TAX PROFESSIONAL RELATIONSHIP <i>Michelle S. Bertolini, Julia L. Higgs and Karen L. Hooks</i>	1
THE SOCIAL NORMS OF TAX COMPLIANCE: SCALE DEVELOPMENT, SOCIAL DESIRABILITY, AND PRESENTATION EFFECTS <i>Donna D. Bobek, Amy M. Hageman and Charles F. Kelliher</i>	37
QUANTIFYING INTUITIONS ABOUT RISK: COMPARING PUBLIC ACCOUNTING FIRM PARTNERS PERCEIVED AS “RISKY” AND “NON-RISKY” <i>Anne Norris, Deborah Saber, David Morrison, Daven Morrison and Greg Trompeter</i>	67
THE EFFECT OF CONGRUENCE BETWEEN CULTURAL TIME ORIENTATION AND BUDGET PLANNING HORIZON ON EMPLOYEES’ SATISFACTION WITH PARTICIPATIVE BUDGETING <i>Zhihong Wang and James E. Hunton</i>	91

INTEGRATED INFORMATION SYSTEMS AND INTERORGANIZATIONAL PERFORMANCE: THE ROLE OF MANAGEMENT ACCOUNTING SYSTEMS DESIGN <i>Andreas I. Nicolaou</i>	117
ARE ENGAGEMENT QUALITY REVIEWS REALLY OBJECTIVE? <i>Ambrose Jones III, Carolyn Strand Norman and Jacob M. Rose</i>	143
DOES THE CHANGE TO PRINCIPLES-BASED ACCOUNTING INCREASE JUROR ASSESSMENTS OF AUDITOR LIABILITY? <i>John T. Sennetti, Charles P. Becker and Howard J. Lawrence</i>	165
WHAT FACTORS INFLUENCE THE NUMBER OF COAUTHORS IN THE PUBLISHED RESEARCH OF THE MOST PRODUCTIVE AUTHORS IN ACCOUNTING LITERATURE? A LONG-TERM STUDY <i>Robert W. Rutledge, Khondkar E. Karim and Alan Reinstein</i>	191

LIST OF CONTRIBUTORS

<i>Charles P. Becker</i>	Business and Finance, Concord University, Athens, WV, USA
<i>Michelle S. Bertolini</i>	Barney School of Business, University of Hartford, West Hartford, CT, USA
<i>Donna D. Bobek</i>	Kenneth G. Dixon School of Accounting, University of Central Florida, Orlando, FL, USA
<i>Amy M. Hageman</i>	Department of Accounting, Kansas State University, Manhattan, KS, USA
<i>Julia L. Higgs</i>	School of Accounting, Florida Atlantic University, Boca Raton, FL, USA
<i>Karen L. Hooks</i>	School of Accounting, Florida Atlantic University, Boca Raton, FL, USA
<i>James E. Hunton</i>	Department of Accountancy, Bentley University, Waltham, MA, USA
<i>Ambrose Jones III</i>	Bryan School of Business and Economics, University of North Carolina Greensboro, Greensboro, NC, USA
<i>Khondkar E. Karim</i>	College of Management, University of Massachusetts Lowell, Lowell, MA, USA
<i>Charles F. Kelliher</i>	Kenneth G. Dixon School of Accounting, University of Central Florida, Orlando, FL, USA
<i>Howard J. Lawrence</i>	Patterson School of Accountancy, University of Mississippi, Oxford, MS, USA

<i>Daven Morrison</i>	Morrison Associates, Palatine, IL, USA
<i>David Morrison</i>	Morrison Associates, Palatine, IL, USA
<i>Andreas I. Nicolaou</i>	Department of Accounting & MIS, Bowling Green State University, Bowling Green, OH, USA
<i>Carolyn Strand Norman</i>	Department of Accounting, Virginia Commonwealth University, Richmond, VA, USA
<i>Anne Norris</i>	College of Nursing, University of Central Florida, Orlando, FL, USA
<i>Alan Reinstein</i>	School of Business Administration, Wayne State University, Detroit, MI, USA
<i>Jacob M. Rose</i>	Department of Accounting and Finance, University of New Hampshire, Durham, NH, USA
<i>Robert W. Rutledge</i>	McCoy College of Business Administration, Texas State University, San Marcos, TX, USA
<i>Deborah Saber</i>	College of Nursing, University of Central Florida, Orlando, FL, USA
<i>John T. Sennetti</i>	Huizenga School of Business & Entrepreneurship, Nova Southeastern University, Ft. Lauderdale, FL, US
<i>Greg Trompeter</i>	Kenneth G. Dixon School of Accounting, University of Central Florida, Orlando, FL, USA
<i>Zhihong Wang</i>	Department of Accountancy, Bentley University, Waltham, MA, USA

REVIEWER ACKNOWLEDGMENTS

The Editor and Associate Editors at AABR would like to thank the many excellent reviewers who have volunteered their time and expertise to make this an outstanding publication. Publishing quality papers in a timely manner would not be possible without their efforts.

Dennis Bline
Bryant University, USA

Peter Booth
University of Technology Sydney,
Australia

Gary Braun
University of Texas El Paso, USA

Joe Brazel
North Carolina State University,
USA

Wayne Bremser
Villanova University, USA

Richard Brody
University of New Mexico, USA

Siew Chan
Washington State University, USA

Mandy Cheng
University of New South Wales,
Australia

Bryan Church
Georgia State University, USA

Anna Cianci
Drexel University, USA

Paul Coram
University of Melbourne, USA

Ronny Daigle
Sam Houston State University,
USA

Stan Davis
University of Tennessee at
Chattanooga, USA

Andy Dill
University of Central Florida,
USA

Ian Eggleton
Victoria University of Wellington,
New Zealand

Mohamed Elbashir
University of Queensland,
Australia

Craig Emby
Simon Fraser University, Canada

Amy Hageman
Kansas State University, USA

Karen L. Hooks
Florida Atlantic University, USA

Susan Ivancevich
University of North Carolina
Wilmington, USA

Marko Järvenpää
University of Jyväskylä

Lorraine Lee
University of North Carolina
Wilmington, USA

Tina Loras
Auburn University, USA

Linda Matuszewski
Northern Illinois University, USA

Robin Pennington
North Carolina State University,
USA

Jacqueline Reck
University of South Florida, USA

Jeff Reinking
University of Central Florida, USA

Joann Segovia
Winona State University, USA

Mike Shaub
Texas A&M University, USA

Kenneth Smith
Salisbury University, USA

Blair Staley
Bloomsburg University of
Pennsylvania, USA

Constantinos J. Stefanou
Alexander TEI of Thessaloniki,
Greece

Marty Stuebs
Baylor University, USA

Steve Sutton
University of Central Florida,
USA

John Sweeney
Washington State University, USA

Greg Trompeter
Boston College, USA

Sandra Vera-Munoz
University of Notre Dame, USA

Brett Wilkinson
Baylor University, USA

Kim Zahller
University of Central Florida, USA

EDITORIAL POLICY AND SUBMISSION GUIDELINES

Advances in Accounting Behavioral Research (AABR) publishes articles encompassing all areas of accounting that incorporate theory from and contribute new knowledge and understanding to the fields of applied psychology, sociology, management science, and economics. The journal is primarily devoted to original empirical investigations; however, literature review papers, theoretical analyses, and methodological contributions are welcome. AABR is receptive to replication studies, provided they investigate important issues and are concisely written, and is receptive to methodological examinations that can potentially inform future behavioral research. The journal especially welcomes manuscripts that integrate accounting issues with organizational behavior, human judgment/decision making, and cognitive psychology.

Manuscripts will be blind-reviewed by two reviewers and an associate editor. The recommendations of the reviewers and associate editor will be used to determine whether to accept the paper as is, accept the paper with minor revisions, reject the paper or to invite the authors to revise and resubmit the paper.

MANUSCRIPT SUBMISSION

Manuscripts should be forwarded to the editor, Donna Bobek, at DBobek@bus.ucf.edu via e-mail. All text, tables, and figures should be incorporated into a Word document prior to submission. The manuscript should also include a title page containing the name and address of all authors and a concise abstract. Also, include a separate Word document with any experimental materials or survey instruments. If you are unable to submit electronically, please forward the manuscript along with the experimental materials to the following address:

Donna Bobek, Editor
Advances in Accounting Behavioral Research
Kenneth G. Dixon School of Accounting
University of Central Florida

P. O. Box 161400

Orlando, FL 32816-1400, USA

References should follow the APA (American Psychological Association) standard. References should be indicated by giving (in parentheses) the author's name followed by the date of the journal or book; or with the date in parentheses, as in 'suggested by Canada (2005)'.

In the text, use the form Hageman *et al.* (2006) where there are more than two authors, but list all authors in the references. Quotations of more than one line of text from cited works should be indented and citation should include the page number of the quotation; e.g. (Phillips, 2001, p. 56).

Citations for all articles referenced in the text of the manuscript should be shown in alphabetical order in the Reference list at the end of the manuscript. Only articles referenced in the text should be included in the Reference list. Format for references is as follows:

For Journals

Dunn, C. L. & Gerard, G. J. (2001). Auditor efficiency and effectiveness with diagrammatic and linguistic conceptual model representations. *International Journal of Accounting Information Systems*, 2 (3): 1–40.

For Books

Ashton, R. H., & Ashton, A. H. (1995). *Judgment and decision-making research in accounting and auditing*. New York, NY: Cambridge University Press.

For a Thesis

Smedley, G. A. (2001). *The effects of optimization on cognitive skill acquisition from intelligent decision aids*. Unpublished doctoral dissertation, University.

For a Working Paper

Thorne, L., Massey, D. W., & Magnan, M. (2000). *Insights into selection-socialization in the audit profession: An examination of the moral reasoning of public accountants in the United States and Canada*. Working paper. York University, North York, Ontario.

For Papers from Conference Proceedings, Chapters from Book etc.

Messier, W. F. (1995). Research in and development of audit decision aids. In R. H. Ashton & A. H. Ashton (Eds.), *Judgment and decision making in accounting and auditing* (pp. 207–230). New York: Cambridge University Press.

MULTIDIMENSIONAL LOCUS OF CONTROL APPLIED TO THE TAXPAYER/TAX PROFESSIONAL RELATIONSHIP

Michelle S. Bertolini, Julia L. Higgs
and Karen L. Hooks

ABSTRACT

This study seeks to further an understanding of taxpayer characteristics. The study presents a multidimensional tax locus of control (LOC) instrument developed from the starting point of a validated LOC instrument from the health-care field. Data collected using the instrument indicate that older taxpayers are more likely to have an external LOC in tax situations, indicated by a greater propensity to defer decision-making to a tax professional, defined as a “powerful other.” As the U.S. population is aging, this information may be helpful to tax practitioners when advising older clients on tax issues and researchers exploring issues related to aging. An additional finding is that taxpayers with more business exposure are less likely to defer to a tax professional. Gender and education play roles in an individual’s internal tax LOC (TaxLOC) beliefs.

INTRODUCTION

Locus of control (LOC) theory offers an explanation about how individuals perceive the impact of their personal actions on environmental outcomes. LOC theory was originally applied by Rotter (1966) using a generalized measure, which has been utilized in various accounting studies (Brownell, 1981; Frucot & Shearon, 1991; Hyatt & Prawitt, 2001; Tsui & Gul, 1996). Rotter's instrument classified individuals as having either an internal or an external LOC. An "internal" believes that outcomes result from personal action. An "external" believes that outcomes result from things outside an individual's control. The concept of external control has been extended since Rotter's early work and is sometimes described as control attributed to factors labeled fate or chance or powerful others – all ways of explaining how influences apart from the individual impact the outcome. Under Rotter's scale, people may display strong internal or external beliefs or may be positioned along a continuum between internal and external. Other disciplines have developed situation-specific LOC instruments that address particular subjects or environments and measure more descriptive subcategories of Rotter's original dichotomous structure (Y.Y. Chung, 2001; Cleveland, Kalamas, & LaRoche, 2005; Spector, 1988; Wallston, 1976, 1978).

Studies on situation-specific LOC in health care demonstrate that as patients age, some adopt powerful other LOC and increasingly delegate health-related decisions to their physicians, although the situation-specific LOC studies do not address why the differences related to age occur (M.E. Lachman, 1986; M.E. Lachman & Leff, 1989).¹ Studies addressing the increased externality in general LOC of the elderly propose declining health (Brothen & Detzmer, 1983), isolation, and inactivity (Lumpkin, 1985) as causal factors. In the situation-specific health-care studies, powerful other is, in essence, a subclassification of the external LOC dimension; the patients apparently believe their physicians are the controlling influence on their health outcomes and, as a result, defer medical decisions to their physicians. The primary purpose of this study is to explore a tax-related question analogous to the one explored in health care: Does a positive correlation exist between taxpayer age and the likelihood of delegating to a tax professional the decisions that may impact personal tax liability?

To examine the proposition that aging tax clients delegate decision-making to their tax professionals just as aging patients do to their physicians, this study adapts Wallston's (1978) Multidimensional Health Locus of Control (MHLC) instrument to the tax domain. Health-care research is the most useful starting place for this study because health-care situation-specific LOC

has been explored extensively, the MHLC instrument has been validated and adapted for use in multiple studies, and the MHLC is sensitive to the impacts of aging on LOC. Furthermore, the MHLC addresses professionals (physicians) as the powerful other. This study also explores other taxpayer attributes, such as, education, business exposure, gender, and taxing authority interaction to identify other possible influences on tax-specific LOC.

Using Wallston's MHLC as a starting point, the authors revise and adapt the instrument to a tax setting. This study reports the analysis of data collected from study respondents. Participants completed the tax LOC (TaxLOC) instrument, the original Rotter Internal/External (I/E) Scale, and questions on demographic information. Factor analysis is used on the TaxLOC instrument responses to determine item loadings for TaxLOC constructs. Consistent with Wallston (1978), the results produce the dimensions of powerful other and internal. This study's results differ from Wallston (1978) in that they do not find a reliable chance/fate dimension in tax-related circumstances.

A multivariate analysis of variance (MANOVA) indicates that age is positively correlated ($p < 0.01$) with the powerful other dimension of TaxLOC, and business exposure shows a negative correlation ($p < 0.05$) with powerful other. Males are more likely than females to show an internal TaxLOC ($p < 0.05$). Three two-way interactions reflect significant correlation ($p < 0.05$) with the internal dimension. Age and business exposure, age and education, and education and business exposure all impact an internal TaxLOC. Important study results are findings that older taxpayers are more likely to defer decisions to their tax professionals, and taxpayers with more business exposure are less likely to defer decisions.

The contributions of this research include introducing a situation-specific LOC instrument into the accounting discipline and demonstrating that LOC theory may be adapted for accounting-specific use. The study findings suggest that, in a tax context, an aging taxpayer may be more likely to have an external LOC, which increases the possibility that the individual may defer tax decisions to the powerful other, that is, the tax professional. In addition, the analysis identifies other demographic factors such as education and business exposure that moderate the effects of aging on TaxLOC. As in public health and medical research, it is useful to understand in a tax environment how knowledgeable professionals and less informed clients interact because of the power imbalance and information asymmetry that exists. The dynamic between the professional and the client may differ based on the characteristics of the client and impact the manner in which the professional conducts interactions with the client. This research contributes

to an understanding of the relationships between tax professionals and clients. The findings may be particularly useful in light of the aging U.S. population.

BACKGROUND AND RESEARCH QUESTIONS

Locus of Control

Rotter (1966) developed the I/E Scale as an instrument to assess whether an individual has an internal or an external LOC. Early LOC theory proposes that individuals make decisions based on either an internal or an external orientation (Rotter, 1966). The I/E Scale is a continuum; an individual is assigned to a dichotomous category based on where he or she falls on the scale. People with an internal orientation believe they personally have a significant influence on outcomes (Hafer & Gresham, 2008; Rotter, 1966). Those with an external LOC orientation believe they have less influence on their environments and are more likely to attribute outcomes to fate, chance, or the influence of others (Rotter, 1966; Wallston, 1978). Research indicates that individuals tend to become more internally oriented through middle age and then often move toward an external LOC orientation (Aldwin & Gilman, 2004; Hale & Cochran, 1986; Lumpkin, 1986).

The primary source of discussions of LOC of different age groups is studies using cross-sectional designs. Even so, based on these studies, a seemingly widely accepted inference exists that an individual's LOC changes with age. Generally, cross-sectional studies indicate that schoolchildren tend to exhibit a more internal LOC at older ages (Nowicki & Strickland, 1973; Payne & Payne, 1989; Penk, 1969; Young & Shorr, 1986). Studies have shown increased measures of internality from college through adulthood (Doherty & Baldwin, 1983), with, perhaps, some level of relatively consistent LOC being attained after high school and by middle age (Jensen, Olsen, & Hughes, 1990). In these studies, the positive association between increased internality and age (at least through middle age) is often discussed in light of the increased competency individuals achieve through growth, education, and life experiences. All of the studies identified in this study that report on LOC of the elderly are consistent in reporting the greatest levels of externality in the elderly.

Within the accounting literature, most LOC research utilizes Rotter's (1966) generalized internal/external framework to measure characteristics of accountants and managers. For example, the audit literature has used LOC

to address auditor behavior (Donnelly, Quirin, & O'Bryan, 2003; Hyatt & Prawitt, 2001; Tsui & Gul, 1996). Donnelly et al. (2003), using the Spector's (1988) work LOC scale, explored whether an individual's LOC affected that individual's reaction to dysfunctional audit behavior, defined as premature sign-off on an audit step, underreporting time, and altering or replacing audit procedures. Auditors who are more accepting of dysfunctional audit behavior are more likely to have an external orientation, and those same auditors had higher turnover intentions. Hyatt and Prawitt (2001), using Rotter's I/E Scale, sought to examine whether LOC and audit structure affect an individual auditor's job performance. The study defined audit structure as policies, tools, and so on to transform audit work and judgment into an audit opinion. The results showed auditors with an internal LOC performed at a higher level in unstructured firms and externals performed at a higher level in structured firms. Tsui and Gul (1996) investigated the interaction between an individual's LOC and ethical reasoning. The study used the Rotter's (1966) LOC instrument and Rest's (1979) Defining Issues Test to measure cognitive moral development skills. Using a case study in which subjects had to resolve an audit conflict involving unrecorded liabilities, the authors found that ethical reasoning moderates LOC and an auditor's willingness to accede to a client's request.

Accounting research has also explored budget participation (Brownell, 1981; Frucot & Shearon, 1991), using LOC as a personality characteristic. Brownell (1981) used Rotter's I/E Scale as a personality variable and moderator for the relationship between budgetary participation and managerial performance. Budgetary participation was shown to have a positive effect on individuals with an internal LOC and an opposite effect on external individuals. Frucot and Shearon (1991) used the La Rosa (1986) LOC instrument to study Mexican managers and their performance and participation in the budgetary process. Although the La Rosa instrument breaks LOC down into five dimensions, the researchers grouped the dimensions to arrive at a basic internal/external LOC structure. The research found results similar to Brownell (1981) with some differences related to cultural issues.

In the tax area, Roberts (1995) sought to identify factors that might influence an Internal Revenue Service (IRS) auditor's assertion that a taxpayer was negligent in preparing his or her tax return or a tax return position. The study used Rotter's LOC instrument (1966) and found no correlation between an auditor's LOC and an assertion of negligence against a taxpayer.

In contrast to the application of LOC in accounting, researchers in other fields have gone further and developed situation- or area-specific LOC instruments. Table 1 displays studies in other disciplines that have developed

Table 1. Select Examples of Nongeneralized LOC Instruments.

Author	Discipline	Locus of Control Instrument	Locus of Control Constructs
Chung (2001)	Management	Sales Work Locus of Control	Internal External: Chance and powerful other
Cleveland et al. (2005)	Environment	Environmental Locus of Control	Internal: Economic motivation and individual recycling efforts External: Biospheric altruism and corporate skepticism
Furnham (1986)	Economic behavior	Economic Locus of Control	Internal External: Chance and powerful other
Lefcourt, Von Baeyer, Ware, and Cox (1979)	Social psychology	Multidimensional-Multiattributinal Causality Scale	Internal: Task-Internal and Person-Internal External: Task-External and Person-External
Levenson (1974)	Human behavior	Levenson Three-Scale Locus of Control	Internal External: Chance and powerful other
Spector (1988)	Organizational psychology	Work Locus of Control	Summation of the following factors: Rotter's LOC, social desirability, job satisfaction, organizational commitment, autonomy, perceived influence, leadership, role stress, job tenure and job shifting potential
Wallston (1978)	Health care	Multidimensional Health Locus of Control	Internal External: Chance and powerful other

a multidimensional LOC instrument for a situation-specific area of study. As apparent from Table 1, many of the situation-specific multidimensional LOC instruments arrive at internal and external constructs, with more than one dimension in the external construct. The most frequently found dimensions of the external construct are chance and powerful other.

Duffy, Shiflett, and Downey (1977) noted increasing evidence of a multidimensional aspect of LOC. Levenson (1981) discussed and differentiated the two-dimensional internal/external LOC structure and the multidimensional three-factor structure of internal, powerful other, and

chance, noting the advantages of the three-factor structure in understanding human psychology. The complex structure is assessed to be valid, and the dimensions of powerful other and chance are viewed to offer additional insight into human behavior.

This study relies on Wallston (1978). As given in Table 1, Wallston (1978) proposed and found three dimensions in a health-care context. Wallston defined people with a powerful other LOC as those who turn over health-care decisions to their physicians, that is, the powerful other. People with Wallston's internal orientation are more likely to make personal health-care decisions independently. Wallston also identified chance as a dimension indicating that some individuals perceive luck, God, or other uncontrollable factors to have a great influence on health-care outcomes. Individuals typically display aspects of all three dimensions, but score relatively higher in one of the categories. Unlike Rotter's I/E Scale, where an individual falls somewhere on a continuum, the Wallston's (1978) MHLC scores each individual on each of the three dimensions.

M.E. Lachman (1986) used the MHLC and showed that LOC specific to the physician/patient relationship, in older relative to younger individuals, tends to be more oriented to the powerful other. Lachman used both Wallston's MHLC instrument (to determine the individual's situation-specific LOC) and Rotter's I/E Scale (to assess generalized LOC). Lachman's findings suggest that no matter an individual's generalized LOC, older individuals show an increased powerful other dimension of health-related LOC. These findings, and the conclusion that individuals change over time, were further validated when M.E. Lachman and Leff (1989) performed similar analysis in a five-year longitudinal study.

In the health-care field, the MHLC has been used to research various aspects of health, behavior, and relationships. In addition to studying physician/patient relationships, the MHLC has been used to explore expectancies of control over health care (Smith et al., 1988; Wallston & Wallston, 1982), biofeedback (Stein & Wallston, 1983), and coping (Masters & Wallston, 2005). Thus, the MHLC is a well-validated and widely used instrument in the health-care field.

Experimental Tax Research

Accounting research investigating tax-related characteristics and behavior is limited. No research to date, of which the authors are aware, uses LOC theory to explore the taxpayer characteristics that may affect taxpayers'

relationships with their tax professionals. One stream of recent research has focused on behaviors and characteristics of tax professionals that may be influenced by tax clients. [Cloyd and Spilker \(1999\)](#) investigated whether tax professionals exhibit confirmation bias in favor of the client's preferred outcome when performing tax research. The study found a positive confirmation bias sufficient to allow inaccurate assessments of the legal position. [Kadous, Magro, and Spilker \(2008\)](#) examined whether high practice risk mitigates the confirmation bias found by [Cloyd and Spilker \(1999\)](#). High practice risk was found to mitigate confirmation bias in information search among tax professionals. In related research, [Kadous and Magro \(2001\)](#) conclude that tax professionals are more likely to take an aggressive position with a low-risk than with a high-risk client, indicating that tax professionals do not evaluate information objectively.

[Stephenson \(2007\)](#) showed that tax preparers assessed their level of client advocacy at a level higher than that of their clients' expectation. This shows a potential for the tax preparer taking a position that exceeds the client's risk tolerance. [Bobek et al. \(2010\)](#) investigated whether client risk and client importance influence a tax professional's advocacy position. The results showed that tax professionals exhibited lower levels of client-specific advocacy for riskier clients and that client importance played a part in offering tax-minimizing positions. This stream of research is relevant to the current study because it indicates that tax professionals may change their behavior in response to client characteristics.

Assessing the interaction between tax professional and taxpayer, [Stephenson \(2007\)](#) found that tax professionals and their clients differed in their views of risk, tax minimization, and accuracy. Tax professionals sought to minimize taxes, thus increasing potential audit risk for the taxpayer. In contrast, clients sought accuracy over tax minimization. [Christensen \(1992\)](#) investigated taxpayers' perceptions of their tax professionals regarding satisfaction and the importance and quality of tax services. Clients were satisfied when expectations were met or exceeded. This stream of research suggests that tax professionals need to understand their clients' preferences regarding tax strategy and decisions.

Experimental research involving taxpayers has also focused on compliance and the regimens that affect compliance ([Carnes & Englebrecht, 1995](#); [J. Chung & Trivedi, 2003](#); [Collins, Milliron, & Toy, 1992](#)). Research by [Schisler and Galbreath \(2000\)](#) relied on attribution theory to explore taxpayers' views on responsibility for tax outcomes. Consistent with attribution theory, the research found that when returns are selected for audit by the IRS, likely considered a negative outcome, taxpayers view the tax

professionals as responsible. However, when the returns are not audited, a positive result, taxpayers attribute responsibility to themselves. Exploring taxpayer preferences regarding tax professionals, Sakurai and Braithwaite (2003) found a majority of taxpayers sought an honest, low-risk, down-to-earth tax practitioner.

Taxpayer Characteristics Research

A number of studies have dealt with taxpayers and their characteristics in relation to taxpayers' use of tax professionals and the cost of tax compliance (Arena, O'Hare, & Stabrianos, 2002; Christian, Gupta, & Lin, 1993; Christian, Gupta, Weber, & Willis, 1994; Dubin, Graetz, Udell, & Wilde, 1992; Long & Caudill, 1987; Slemrod & Sorum, 1984). All of these studies used certain demographic and tax return complexity issues to assess the likelihood of the taxpayer using a tax preparer or in determining the cost of compliance. For investigating the use of tax preparers, a majority of the studies used IRS panel data to obtain certain demographic and return complexity information. The authors found that age, marriage, and number of dependents increased the use of tax preparers. In addition, return complexity also played an important role, with self-employed individuals and those with more complex returns more likely to use a tax preparer. Christian et al. (1994) also found a relationship between a taxpayers' prepayment position and the use of a tax preparer. The prepayment position of taxpayers using tax preparers was found to be lower, while the taxpayer's refunds were relatively larger. The authors apply transaction cost theory and posit that tax preparers use the differences in prepayments and refunds to justify fees.

In investigating taxpayers' compliance costs, researchers again look to demographics and return complexity to determine the monetary cost of compliance. Slemrod and Sorum (1984) investigated age, gender, education, income, employment status, occupation, and wage rates along with return complexity to determine costs. The results showed that older individuals were more likely to use a tax preparer, while the inverse held for education. More educated taxpayers were less likely to use a tax preparer. Arena et al. (2002) found similar results, but also found that return complexity led to increased tax preparer usage.

In summary, the research to date addressing relationships between tax professionals and clients shows that these relationships are important when considering quality of service and tax professionals' behavior. In addition, the research shows increased use of tax preparers as individuals' age.

However, the research identified and summarized here does not help to illuminate how tax professionals might best respond based on differences in taxpayer characteristics such as LOC.

Research Question Development

Other disciplines have developed situation-specific LOC instruments. Consequently, the expectation is that one can be used in a tax setting. As accounting research to date has measured LOC only using a generalized instrument, this study relies heavily on research in other disciplines in exploring the situation-specific relationship between tax professionals and clients. As stated previously, health-care research provides substantial theoretical and applied support because it is fairly well developed, has a widely used and validated instrument, addresses the professional/client relationship, and considers age. This study explores how age may affect LOC influences on the tax professional and client relationship. Wallston's MHLC structure and instrument as applied to health care most closely parallels this topic. Specifically, this study assumes that the relationship between a physician and a patient is parallel to a tax professional and client and that the MHLC can be adapted for use in a tax setting. Both physicians and tax professionals have specific knowledge not known to the general public (Johnson, 1977). Therefore, the lay public uses the services of both physicians and tax professionals. These parallels lead to the first research question.

RQ1. *Does a multidimensional locus of control belief structure exist related to taxes that parallels the established and documented MHLC?*

As reported in the health-care literature, M.E. Lachman (1986) and M.E. Lachman and Leff (1989) showed that aging has a profound effect on the relationship between a physician and a patient. As individuals age, they become more external in their LOC belief structures and are more likely to defer medical decisions to the powerful other, the physician. A similar situation may exist between taxpayers and tax professionals. Using an adaptation of the MHLC allows this study to be guided by health-care research exploring the impact of aging on LOC. Tax professionals, like physicians, have specialized knowledge and clients seek advice from these professionals. The effect of client aging on the professional/client relationship found in the health-care setting may also manifest in the tax professional/taxpayer relationship, leading to the second research question:

RQ₂. *Do older taxpayers have a more situation-specific powerful other TaxLOC about their relationships with their tax professionals?*

From an exploratory perspective, this study questions whether, in addition to the LOC variables addressed in health research, other demographic characteristics may influence TaxLOC. These include business exposure, education, college business courses, primary relationship status, taxing authority experience, and gender. The expectation is that with respect to financial matters and more specifically tax matters, beliefs may be shaped by personal characteristics and experiences. On the basis of this, greater business exposure, general education, and number of college business courses may represent characteristics likely to result in a more internal TaxLOC. Additionally, gender might influence TaxLOC. On the basis of prior LOC research (Doherty & Baldwin 1985; Linder 1986; Reed, Kratchman, & Strawser, 1994; Semykina & Linz, 2007), males are expected to be more internally focused than females in their tax situation-specific LOC. Along with gender, primary relationship status may impact TaxLOC, but there is no research support regarding directionality. History with a taxing authority may also influence TaxLOC, although, again, given the lack of prior research, there is no expectation regarding the direction of the influence.

RQ₃. *Do other demographic variables affect the dimensions of an individual's situation-specific LOC relationship with a tax professional?*

In addition to the three research questions, interactions of the variables are also of interest.

METHOD

Instrument Development and Confirmation

This study began with development of the multidimensional TaxLOC instrument (TaxLOC).² As Rotter's I/E Scale deals with a generalized LOC, specialized LOC instruments from other disciplines that most closely align with this study's area of interest, as given in Table 1, were explored for use. As stated previously, those of greatest interest address the relationship between a professional and the individual receiving the professional service. The MHLC was selected as an appropriate starting point based on its extensive use as a well-validated situation-specific LOC instrument.³ An important characteristic that influenced the selection of the MHLC for adaptation, rather than

any other possible choices, is that in health-care studies, the MHLC was sensitive to age as an independent variable predicting LOC.

Wallston's MHLC instrument was appropriate for use in this study because it explores the unique relationship of a health-care provider and patient. The patient has to place a certain amount of trust in the health-care provider because of the knowledge asymmetry that exists between the provider and the patient. In a parallel context, this trust must also exist between a tax professional and his or her client. Other multidimensional LOC instruments do not look at this type of relationship but instead focus more on subjects' LOC in different situations.

In modifying the MHLC for tax situations, changes were limited as much as possible. The health-related terms "sick," "health," "well-being," and "ill" were replaced with the term "tax outcome." The MHLC items were phrased to indicate positive or negative health outcomes by use of words such as sick or health. To be consistent with this approach, the adjectives "positive" or "negative" were added to "tax outcome." Given the possible variability in individual interpretations of what may be a positive or negative tax outcome, the instrument leaves it to the respondents to define positive or negative. However, based on prior research findings that taxpayers desire an accurate tax return and attribute being selected for an IRS audit as a negative outcome resulting from tax preparer actions (Schisler & Galbreath, 2000; Stephenson, 2007), the following are offered in the instrument instructions as possible examples of positive outcomes: avoiding interaction with the IRS, filing a legitimate return on a timely basis with honest information, and minimizing your tax liability. In addition, the following possible examples of negative outcomes are provided: being audited by the IRS, filing a return based on uncertain information with errors or filing late, and paying more taxes than is legally required.

Another modification made to the MHLC for this study was replacing all health professional terms with "tax professional." As with the adjectives positive and negative, discussed previously, the instrument does not define tax professional but leaves interpretation to the participant. The resulting TaxLOC instrument includes the same 18 items as the MHLC. The TaxLOC instrument identifies the items with the same item identification numbers as the MHLC. The TaxLOC instrument is shown in the [appendix](#).

A series of steps were performed to assure that this study's adaptation of the original instrument did not change its usefulness and validity. The steps were intended to confirm that the revised instrument functions similarly with the validated MHLC instrument. Straub (1989) note that, at least in the MIS field, many accounting researchers who develop instruments do not adequately address instrument validation. Their criticism extends to both

the development of new instruments and the adaptations of previously validated instruments, as in this study. Straub (1989) present a four-phase process to assure content and construct validity and reliability that is used in this study for guidance in confirming that the adaptation of the MHLC retains its validity from its previous health-care-specific form. On the basis of the Straub and Carlson steps, the study conclusion is that the adapted TaxLOC instrument continues to be understandable and to measure situation-specific LOC constructs.

Straub and Carlson's phase one consists of pretesting the instrument by personally interviewing participants, rather than having them complete the instrument independently. This step allows for revisions before the instrument is exposed to more participants. Phase two seeks to confirm the instrument across subjects through a comparison of personal interviews versus pen and paper completion of the instrument. On the basis of the extensive use and validation of the MHLC instrument (Wallston, 2005), this study used a combined step for Straub and Carlson's phases one and two. The pretest was carried out with two individuals, ages 45 and 76, who completed the TaxLOC. One of the researchers then interviewed each individual regarding interpretation of the questions. On the basis of the pretest participants' feedback, no changes to the instrument were required and the process moved on to Straub and Carlson's phase three.

Phase three consists of a pilot test. Ten graduate students took the survey during a regularly scheduled class. Participation was voluntary, although given the method it was not anonymous. A Ph.D. student not associated with the class administered the pilot. The Ph.D. student was instructed to report to the researchers on questions regarding the instrument but to keep student identities confidential. None of the students had questions regarding the instructions or instrument. A factor analysis was then performed on the pilot responses of the TaxLOC instrument. The items loaded on four distinct factors.⁴ None of the items from the pilot results loaded on more than one factor. This result is consistent with the MHLC constructs.

The final phase in confirming the appropriateness of the TaxLOC instrument was performing a factor analysis on results from the 18-item TaxLOC instrument of the full sample. (For factor structure analysis, 99 survey responses were used, although not all of these were subsequently analyzed using MANOVA, as explained later.) The instrument packets were distributed to three distinct sample groups, described in the following section on Data Collection. Using a rotated factor analysis (Equamax with Kaiser normalization rotation method), the full sample results also loaded on four factors. However, the Cronbach's alpha on two of the factors was below 0.5.⁵

TaxLOC instrument items with low loadings were removed, that is, those with a factor loading below 0.50, based on the expectation that any loading over 0.50 is important (Hair, Anderson, Tatham, & Black, 1992). Eight items were deleted.⁶ Using the revised data, a second factor analysis (Equamax with Kaiser normalization rotation method) produced two important factors, powerful other (POTLOC) and internal (ITLOC). Cronbach's alpha is 0.835 for POTLOC and 0.813 for ITLOC. A Cronbach's alpha of over 0.60 or greater is considered sufficient for exploratory research, such as this study; 0.80 or more is expected for established research streams (Hair et al., 1992). Another indicator that the two-factor structure identified is appropriate for this study's data is that the eigenvalue for the potential third factor is 0.831, which is below the typical standard for inclusion (Hair et al., 1992).⁷

Wallston's health-care data produced a three-factor LOC structure (internal, fate or chance, powerful other), and this study produced a two-factor TaxLOC structure (internal and powerful other). The difference may exist because Wallston explored health and medicine in which natural law is a controlling element. In the health arena, uncontrollable outcomes may be determined by randomness or natural law. For tax situations, this natural science-related randomness does not exist. An individual's tax position is determined by laws and regulations produced by society and enforced by the government. Table 2 presents the TaxLOC factor structure and loadings for the questions that were used to measure ITLOC and POTLOC. Five questions were used to measure each factor, with the score for each question ranging from 1 to 6 giving a minimum score of 5 and a maximum score of 30 for each factor.

The final instrument package used in the study included the TaxLOC instrument, Rotter's I/E Scale, and a demographic questionnaire.⁸ The Rotter I/E Scale used was the original, well-validated 29-item instrument (Rotter, 1966). The demographic questionnaire collected information on age, education, business exposure, relationship status, taxing authority experience, and gender.

Data Collection

Survey instruments were distributed to three distinct groups. The first group consists of tax clients of small public accounting firms located in the southeastern United States. The second group consists of graduate accounting students from a southeastern state university. The third group consists of adults, mostly older members of the following: a country club, an investment group, a computer club, and a church. These various groups were used to gain

Table 2. Multidimensional Tax Locus of Control Adaptation of the Wallston Model.

Item Number ^a	Question	Factor Loading
	Factor: Internal Tax Locus of Control (ITLOC) [Cronbach's alpha: 0.813]	
6	I am directly responsible for my tax outcomes	0.731
8	Whatever goes wrong with my tax situation is my own fault	0.623
12	The tax outcome I experience depends on whether I take care of my tax situation	0.727
13	When I have a negative tax outcome, I know it is because I have not been taking care of my tax situation property	0.509
17	I can pretty much maintain positive tax outcomes by taking good care of my tax situation	0.543
	Factor: Powerful Other Tax Locus of Control (POTLOC) [Cronbach's alpha: 0.835]	
3	If I see an excellent tax professional regularly, I am less likely to have a negative tax outcome	0.667
5	I can only maintain my tax outcomes by consulting tax professionals	0.576
10	Tax professionals keep me having positive tax outcomes	0.593
14	The type of assistance I receive from other people is what is responsible for how well I recover from a negative tax situation	0.713
18	Following my tax professional's guidance to the letter is the best way for me to maintain a positive tax outcome	0.831

^aSee the Tax Locus of Control instrument in the appendix.

access to individuals with different population demographics. A contact-person approach was used to distribute the data collection materials. The data collection materials included the instrument package, instructions, and a self-addressed stamped envelope to return the materials to the researchers. The data collection materials were designed to be anonymous and indistinguishable between the three respondent subgroups.

The small public accounting firms distributed the data collection materials to approximately 150 individual tax clients at the time taxpayers visited offices to leave their information for tax return preparation. The firms advised their clients to take the documents home to complete. The involvement of the accounting firms ended with delivering the packets to clients. Due to the method of delivery and complete anonymity of the responses, the response rate for this subgroup is unknown. Participation of all contact individuals and respondents was voluntary.

One of the researchers distributed the data collection materials to a group of 30 graduate accounting students during a regularly scheduled class period. The explanation was that the students' basic assumption in responding should be that they were seeing a tax professional to complete their returns or were seeking consultation on a specific tax issue.⁹ The students took the instruments home to complete and returned them in self-addressed stamped envelopes, thus protecting the confidentiality of student identities. Participation was voluntary and no credit was awarded for participation. As stated previously, the packets and return envelopes were identical among the three subgroups of participants. Although the exact number of students who received data collection packets is known to be 30, the response rate of the student subgroup is unknown.

The remaining participants received their packets from members of the various organizations who were recruited by the researchers. Due to the method of distribution, the researchers are unsure of the exact number of packets ultimately distributed to this group. Again, participation of the contact individuals and respondents was voluntary, and the anonymous responses were mailed back to the researchers to protect participant confidentiality, and response rate of the subgroup could not be determined.

Analysis

The initial analysis was a MANOVA using the following as independent variables: age, education level, business exposure, taxing authority experience, number of business college courses taken, Rotter's I/E Scale result, and gender.¹⁰ The dependent variables were factors from the TaxLOC instrument: POTLOC and ITLOC. All of the responses to survey items of each factor were summed to produce the value of each dependent variable. Nonsignificant independent variables were dropped from later analysis. After a second MANOVA analysis of main and interaction variables, post hoc analysis of variance (ANOVA) with Bonferroni correction was conducted.

RESULTS

Descriptive Statistics

A total of 99 fully completed instruments were received; however, the responses were partitioned into two data sets: individuals who (1) have and

(2) have not used a tax professional's services. Although all responses were used for the factor analysis, only those responses indicating prior use of a tax professional's services were analyzed further, resulting in a sample size for the MANOVA of 76. Table 3 presents demographic data of the group of respondents who have used a tax professional's services. The sample includes 40 males and 36 females. The average age was 54.16 years.

To facilitate interpretation of the MANOVA analysis, rather than using age as an integer form continuous variable, respondents were assigned to an age classification based on reported year of birth. The age classifications are as follows: after 1963 (GenX, GenY, and Millennials), between 1944 and 1963 (Baby Boomers), and before 1944 (World War II generation).¹¹

Education level is captured with classifications for high school, associate degree, bachelor's degree, graduate degree, and other. All the individuals indicating "other" had actually completed various types of graduate degrees. The study's participant pool was highly educated with 31 holding

Table 3. Descriptive Statistics and Demographic Information ($n = 76$).

Continuous variables				
Variable	Mean	Minimum	Maximum	Standard Deviation
<i>Age</i>	54.16	22	87	19.621
Number of college-level business courses	6.07	0	42	8.209
Categorical variables				
	Frequency		Frequency	
<i>Age by classification</i>			<i>Education level</i>	
Born after 1963	21		High school	9
Born between 1944 and 1963	26		Associates degree	5
Born before 1944	29		Bachelor degree	31
			Graduate degree	28
<i>Gender</i>			Other	3
Male	40		<i>Primary relationship status</i>	
Female	36		Single	15
			Married	46
<i>Business exposure</i>			Divorced	6
0	15		Widowed	9
1	34		<i>Taxing authority experience</i>	
2	24		None	32
3	3		Some	44

bachelor's degrees and 31 having graduate degrees. The large percentage of highly educated subjects in the sample is not surprising given the method through which participants were contacted and is consistent with the population subset expected to use tax professionals' services.

In addition to basic demographic data, the survey instrument collected information about business exposure, primary relationship status, number of college-level business courses taken, gender, and interactions with taxing authorities. Business exposure was investigated with three questions requiring yes or no responses. One point was assigned for each yes answer, and zero points were assigned for each no answer. The summation of these point values was the business exposure variable.¹² The mean was an aggregate score of 1.2 on the business exposure variable, with the majority of respondents having at least some business exposure or competency. Five taxing authority interaction questions focused primarily on the IRS and include, for example, whether the respondent has received correspondence from or been audited by the IRS. The interaction with taxing authorities variable was assigned a value of one if the individual answered yes to any of five questions posed.¹³ In contrast to the business exposure variable, taxing authority responses were not aggregated because the expectation is that after an individual has any interaction with the IRS, additional interactions should have little or no incremental effect.¹⁴ Forty-four respondents reported interaction with taxing authorities and 32 did not. The reported range on responses to a descriptive data question about the number of college business courses taken is from 0 to 42. The mean number of courses reported was 6.07.

Rotter's (1966) I/E Scale questionnaire was also used for data collection. Rotter's LOC questionnaire contains 29 questions, including 6 distracter questions. Consequently, 23 are actually used to determine an individual's I/E Scale. Using Rotter's scoring system, each question has one answer that is awarded one point while the alternative answer is awarded zero points. With a potential point total ranging from 0 to 23, the scale midpoint is 11.5. Those with a lower (higher) aggregate score are considered more internal (external). The study's sample's mean on the Rotter I/E Scale was 9.18. Thus, the overall sample mean tends toward an internal generalized LOC.

Manova

Independent variables that were significant in the initial MANOVA are the following: age classification, business exposure, education, and gender. As

stated previously, the original exploratory MANOVA included Rotter's I/E Scale and other demographic data as independent variables. Note that Rotter's I/E Scale was not significant for either POTLOC or ITLOC.¹⁵ This suggests that the TaxLOC instrument measures something other than generalized, dichotomous internal and external LOC.

The four significant independent variables were analyzed further. Based on previous research (Lachman, 1986; Lachman & Leff, 1989) that found that age plays a role in an individual's situation-specific LOC when dealing with the physician/patient relationship, age was expected to be important in the taxpayer/tax professional relationship. The expectation was age would be positively associated with a stronger powerful other orientation. The second important variable, business exposure, likely impacts the extent of knowledge asymmetry in the relationship and was therefore expected to be positively related to an internal orientation or to have an inverse relationship with the powerful other orientation. A positive relationship between business exposure and an internal LOC would likely be related to a greater perceived personal efficacy in tax decisions. An inverse relationship to powerful other would likely indicate that more business exposure reduces the information asymmetry between taxpayer and tax professional. Education level is expected to have the same impacts as business exposure. Last, gender was expected to play a role because prior research has shown that males have more internal LOC than females.

A second MANOVA was conducted including the four independent variables previously found to be significant (age classification, business exposure, education, and gender), selected independent variable interaction terms, and the same two dependent variables, POTLOC and ITLOC. Table 4 summarizes the results of the second MANOVA (panel A) and post hoc ANOVAs (panels B and C) performed to determine the causes of significance in the MANOVA.¹⁶ Table 5 displays means for the variables found to be significant.

For example, to determine the cause of significance for the age classification variable, post hoc ANOVAs with Bonferroni adjustment used the two TaxLOC dimensions of powerful other and internal LOC as dependent variables and age classification as the independent variable. Table 4 displays that age classification is significant for the POTLOC dependent variable, but not for the ITLOC dependent variable. Significant differences were found between the younger and the older cohort groups. Table 5, panel A, summarizes the means for POTLOC by age categories and displays the directionality of the difference. The means for powerful other related to those born after 1963, those born between 1944 and 1963, and those born before 1944 were 18.417, 20.929, and 24.023, respectively. The means comparison

Table 4. MANOVA and ANOVA Test Results.

Panel A: MANOVA Results				
Independent variables	Wilk's lambda	df	F-value	p-value
Age classification	0.669	4,58	3.230	0.018
Business exposure	0.624	6,58	2.571	0.028
Gender	0.813	2,29	3.328	0.050
Education	0.804	8,58	0.834	ns
Age classification × gender	0.919	4,58	0.623	ns
Age classification × education	0.580	8,58	2.270	0.035
Business exposure × gender	0.835	4,58	1.366	ns
Business exposure × education	0.502	10,58	2.388	0.019
Age classification × business exposure	0.570	8,58	2.530	0.029

Panel B: ANOVA Results – Dependent Variable: Powerful Other		
Independent variables	F-value	p-value
Age classification	6.799	0.004
Business exposure	4.209	0.013
Gender	0.075	ns
Education	0.623	ns
Age classification × gender	0.302	ns
Age classification × education	0.829	ns
Business exposure × gender	0.859	ns
Business exposure × education	1.156	ns
Age classification × business exposure	2.670	0.051

Panel C: ANOVA Results – Dependent Variable: Internal		
Independent variables	F-value	p-value
Age classification	0.572	ns
Business exposure	1.184	ns
Gender	6.885	0.014
Education	1.111	ns
Age classification × gender	0.901	ns
Age classification × education	4.353	0.007
Business exposure × gender	1.846	ns
Business exposure × education	4.119	0.006
Age classification × business exposure	3.155	0.028

Table 5. Mean Values of the Dependent Variables.

Panel A: Mean Values of Powerful Other LOC by Age Classification				
Age classification				
	Born after 1963	Born 1944–1963	Born before 1944	
Powerful other LOC	18.417	20.929	24.023	

Panel B: Mean Values of Powerful Other LOC by Levels of Business Exposure				
Business exposure				
	Least (0)	(1)	(2)	Most (3)
Powerful other LOC	22.091	20.394	22.940	17.333

Panel C: Mean Values of Internal LOC by Gender		
Gender		
	Males	Female
Internal LOC	23.069	22.250

Panel D: Mean Values of Internal LOC by Age Classification and Education				
Internal LOC				
Age classification				
	Born after 1963	Born 1944–1963	Born before 1944	
Education	High school	–	21.250	23.750
	Associate	23.000	23.000	28.500
	Bachelor	23.026	24.700	22.300
	Graduate	22.833	24.400	22.500
	Other	–	22.000	17.000

Panel E: Mean Values of Internal LOC by Business Exposure and Education					
Internal LOC					
Business Exposure					
		Least (0)	(1)	(2)	Most (3)
Education	High school	23.250	24.250	24.250	–
	Associate	–	24.833	23.000	–
	Bachelor	18.000	23.194	26.600	29.000
	Graduate	21.300	24.800	20.900	22.000
	Other	–	22.000	12.000	–

Table 5. (Continued)

Panel F: Mean Values of Powerful Other LOC by Age Classification and Business Exposure				
Powerful other LOC				
		Age classification		
Business exposure		Born after 1963	Born 1944–1963	Born before 1944
	Least (0)	16.750	25.333	25.000
	1	18.389	17.500	23.344
	2	20.125	22.200	25.933
	Most (3)	–	18.000	16.000

Panel G: Mean Values of Internal LOC by Age Classification and Business Exposure				
Internal LOC				
		Age classification		
Business exposure		Born after 1963	Born 1944–1963	Born before 1944
	Least (0)	19.250	22.333	20.250
	(1)	24.361	22.125	23.688
	(2)	24.500	21.700	22.900
	Most (3)	–	27.000	19.000

Mean values of the dependent variables are based on a range of response from 5 (indicating a low score on the dependent variable) to 30 (indicating a high score on the dependent variable).

indicates that individuals in the older cohorts have, respectively, a greater powerful other orientation, displaying a trend.

As the data are cross-sectional rather than longitudinal, two explanations may be relevant to the age-related differences. First, there may be fundamental differences among these three cohort groups that remain stable across time. An alternative explanation is that people change as they age and exhibit a greater powerful other LOC orientation in tax situations. However, only a longitudinal analysis can definitively support this explanation. The second explanation of a change that occurs with aging is consistent with Lachman's (1989) findings based on longitudinal data, that individuals become more external as they age. A caveat is that Lachman's longitudinal study spanned only five years. However, as stated previously, multiple cross-sectional studies have also found greatest levels of externality in the elderly. As this study is cross-sectional, stating definitively which explanation is appropriate is impossible. The analysis clearly indicates that the respectively older cohorts in

the sample have relatively stronger POTLOC but does not provide information on the reason for the observed difference.

Business exposure also shows significance for the POTLOC dependent variable. The follow-up ANOVA with business exposure as the independent variable shows a statistically significant difference in means for the POTLOC dimension between those who have low levels of business exposure and those who have significant business exposure. Means given in Table 5, panel B, do not show a clear trend but indicate that individuals with the most business exposure display the lowest powerful other orientation (mean of 17.333). The results support the proposition that an individual with more business exposure has a lower POTLOC focus in dealing with his or her tax professional than an individual with less business exposure.

Gender is significant for the ITLOC dependent variable. A comparison of means given in Table 5, panel C, 23.069 for males and 22.250 for females, shows that males are more internal than females. This result is consistent with other LOC literature.

The education variable was not significant on a stand-alone basis for either POTLOC or ITLOC. However, education is significant for the ITLOC dependent variable in the context of interactions with two other variables: age classification and education, and business exposure and education. Note that the main effects significance for age classification and business exposure are for the POTLOC dependent variable. This contrasts with significance for ITLOC for the interaction terms. Regarding the interaction of age classification and education, as can be seen from a comparison of the means mentioned in Table 5, panel D, no clear trend emerges. In the 1944–1963 birth cohort, those with bachelor's and graduate degrees are most internal. In the earlier 1944 birth cohort, those with a high school education or an associate degree are most internal. The authors cannot offer any definitive explanation for the different impacts of education in the three age cohorts. The business exposure and education interaction, while significant for the internal LOC dependent variable, also does not produce any clear and interpretable pattern of means (given in Table 5, panel E).

Finally, the interaction of business exposure and age classification is significant for both the POTLOC and the ITLOC dependent variables. As seen from the means mentioned in Table 5, panels F and G, with the exception of the youngest cohort, a consistent pattern is absent. These interactions suggest something other than an easily interpretable relationship between increased age and business exposure and differences in TaxLOC.

This study is subject to the common limitation of experimental studies, that uncertainty exists regarding the generalizability of the results to groups

beyond the present sample. Although data were collected from multiple sources to include individuals with varying personal and demographic characteristics, the sample was a convenience sample, lacking proven randomness, and therefore generalizability is not certain. An additional limitation of the study findings, previously stated, is an inability to assess whether the observed differences associated with age are due to the aging process or the observed characteristics differ only in the specific generational cohorts examined. In other words, the study results may indicate differences that do not occur with aging but occur because of cultural or environmental differences affecting people of a certain age at a specific point in time. This limitation does not preclude the study's findings from being relevant to the current population of taxpayers in the age groups studied.

DISCUSSION AND CONCLUSION

The first research question is whether a multidimensional LOC belief structure exists for tax situations that parallels the health-care LOC established and documented using the MHLC. The results suggest that similar structures exist. However, the MHLC is documented as three-dimensional, and this study finds a two-dimensional TaxLOC structure with internal and powerful other dimensions. Possibly, the fate dimension does not come into play because laws and regulations influence the tax environment; laws and regulations are not random forces; thus, fate does not influence the tax-specific LOC structure. However, further research on the fate-chance dimension, using different samples, may be valuable.

This study's findings regarding the tax situation-specific LOC structure contribute to a better understanding of influences that can impact the relationship between taxpayer and tax professional. [Sharma \(1997\)](#) and [Almer, Higgs, and Hooks \(2005\)](#) discuss the special characteristics of professionals that create value for clients in their relationships with their financial professionals. Consistent with the descriptions of the unique professional and client interaction in [Sharma \(1997\)](#) and [Almer et al. \(2005\)](#), the current study's findings indicate that the tax professional and client relationship may also have special characteristics possibly reflecting power and information asymmetry. Applying the theories about professionals and clients to these findings results in a proposition that older tax clients, females, and those without significant business exposure may be more likely to respond to a tax professional as a powerful other. These clients may delegate decision-making to their tax professionals because they perceive a

knowledge or power asymmetry in the relationships. This proposition highlights the need for more research on how tax professionals may best manage relationships with clients, particularly clients who may be more likely to pass their decision-making responsibility to their tax accountants by naively accepting all of the tax professional's recommendations.

The findings provide information related to the second research question, which is whether older taxpayers have a more situation-specific powerful other TaxLOC about their relationships with their tax professionals. The results show that currently, the older cohort, composed of those born before 1944, has a stronger powerful other TaxLOC than other age groups. Furthermore, based on the post hoc ANOVA and a comparison of the means, a trend exists showing that individuals in increasingly older age categories have LOC orientation reflecting respectively stronger POTLOC. Therefore, generally, this study indicates that for the current living population of taxpayers, being older is associated with a stronger powerful other TaxLOC.

The final research question examines whether other demographic variables affect the dimensions of an individual's situation-specific LOC relationship with a tax professional. The study ultimately explored age and three other important taxpayer characteristics: business exposure, education, and gender. The MANOVA analysis indicates that, in addition to age classification, two of the characteristics, business exposure and gender, display main effects associated with LOC. Education has an effect only in interaction with other variables. In a tax situation, persons with the most business exposure have a lower powerful other score. The result is as expected and consistent with "professional as agent" theories (Almer et al., 2005; Sharma, 1997), because as individuals accrue business exposure, the knowledge and the power asymmetries between the lay client and the tax professional are reduced. Thus, the client with more business knowledge is less likely to have a high tax-related powerful other LOC. The gender difference between males and females is consistent with prior studies in that males in the study's sample are more internal than females. The means of the interaction variables that show significance do not produce clear trends, indicating the complexity of impact of these variables in real human activities.

Benefits of the Study to Scholarship

This study imports research tools and findings from the health-care discipline and applies them in a financial environment. The confirmed results of the

TaxLOC indicate that the MHLC instrument may be modified and used in a financial setting. Using a similar modification method, the instrument has the potential to be useful in exploring relationships between financial accountants, auditors, or business advisors and their clients. This would allow research to use a specific LOC scale for participants rather than the general LOC I/E Scale.

This study's results expand accounting and LOC knowledge by introducing a situation-specific LOC measure dealing with the relationship between a tax professional and client as well as by extending the LOC measure to the client. Understanding the relationship between professional and client is important for research to explore client decision-making processes. If a client has a tax situation-specific internal LOC, it is reasonable to expect that the client will assess the various options and make a decision based on his or her personal assessment. However, if the client has a tax situation-specific powerful other LOC, these findings suggest that client decision-making is more likely to adhere to whatever recommendation the professional makes because the client delegates decision-making to the professional.

The TaxLOC instrument is different from both the generalized and the other situational-specific LOC instruments. Rotter's generalized I/E Scale focuses on the basic constructs of internal and external. Other situational-specific instruments arrive at multiple constructs, with internal, powerful other, and chance being those most frequently represented. However, other than Wallston's MHLC, situation-specific LOC instruments do not deal with a trusted relationship. This study's revision of Wallston's MHLC arrived at a TaxLOC instrument that addresses the relationship between a tax professional and a client. Thus, the TaxLOC instrument is a situation-specific LOC instrument targeting a trust relationship.

Finally, this study provides research value because a majority of the accounting research on taxpayers has grouped all participants into a single category, that is, taxpayer. The situation-specific TaxLOC instrument developed in this study allows researchers to consider taxpayers' differing personal profiles. Specifically, researchers can assess whether research outcomes differ based on taxpayer characteristics, such as tax situation-specific LOC and the typical demographic characteristics that accompany a particular TaxLOC.

Benefits of the Study to Practice

Tax professionals can use the information from this study to better understand and assist their clients. As a corollary to better service, the

knowledge gained from this study can help tax professionals consider, and perhaps reduce, their legal risks. Most tax professionals know basic demographic information about their clients. Tax professionals can use the demographic information when assessing whether their clients understand the importance of tax decisions, particularly when there are potential tax risks. An example might be one where there is a questionable treatment of a specific type of income, for example, capital gains versus ordinary income. The tax professional may take an aggressive position, but the taxpayer may have a conservative, risk-averse personality type. It may be that if the client has a powerful other situation-specific LOC, he or she might, contrary to personal preference, simply follow the tax professional's advice.¹⁷

A client who understands the issues and makes decisions based on the level of risk he or she is willing to accept, rather than unquestioningly following the tax professional's advice, is more likely to be a satisfied client. An additional benefit to the tax professional is the potential for lowering the risk of litigation that may exist when the client does not fully understand or accept personal responsibility for decisions. Being sure that the client understands the risks and potential downsides of tax decisions is an important factor in protecting the tax professional from client dissatisfaction and litigation.

The practical importance of the study's findings is that clients may over-rely (powerful other LOC) or under-rely (internal LOC) on a tax professional's advice. This has the potential to cause information or power asymmetry and unintended consequences. IRS Circular 230 §10.33 suggests that tax professionals operate under a set of best practices which includes communicating clearly with the client, establishing the client's needs, obtaining the facts and issues surrounding the client's request for service, and advising the client of the importance of the advice given. Although not explored here, this study's findings, reflecting LOC profile tendencies of older taxpayers, may relate to whether a tax professional's communications are successful in achieving the goals of the IRS best practice recommendations.

The aging of the world's population, and particularly current demographic shifts in the United States, highlights the cultural and business impacts of the generational cohorts that are leaving middle age and moving through retirement years. According to the 2003 U.S. Census Bureau, the number of people ages 65 and older is expected to increase from 35 to 71 million between 2000 and 2030.¹⁸ That same U.S. Census Bureau document quotes The Center for Disease Control noting that the aging population "will have dramatic consequences for public health, the health-care financing and delivery systems, informal care giving, and pension systems" (Centers for Disease Control and Prevention [CDC], 2003, p. 106). Research

results related to client age, such as those presented here, may be important in financial professional service settings as the overall U.S. population ages and financial impacts of older individuals' activities grow.

Our findings add to the body of literature on taxpayer behavior; this study highlights the underexplored area of taxpayer and tax professional relationships and highlights the importance of those relationships when the client is older. These relationships are worthy of more study using professionalism and LOC theories.

NOTES

1. In this chapter, for brevity, we use the term physician to include all health-care professionals.

2. IRB approval was obtained from the authors' university before commencing instrument development and all other aspects of the study.

3. In developing the MHLC, Wallston initially prepared two versions, A and B. A third version, C, was developed later. In 1978, Wallston reported validation results showing both versions A and B as equally effective in determining an individual's health-care LOC. Form B, which contains 18 items categorized as three distinct dimensions labeled internal, chance or fate, and powerful other, was modified for this study. Each dimension is assessed with six items.

4. Factor analysis seeks to correlate certain similar responses to questions based on interrelationships between the questions. The analysis uses interdependence techniques in which all variables are simultaneously considered (Hair et al., 1992).

5. The two factors that had a Cronbach's alpha below 0.5, 0.435, and 0.459, respectively, were correlated to the fate questions contained within the instrument. Eight items were removed from the final analysis as a result of eliminating the two factors with low Cronbach's alpha.

6. Item numbers deleted were 1, 2, 4, 7, 9, 11, 15, and 16. Six of the eliminated items were specified as fate related in the MHLC. Two of the eliminated items, numbers 1 and 7, were classified as powerful other in the MHLC and may have become ambiguous as a result of the adaptation from health to tax. The final tax versions of these items seem likely to have been perceived by respondents to relate to either powerful other or fate.

7. Two additional statistical tests were used to confirm that factor analysis was in fact the correct methodology for these variables. First, because the respondents provided both predictor and criterion variables, Harman's Single Factor test was used to determine whether there was an issue of common method variance (Jones, Norman, & Wier, 2010). As part of the test, an exploratory factor analysis was performed on the 10 selected variables using an unrotated factor solution. If common method variance is present, a single factor will account for a majority of the variance. The test revealed two factors, which is in-line with the model, and neither of the factors explained more than 29.30% of the variability. The total variability explained by the two factors was 55.97%. The second test was the Kaiser-Meyer-Olkin

Measure of Sampling Adequacy, which measures whether factor analysis is the correct method for data reduction. Any measurement above 0.5 is considered sufficient to state that factor analysis is the correct statistical tool (Newcastle University, 2007). In the current study the measurement was 0.746, which confirms that factor analysis was the correct methodology. Thus, this study's data produces a final situational-specific TaxLOC structure consisting of two factors, POTLOC and ITLOC, with each measured by five items.

8. The instruments were included in the packet in a consistent order beginning with the TaxLOC instrument, followed by the demographic data questions, and then the Rotter's I/E Scale. Although the demographic questions were intended to serve as a distracter, the order of the instruments was not manipulated and the effectiveness of the distraction was not measured or analyzed. The extent to which the study results may have been affected by the order in which the participants completed the various instrument items is unknown. Because the participants did not complete the research instrument in a controlled setting, the reliability of any distracter or order manipulation, had one been used, could not have been determined.

9. Regardless of these instructions, in analyses other than the factor analysis, only responses from participants who currently or previously used the services of a tax professional were included. As a result, only 76 responses were used in the MANOVA and post hoc ANOVAs although a total of 99 completed packets were collected. This is described in the section on Results.

10. Primary relationship status was not used for analysis because it was highly correlated with age in the sample.

11. The analysis was also conducted using age as a continuous variable, and while still significant, the results were not as strong nor were they as clearly interpretable. As LOC may change in different directions at different stages of life, analyzing age using generational categories is more effective for this study.

12. Business exposure is coded with an additive value of one for a positive response to each of the following questions:

1. Do you follow the financial news?
2. Do you invest in the securities market?
3. Do you manage the finances in your household?

13. For taxing authority interaction, if the participant answered yes to any of the following questions, a value of one is assigned; otherwise, the variable is assigned a value of zero.

1. Have you ever received correspondence from the IRS regarding your tax return?
2. Have you ever been assessed a penalty by the IRS, that is, late filing, late payment, or other type of penalty?
3. Has the IRS required you to send in additional documentation to support a number on your tax return?
4. Has the IRS ever conducted an audit on one of your tax returns or required either you or your representative to interact face to face with the IRS for either a business or personal tax return?

5. Have you had any interaction similar to questions 2 through 4 with state or local taxing authorities?

14. Analysis was performed using this characteristic as a dichotomous and an aggregated variable (based on the number of positive answers to the questions) with no change in results.

15. Additionally, Rotter's I/E Scale was not significantly correlated (<0.05) with either powerful other or internal LOC measures.

16. This MANOVA was also conducted with the full data set of 99 participants and produced similar results.

17. As this study was exploratory in nature, it did not test whether a client would, in fact, follow the tax professional's advice if it conflicted with the client's personal risk tolerance. Future research could use the TaxLOC instrument in an experiment to determine whether an individual with a powerful other LOC would follow a tax professional's advice, even if the advice was contrary to the client's personal level of risk tolerance.

18. <http://www.cdc.gov/MMWR/preview/mmwrhtml/mm5206a2.htm>

REFERENCES

- Aldwin, C. M., & Gilman, D. F. (2004). *Health, illness and optimal aging*. London: Sage.
- Almer, E. D., Higgs, J. L., & Hooks, K. L. (2005). A theoretical framework of the relationship between public accounting firms and their auditors. *Behavioral Research in Accounting*, 17, 1–22.
- Arena, P. J., O'Hare, F., & Stabrianos, M. P. (2002). Measuring taxpayer compliance burden: A microsimulation approach. *Proceedings of the National Tax Association*.
- Bobek, D. D., Hageman, A. M., & Hatfield, R. C. (2010). The role of client advocacy in the development of tax professionals' advice. *Journal of the American Taxation Association*, 32, 25–51.
- Brothen, T., & Detzmer, D. (1983). Perceived health and locus of control in the aged. *Perceptual and Motor Skills*, 56, 946.
- Brownell, P. (1981). Participation in budgeting, locus of control and organizational effectiveness. *The Accounting Review*, 56(October), 844–860.
- Carnes, G. A., & Englebrecht, T. D. (1995). An investigation of the effect of detection risk perceptions, penalty sanctions, and income visibility on tax compliance. *The Journal of the American Taxation Association*, 17(1), 26–41.
- Centers for Disease Control and Prevention. (2003). Public health and aging: Trends in aging — United States and worldwide. *MMWR Weekly*, 52(06), 101–106.
- Christensen, A. L. (1992). Evaluation of tax services: A client and preparer perspective. *The Journal of the American Taxation Association*, 14(2), 60–87.
- Christian, C. W., Gupta, S., & Lin, S. (1993). Determinants of tax preparer usage: Evidence from panel data. *National Tax Journal*, 46, 487–503.
- Christian, C. W., Gupta, S., Weber, G. J., & Willis, E. (1994). The relation between the use of paid preparer and taxpayer's prepayment position. *Journal of the American Taxation Association*, 16(1), 17–40.

- Chung, J., & Trivedi, V. U. (2003). The effect of friendly persuasion and gender on tax compliance behavior. *Journal of Business Ethics*, 47(2), 133–145.
- Chung, Y. Y. (2001). The development and evaluation of the sales work locus of control scale. *International Journal of Management*, 18(2), 231–240.
- Cleveland, M., Kalamas, M., & LaRoche, M. (2005). Shades of green: Linking environmental locus of control and pro-environmental behaviors. *The Journal of Consumer Marketing*, 22(4/5), 198–212.
- Cloyd, C. B., & Spilker, B. C. (1999). The influence of client preferences on tax professionals' search for judicial precedents, subsequent judgments and recommendations. *The Accounting Review*, 74(3), 299–322.
- Collins, J. H., Milliron, V. C., & Toy, D. R. (1992). Determinants of tax compliance: A contingency approach. *The Journal of the American Taxation Association*, 14(2), 1–18.
- Doherty, W. J., & Baldwin, C. (1983). Locus of control in the 1970 divergence of sexes. *Journal of Personality and Social Psychology*, 48, 1048–1053.
- Doherty, W. J., & Baldwin, C. (1985). Shifts and stability in locus of control during the 1970's: Divergences of the sexes. *American Psychologist* (April), 1048–1053.
- Donnelly, D., Quirin, J., & O'Bryan, D. (2003). Auditor acceptance of dysfunctional audit behavior: An explanatory model using auditors' personal characteristics. *Behavioral Research in Accounting*, 15, 87–110.
- Dubin, J. A., Graetz, M., Udell, A., & Wilde, L. L. (1992). The demand for tax return preparation services. *The Review of Economics and Statistics*, 74, 75–82.
- Duffy, P. J., Shiflett, S., & Downey, R. G. (1977). Locus of control – Dimensionality and predictability using Likert scales. *Journal of Applied Psychology*, 62(2), 214–219.
- Frucot, V., & Shearon, W. (1991). Budgetary participation, locus of control, and Mexican managerial performance and job satisfaction. *The Accounting Review*, 66(1), 80–99.
- Furnham, A. (1986). Economic locus of control. *Journal of Human Relations*, 39, 29–43.
- Hafer, J., & Gresham, G. (2008). Lucks' role in business success: Why it's too important to leave to chance. *Journal of Behavioral and Applied Management*, 9(3), 295–316.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). *Multivariate data analysis with readings* (3rd ed.). New York, NY: Macmillan Publishing Company.
- Hale, W. D., & Cochran, C. D. (1986). Locus of control across the adult life span. *Psychological Reports*, 59(1), 311–313.
- Hyatt, T. A., & Prawitt, D. F. (2001). Does congruence between audit structure and auditors' locus of control affect job performance. *The Accounting Review*, 76(2), 263–274.
- Jensen, L., Olsen, J., & Hughes, C. (1990). Association of country, sex, social class, and life cycle to locus of control in western European countries. *Psychological Reports*, 67, 199–205.
- Johnson, T. (1977). Professions and the class structure. In: R. Scase (Ed.), *Industry society: Class, cleavage and control*. New York, NY: St. Martin's Press.
- Jones, A., III., Norman, C. S., & Wier, B. (2010). Healthy lifestyle as a coping mechanism for role stress in public accounting. *Behavioral Research in Accounting*, 22(1), 21–41.
- Kadous, K., & Magro, A. M. (2001). The effects of exposure to practice risk on tax professionals' judgments and recommendations. *Contemporary Accounting Research*, 18(3), 451–475.
- Kadous, K., Magro, A. M., & Spilker, B. C. (2008). Do effects of client preference on accounting professionals' information search and subsequent judgments persist with high practice risk? *The Accounting Review*, 83(1), 133–156.

- Lachman, M. E., & Leff, R. (1989). Perceived control and intellectual functioning in the elderly: A 5 year longitudinal study. *Developmental Psychology, 25*(5), 722–728.
- Lachman, M. E. (1986). Locus of control in aging research: A case for multidimensional and domain-specific assessment. *Psychology and Aging, 1*, 34–40.
- La Rosa, J. (1986). Escala de Locus de Control y Autoconcepto: Construcción y Validación. *Tesis de Doctorado*. Universidad Nacional Autónoma de México.
- Lefcourt, H. M., Von Baeyer, C. L., Ware, E. E., & Cox, D. J. (1979). The multidimensional-multiautributional causality scale. *Canadian Journal of Behavioral Science, 11*, 286–304.
- Levenson, H. (1974). Distinctions within the concept of internal-external control. *Journal of Personality Assessment, 38*, 377–383.
- Levenson, H. (1981). *Research with the locus of control construct (Vol. 1): Assessment methods* (pp. 15–63). New York, NY: Academic Press.
- Linder, F. (1986). Locus of control, values, and political perspectives. *Perceptual and Motor Skills* (February), 278.
- Long, J. E., & Caudill, S. B. (1987). The usage and benefits of paid tax return preparation. *National Tax Journal, 40*(1), 35–46.
- Lumpkin, J. R. (1985). Health versus activity in elderly persons locus of control. *Perceptual and Motor Skills, 60*, 288.
- Lumpkin, J. R. (1986). The relationship between locus of control and age: New evidence. *Journal of Social Behavior and Personality, 1*, 245–252.
- Masters, K. S., & Wallston, K. A. (2005). Canonical correlation reveals important relations between health locus of control, coping, affect and values. *Journal of Health Psychology, 10*, 719–731.
- Newcastle University. (2007, July). How to perform and interpret factor analysis using SPSS. Retrieved at <http://www.ncl.ac.uk/iss/statistics/docs/factoranalysis.php>. Accessed on 30 May, 2011.
- Nowicki, S., & Strickland, B. R. (1973). A locus of control scale for children. *Journal of Consulting and Clinical Psychology, 40*, 148–155.
- Payne, B. D., & Payne, D. A. (1989). Sex, race and grade differences in the locus of control orientations of at-risk elementary students. *Psychology in Schools, 26*(January), 84–88.
- Penk, W. E. (1969). Age changes and correlates of internal-external locus of control scale. *Psychology Reports, 25*, 856.
- Reed, S. A., Kratchman, S. H., & Strawser, R. H. (1994). Job satisfaction, organizational commitment, and turnover intentions of United State accountants: The impact of locus of control and gender. *Accounting, Auditing and Accountability Journal, 7*(1), 31–58.
- Rest, J. R. (1979). *Development in judging moral issues*. Minneapolis, MN: University of Minnesota Press.
- Roberts, M. L. (1995). The influence of contextual factors on IRS agents' assessments of taxpayer neglect. *Journal of the American Taxation Association, 17*, 62–77.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, 1–28.
- Sakurai, Y., & Braithwaite, V. (2003). Taxpayers' perceptions of practitioners: Finding one who is effective and does the right thing? *Journal of Business Ethics, 46*(4), 375–387.
- Schisler, D. L., & Galbreath, S. C. (2000). Responsibility for tax outcomes: An attribution theory approach. *Advances in Taxation, 12*, 173–204.

- Semykina, A., & Linz, S. (2007). Gender differences in personality and earnings: Evidence from Russia. *Journal of Economic Psychology, 28*(3), 387.
- Sharma, A. (1997). Professional as agent: Knowledge asymmetry in agency exchange. *The Academy of Management of Review, 22*(3), 758–798.
- Slemrod, J., & Sorum, N. (1984). The compliance cost of the U. S. individual income tax system. *National Tax Journal, 37*, 461–474.
- Smith, R. A. P., Woodward, N. J., Wallston, B. S., Wallston, K. A., Rye, P., & Zylstra, M. (1988). Health care implications of desire and expectancy for control in elderly adults. *Journals of Gerontology, 43*, 1–7.
- Spector, P. E. (1988). Development of the work locus of control scale. *Journal of Occupational Psychology, 61*, 335–340.
- Stein, M., & Wallston, K. A. (1983). Biofeedback and locus of control: Some considerations for future research. *American Journal of Clinical Biofeedback, 6*, 40–45.
- Stephenson, T. (2007). Do clients share preparers' self-assessment of the extent to which they advocate for their clients? *Accounting Horizons, 21*(4), 411–422.
- Straub, D. W. (1989). Validating instruments in MIS research. *MIS Quarterly, 13*(2), 147–169.
- Tsui, J. S. L., & Gul, F. A. (1996). Auditors' behaviour in an audit conflict situation: A research note on the role of locus of control and ethical reasoning. *Accounting, Organizations & Society, 21*(1), 41–51.
- Wallston, K. A. (1976). Development and validation of the health locus of control (HLC) scale. *Journal of Consulting and Clinical Psychology, 44*, 580–585.
- Wallston, K. A. (1978). Development of the multidimensional health locus of control. *Health Education Behavior, 6*, 160–170.
- Wallston, K. A. (2005). The validity of the multidimensional health locus of control scales. *Journal of Health Psychology, 10*, 623–631.
- Wallston, K. A., & Wallston, B. S. (1982). Who is responsible for your health? The constructs of health locus of control. In: G. S. Sanders & J. Suls (Eds.), *Social psychology of health and illness* (pp. 65–95). Hillsdale, NJ: Erlbaum.
- Young, T. W., & Shorr, D. N. (1986). Factors affecting locus of control in school children. *Genetic, Social, and General Psychology Monographs, 112*(4), 405–417.

APPENDIX: TAX LOCUS OF CONTROL INSTRUMENT

Each item below is a brief statement about your tax attitudes with which you may agree or disagree. The questions include the terms “tax outcome” and “tax situation.” A positive or negative tax outcome or tax situation might mean different things to different people and you should respond to each statement based on what it means for you. Examples of what you might interpret the terms to mean are

Positive	Negative
Avoiding interaction with the IRS	Being audited by the IRS
Filing a legitimate return on a timely basis with honest information	Filing a return based on uncertain information, with errors or filing late
Minimizing your tax liability	Paying more taxes than are legally required

Below each statement is a scale that ranges from strongly disagree (1) to strongly agree (6). For each item, we would like you to circle the number that represents the extent to which you agree or disagree with that statement. The more you agree with a statement, the higher will be the number you circle. The more you disagree with a statement, the lower will be the number you circle. Please make sure that you answer **EVERY ITEM** and that you circle **ONLY ONE** number per item. This is a measure of your personal beliefs; obviously, there are no right or wrong answers. Your response is completely anonymous.

1. If I have a negative tax outcome, I have the power to correct my tax situation.
2. Often I feel that no matter what I do, if I am going to have a negative tax outcome, I will have a negative tax outcome.
3. If I see an excellent tax professional regularly, I am less likely to have negative tax outcomes.
4. It seems that my tax outcomes are greatly influenced by accidental happenings.
5. I can only maintain my tax outcomes by consulting tax professionals.
6. I am directly responsible for my tax outcomes.
7. Other people play a big part in whether I have a negative or positive tax outcome.

8. Whatever goes wrong with my tax situation is my own fault.
9. When I have a negative tax outcome, I just have to let the situation run its course.
10. Tax professionals keep me having positive tax outcomes.
11. When I have a positive tax outcome, I am just plain lucky.
12. The tax outcome I experience depends on whether I take care of my tax situation properly.
13. When I have a negative tax outcome, I know it is because I have not been taking care of my tax situation properly.
14. The type of assistance I receive from other people is what is responsible for how well I recover from a negative tax situation.
15. Even when I take care of my tax situation it is easy to have a negative tax outcome.
16. When I have a negative tax outcome, it's a matter of fate.
17. I can pretty much maintain positive tax outcomes by taking good care of my tax situation.
18. Following my tax professional's guidance to the letter is the best way for me to maintain a positive tax outcome.

Each question is measured using the following scale:

1	2	3	4	5	6
Strongly Disagree	Moderately Disagree	Slightly Disagree	Slightly Agree	Moderately Agree	Strongly Agree

THE SOCIAL NORMS OF TAX COMPLIANCE: SCALE DEVELOPMENT, SOCIAL DESIRABILITY, AND PRESENTATION EFFECTS

Donna D. Bobek, Amy M. Hageman
and Charles F. Kelliher

ABSTRACT

In this study, we develop reliable scales for measuring taxpayers' social norms toward tax compliance and explore the effect of social desirability bias and several methodological issues that may affect behavioral tax and accounting studies. This study provides theoretical specificity to a potentially "decisive" (Alm & McKee, 1998) influence on tax compliance by drawing on Cialdini and Trost's (1998) taxonomy of social norms in developing our scale items. We describe in detail the methods that we used to develop these scales. On the basis of the responses of 218 experienced taxpayers, our results identify four separate social norm dimensions that correspond with the four social norm constructs identified by Cialdini and Trost. We also consider the effect of social desirability bias and find that these effects are mild for experienced taxpayers and are not directly related to compliance intentions. Finally, we also manipulate both the order of the

Advances in Accounting Behavioral Research, Volume 14, 37–66

Copyright © 2011 by Emerald Group Publishing Limited

All rights of reproduction in any form reserved

ISSN: 1475-1488/doi:10.1108/S1475-1488(2011)0000014005

items presented in the experiment and the form (online or paper-based) of the experimental instrument. While order and form effects do not interfere with the interpretation of the influence of social norms on tax compliance, we do find a significant presentation order effect driven by the paper condition, which suggests that online data collection may be preferable to uncontrolled paper and pencil administration.

INTRODUCTION

Taxpayer noncompliance remains one of the greatest and most important problems in the area of taxation (Weisbach & Plesko, 2007). In the United States, the Internal Revenue Service (IRS) estimates that it loses nearly \$350 billion annually due to taxpayer noncompliance, with the majority of the loss due to individual income tax under-reporting. The IRS focuses its resources on detection and enforcement, but recognizes that closing the tax gap (i.e., the difference between what should be paid in taxes and what is actually collected by the IRS) requires increased voluntary compliance (U.S. Department of the Treasury, 2009). Traditionally, studies of tax compliance have concentrated on the effect of economic variables such as audit rates and penalty structures; however, the current level of tax compliance in the United States is higher than would be expected based on these economic models alone (for a review, see Andreoni, Erard, & Feinstein, 1998). Andreoni et al. (1998, p. 885) therefore call for more research exploring the “diverse psychological, moral and social influences” on tax compliance behavior, to help explain and understand the current level of compliance.

Partially in response to this call, over the past decade and a half, accounting and economics researchers (e.g., Alm, McClelland, & Schulze, 1999; Bobek & Hatfield, 2003; Bobek, Roberts, & Sweeney, 2007; Blanthorne & Kaplan, 2008; Davis, Hecht, & Perkins, 2003; Hanno & Violette, 1996; Kaplan, Newberry, & Reckers, 1997; Pommerehne, Hart, & Frey, 1994; Scholz & Pinney, 1995) have begun to consider both moral and social influences on taxpayer behavior. In general, the moral and social influences modeled in these studies have been shown to have a significant effect on compliance intentions and behavior. However, consensus regarding the exact nature of these social and moral influences is lacking. This study draws on a comprehensive theory of social norms to synthesize the measurement items used in prior research for the purposes of developing specific scales to measure social norms that can be used to further tax compliance research and

ultimately help improve tax compliance. We also consider the effect of social desirability bias and the influence of presentation order and presentation form on the reliability of our scales. Specifically, we rely on Cialdini and Trost's (1998) comprehensive discussion of social norm theory. Social norms are defined as "rules and standards that are understood by members of a group, and that guide and/or constrain social behavior without the force of laws" (Cialdini & Trost, 1998, p. 152). They identify four categories of influences: general societal expectations of behavior (injunctive norms), expectations of valued others for one's own behavior (subjective norms), one's own expectations for proper behavior (personal norms), and standards that develop out of observation of others' behavior (descriptive norms). Thus, they include not only outside social influences but also an individual's personal moral "compass" in their definition of social norms. Individuals' own ethical norms (personal norms) are included as part of social norm theory because of the interrelationships among these different constructs.¹

Two hundred and eighteen experienced taxpayers participated in our study. These participants responded to an experimental instrument which included a hypothetical tax compliance dilemma, 36 different items measuring the four different types of social norms, other attitudinal and control variables, and an overclaiming scale (Randall & Fernandes, 1991) designed to assess the extent of social desirability bias exhibited by the participants. The presentation order (hypothetical tax compliance scenario early vs. later in the instrument) and the presentation format (paper vs. online) were manipulated in a between-subject experimental design.

Our results identified four separate and theoretically valid social norm constructs that are significantly related to tax compliance intentions. Furthermore, the results of our social desirability inquiry suggest that, for experienced taxpayers, our measure of social desirability tendencies was not a significant influence on tax compliance intentions; however, it was related to two of the social norm constructs. Finally, we did identify a significant order effect related to participants' compliance intentions that was driven by participants in the paper condition. Specifically, those participants in the paper condition who responded to the hypothetical tax compliance scenario early in the experimental instrument had higher tax compliance intentions than all other conditions. Interpretation of the influence of social norms on tax compliance was not affected, however. We interpret this finding to suggest that online administration may provide a more stable platform than uncontrolled paper and pencil administration.

This study contributes to the tax compliance literature by developing reliable scales that measure theoretically valid social norm constructs that

are significantly related to tax compliance intentions. The social norm scales that are developed in this study contribute to future tax research by providing a more precise means of measuring the four different social norm constructs, which in turn significantly influence tax compliance. Prior research has noted the importance of social norms in the tax compliance decision, but has not differentiated between the four types of social norms or measured them with a great deal of specificity (e.g., Alm & McKee, 1998). Thus, we believe our scale development is a significant contribution to the tax compliance literature and provides researchers with a tool to be used in future tax compliance research. The detailed consideration of how these norms affect tax compliance, the antecedents of social norms, and the interrelationships among the norm constructs are explored in a related study (Bobek, Hageman, & Kelliher, 2011).

Further, this study also makes several contributions to the behavioral accounting literature more generally by exploring social desirability bias and presentation order and form effects. Specifically, we introduce a measure to control for social desirability, called the overclaiming scale (Randall & Fernandes, 1991). Second, our results suggest that online administration of an un-controlled experimental instrument appears to provide for more stable responses. Lastly, our findings also emphasize the need to tailor scale items to measure the specific behavior being examined (as opposed to only measuring general attitudes and beliefs), consistent with attitude theory recommendations (Ajzen & Fishbein, 1980). These findings will assist future behavioral researchers in making experimental design choices.

The remainder of this chapter is organized as follows. In the next section, we discuss social norm theory, the measurement of social norms in prior tax research, and social desirability bias. Next, we report our results. Finally, we draw conclusions and offer suggestions for future research.

THEORY AND PRIOR RESEARCH

Social Norm Theory

Social norms are a social psychology construct that are defined as “rules and standards that are understood by members of a group, and that guide and/or constrain social behavior without the force of law” (Cialdini & Trost, 1998, p. 152). The four different social norm constructs are injunctive norms, descriptive norms, subjective norms, and personal norms. Table 1 presents a definition and explanation of the relevance of each of the social norm

Table 1. Summary Table of Social Norm Constructs.

	Descriptive Norms	Injunctive Norms	Subjective Norms	Personal Norms
Description	What one perceives that other people <i>do</i> in a given situation. Watching others provides information about what is “normal” in a novel or ambiguous situation.	The perception of what most people think others <i>should</i> do in a given situation. They specify what should be done and are the moral rules of the group.	A person’s perception about what those who are important to him think he should or should not do in a given situation. They are one’s perceptions of the injunctive norms held by the people whose opinions matter most to them.	Self-based standards or expectations for behavior that flow from internalized values.
Social goal being achieved by conforming to norm.	Effective action – the desire to be accurate in one’s choices and behaviors.	Building and maintaining social relationships.	Building and maintaining social relationships.	Managing self-concept.
When will they matter?	Most likely to use evidence of others’ behavior to decide most effective course of action when the situation is novel, ambiguous, or uncertain, and especially when the source of reference is similar to us.	They motivate behavior by promising social rewards or punishments. Need not be expressed in order to direct behavior. These norms might be more powerful when they are made salient.	When people are motivated to comply with the norms of “referent” others.	Enforced through the anticipation of self-enhancement or self-deprecation (“self-reinforcing”).

Table 1. (Continued)

	Descriptive Norms	Injunctive Norms	Subjective Norms	Personal Norms
Why norm may influence tax compliance	If tax compliance is viewed as “normal,” then descriptive norms will lead to compliance. If situation is ambiguous or novel, then it may matter what one thinks others do. High profile cheaters or compliers may cause imitation.	To the extent one has a desire to conform to societal norms, injunctive norms should influence behavior.	Influence of important others might affect likelihood of feeling guilty. Conformity is greater with friends, family and/or “similar” others. More likely to share tax compliance choices with referent others.	Complying with tax laws is the <i>legal</i> course of action. To the extent that individuals also believe it is the <i>moral</i> course of action, they should be more likely to comply.

Source: Bobek et al. (2007).

constructs to tax compliance research. Specifically, *injunctive norms* specify what *should* be done and are therefore the moral rules of the group. *Descriptive norms*, on the contrary, are standards that develop out of observation of how others *actually* behave in particular situations. *Subjective norms* relate specifically to the expectation of *important* others (e.g., family, friends, coworkers, and so forth). Finally, *personal norms* are one's *own* expectations for behavior, which may arise from the internalization of injunctive norms.²

While prior researchers (e.g., Blanthorne & Kaplan, 2008; Bobek et al., 2007; Davis et al., 2003; Wenzel, 2004, 2005) have examined the influence of social norms on tax compliance behavior, many open questions remain regarding their effect. Almost all prior tax research that has included a social or personal norm construct has identified a significant influence on tax compliance from these constructs. However, there has been a lack of specificity regarding the identity, emergence, and transmission of these norms. Alm and McKee (1998, p. 270) concluded that, "social norms play an important, perhaps a decisive role in tax compliance". Furthermore, Wenzel (2004, p. 215) suggested that greater refinement of the social norm constructs and their role on tax compliance behavior was needed to understand their potentially "substantial impact." Thus, a more rigorous and comprehensive measure of social norms is needed.

The primary goal of this study is to provide additional specificity regarding the measurement of the social norms of tax compliance. Our experimental procedures therefore rigorously investigate four separate constructs that map to the social norm theory as articulated by Cialdini and Trost (1998). In constructing a scale for taxpayer social norms, we followed the guidelines of Mason and Levy (2001) on the development of measures for latent constructs in accounting research.

Social Norm Measures

The first step in developing our scale was to generate a sample of potential item measures based on an extensive literature search of prior tax compliance studies that have *measured* social norms. These items were then categorized in terms of the four types of social norms from the Cialdini and Trost (1998) framework. We further categorized the items as either a general item (e.g., *Most people think it is morally wrong to engage in tax evasion behavior*) or a scenario-specific item (e.g., *Most people would feel justified in taking the additional \$2,000 deduction*). Prior research has measured both social norms in general and social norms related to a specific compliance decision. Attitude

theory (Azjen & Fishbein, 1980) suggests that attitudes about a specific behavior will have more predictive power than general attitudes. Similarly, we posit that social norms regarding a specific tax compliance decision may well be superior predictors of tax compliance intentions. However, since prior research has considered both general and scenario-specific items, we also include both in this study.

When categorizing items from prior research, we endeavored to find at least four different scenario-specific items for each norm along with at least four general items for each norm. As shown in Table 2, most items originated from prior research; however, we developed additional items for some categories (particularly for descriptive norms) to achieve our four-item minimum. Additionally, some items were modified from prior research. For example, we included a general injunctive norm item from Wenzel (2004), "Do most people think they should honestly declare cash earnings on their taxes?" but also included a modified version of the item to measure general subjective norms, "The people closest to me (for example, my family and/or my friends) think they should honestly report cash earnings on their tax return."

Two of the researchers independently categorized the items and were in agreement regarding which items measured each of the social norm constructs. This resulted in an instrument with 36 items measuring the different types of social norms. Thirty-one (31) of the items were measured on a seven-point Likert-type scale with participants indicating their degree of agreement (1 = strongly agree and 7 = strongly disagree). Five of the items were on an 11-point scale as they measured, for example, the percentage of taxpayers who engaged in a specific behavior. The response scale for these items was anchored with 0% and 100%, with 10 percentage point increments in between. Many of the items were reverse-coded to ensure a consistent influence on compliance. The first three columns of Table 2 present a list of these social norm items categorized by the type of norm and the original source of the item, as well as whether the item was intended to measure scenario-specific norms or general social norms.

Social Desirability Bias

Randall and Fernandes (1991, p. 805) describe social desirability bias as "the tendency of individuals to deny socially undesirable traits and behaviors and to admit to socially desirable ones." This potential bias is of concern in social norm and tax compliance research because individuals might be inclined to minimize reports of socially undesirable behavior. Prior researchers have used a number of different scales to measure social desirability, such as impression

Table 2. Summary of Social Norm Items.

Type of Norm	Items	Source	Retained in Pilot Study	Retained in Final Scale Development
Personal norms – general	How much of a moral obligation – that is, an obligation based on your own personal feelings of what’s right and wrong – do you feel to be completely honest in filling out your tax return?	Scholz and Pinney (1995)		X
Personal norms – general	I think I should honestly report cash earnings on my tax return.	Wenzel (2004)	X	
Personal norms – general	I think it is acceptable to overstate tax deductions on my tax return (R).	Wenzel (2004)	X	X
Personal norms – general	I think working for cash payments without paying tax is a trivial offense (R).	Wenzel (2004)		
Personal norms – specific	Would you feel guilty if you took the additional \$2,000 deduction?	Stalans, Kinsey, and Smith (1991)	X	X
Personal norms – specific	Would you feel ashamed if you took the additional \$2,000 deduction?	Modified from Stalans et al. (1991) and Murphy (2004)	X	X
Personal norms – specific	Would you feel justified in taking the additional \$2,000 deduction? (R).	Modified from Stalans et al. (1991)		X
Personal norms – specific	Would you feel pleased in taking the additional \$2,000 deduction? (R).	Modified from Stalans et al. (1991)	X	X
Personal norms – specific	Would you be afraid you would get caught if you took the additional \$2,000 deduction?	Modified from Stalans et al. (1991)		
Personal norms – specific	You would definitely think it was not okay (i.e., morally right) to deduct the additional \$2,000.	Modified from Hanno and Violette (1996)		
Subjective norms – general	Do you get the general impression from people you work with that they think it is acceptable for people to pay less income taxes than they legally owe? (R).	Stalans et al. (1991)	X	

Table 2. (Continued)

Type of Norm	Items	Source	Retained in Pilot Study	Retained in Final Scale Development
Subjective norms – general	The people closest to me (e.g., my family and/or my friends) think they should honestly report cash earnings on their tax return.	Modified from Wenzel (2004)		
Subjective norms – general	The people closest to me (e.g., my family and/or my friends) think that it is acceptable to overstate tax deductions on their tax return (R).	Modified from Wenzel (2004)		
Subjective norms – general	The people closest to me (e.g., my family and/or my friends) think that working for cash payments without paying tax is a trivial offense (R).	Modified from Wenzel (2004)		
Subjective norms – specific	Most people you know would definitely not think it is okay (i.e., morally right) to deduct the additional \$2,000.	Modified from Hanno and Violette (1996)	X	X
Subjective norms – specific	Your co-workers would definitely not think it is okay (i.e., morally right) to deduct the additional \$2,000.	Modified from Hanno and Violette (1996)	X	X
Subjective norms – specific	Your family would definitely not think it is okay (i.e., morally right) to deduct the additional \$2,000.	Modified from Hanno and Violette (1996)		X
Subjective norms – specific	Your friends would definitely not think it is okay (i.e., morally right) to deduct the additional \$2,000.	Modified from Hanno and Violette (1996)	X	X
Subjective norms – specific	Your spouse/significant other would definitely not think it is okay (i.e., morally right) to deduct the additional \$2,000.	Modified from Hanno and Violette (1996)		X

Table 2. (Continued)

Type of Norm	Items	Source	Retained in Pilot Study	Retained in Final Scale Development
Injunctive norms – general	Most people in the U.S. think they should honestly report cash earnings on their tax returns.	Wenzel (2004)	X	
Injunctive norms – general	Most people in the U.S. think it is acceptable to overstate tax deductions on their tax returns (R).	Wenzel (2004)		
Injunctive norms – general	Most people in the U.S. think working for cash payments without paying tax is a trivial offense (R).	Wenzel (2004)		
Injunctive norms – general	Most people think it is morally wrong to engage in tax evasion behavior.	Modified from Kaplan, Reckers, and Roark. (1988)		
Injunctive norms-specific	Most people would feel guilty if they took the additional \$2,000 deduction.	Modified from Stalans et al. (1991)	X	X
Injunctive norms – specific	Most people would feel ashamed if they took the additional \$2,000 deduction.	Modified from Stalans et al. (1991)	X	X
Injunctive norms – specific	Most people would feel justified in taking the additional \$2,000 deduction (R).	Modified from Stalans et al. (1991)		
Injunctive norms – specific	Most people would feel pleased in taking the additional \$2,000 deduction (R).	Modified from Stalans et al. (1991)		
Injunctive norms – specific	Most people would be afraid they would get caught if they took the additional \$2,000 deduction.	Modified from Stalans et al. (1991)		X
Descriptive norms – general	In your opinion, what percentage of taxpayers at your income level deliberately pay less taxes than they legally owe? (R).	Modified from Scholz and Pinney (1995) and Scholz and Lubell (1998)	X	X

Table 2. (Continued)

Type of Norm	Items	Source	Retained in Pilot Study	Retained in Final Scale Development
Descriptive norms – general	In your opinion, what percentage of taxpayers at your income level carelessly, but unknowingly pay less taxes than they legally owe? (R).	Modified from Scholz and Pinney (1995) and Scholz and Lubell (1998)	X	X
Descriptive norms – general	Tax evasion in the U.S. is widespread (R).	Developed for the study.		
Descriptive norms – general	From what I have observed, most people in the U.S. would do anything to avoid paying taxes (R).	Developed for the study.		
Descriptive norms – specific	Given the opportunity, what percentage of U.S. taxpayers do you think would take the additional \$2,000 automobile deduction? (R).	Developed for the study.	X	
Descriptive norms – specific	Given a similar opportunity, do you think the average U.S. taxpayer would deduct the additional \$2,000? (R).	Developed for the study.		
Descriptive norms – specific	Given the opportunity, what percentage of U.S. taxpayers at your income level do you think would take the additional \$2,000 automobile deduction? (R).	Developed for the study.	X	X
Descriptive norms – specific	What percentage of the additional \$2,000 in automobile expenses do you think is the closest to the amount the average U.S. taxpayer would deduct? (R).	Developed for the study.		X

Note: (R) = response was reverse scored.

management, self-deception enhancement, and narcissism. However, most of these scales suffer from a lack of reliability and an inability to differentiate between socially desirable responses and true responses (Barger, 2002; Leite & Beretvas, 2005; Randall & Fernandes, 1991). In contrast, an *overclaiming* scale assesses the extent to which individuals claim knowledge of something that does not exist. This scale is strongly related to individuals' self-deception enhancement (Randall & Fernandes, 1991), indicating that individuals who overclaim have less self-insight. Furthermore, measuring an individual's tendency to overclaim captures individuals' tendencies to claim knowledge that is in actuality not true. We therefore study whether a greater tendency to overclaim affects participants' proclaimed tax compliance intentions and/or their social norms to attest whether these measures could suffer from potential social desirability bias.

Presentation Order and Form Effects

We investigate presentation order (compliance scenario early or late) and presentation form (paper-based or online) effects to ensure that our study results are robust to these variations and to provide experimental design guidance to future behavioral researchers investigating a sensitive subject such as tax compliance. Previous accounting research has demonstrated that auditors and tax professionals may be influenced by the order of experimental information (e.g., Asare, 1992; Kennedy, 1993; Pei, Reckers & Wyndelts, 1992). Similarly, the placement of the tax compliance scenario within the experimental instrument might possibly affect how participants respond. We therefore manipulate the order of the scenario placement to study whether order effects may exist in a tax compliance setting.

Second, the use of online and Internet-based surveys has only become prevalent in the past decade (Bryant, Hunton, & Stone, 2004). Given that pre-2000 tax compliance research was almost uniformly conducted with paper-based instruments, we investigate whether responses differ between these two forms. Prior research on tax professionals has established that in-laboratory and out-of-laboratory experiments have similar convergent validity (Alexander, Blay, & Hurtt, 2006). However, prior research has not specifically examined whether the difference in experimental form matters in tax compliance research, particularly when participants are answering questions regarding their propensity to engage in illegal behavior (tax evasion) and their related social norms. Thus, we consider potential differences in experimental form.

RESEARCH METHODS

Experimental Task and Procedures

Participants responded to a hypothetical tax compliance dilemma in this study. The dilemma was adapted from Kaplan et al. (1997) and Bobek and Hatfield (2003), who used a similar scenario. To increase participants' ability to relate to the hypothetical taxpayer, the scenario included additional contextual details to make the taxpayer seem more true-to-life; this also was designed to help reduce the "better than average" effect that individuals often exhibit when evaluating a comparison target (Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995). The text of the tax compliance dilemma is presented below:

Larry Brown is a 40-year old married father of two boys. He has his own small business and his wife, Laura, works part-time. In addition to working in his business, he is the coach for his son's Little League baseball team.

Larry prepares his own income tax return. Larry uses his personal automobile for both business and personal reasons. The tax laws provide that automobile expenses are deductible to the extent the automobile is used for business. In preparing his income tax return, Larry determines that the automobile was used 60% for business. However, Larry also calculates if he FALSELY claimed it was used 80% for business, his deduction would increase by \$2,000, and he would save \$500 in taxes.

The primary dependent variable is *taxpayer compliance intentions*, measured as the response to, "Do you think that if you were in similar situation as Larry, YOU would deduct the ADDITIONAL \$2,000?" Participants responded using a 7-point Likert-type scale, ranging from "7 = very unlikely" to "1 = very likely" that participants would evade taxes by taking the additional deduction. Thus, higher scores indicate more favorable tax compliance intentions. While not the primary focus of this study, we use this variable to establish that social norms have a pronounced influence on tax compliance and also use the variable to test the robustness of our social norm scales with regard to social desirability, order, and form effects.

To assess whether participants' responses are affected by the form (i.e., paper vs. online) of the questionnaire, one-half of the potential participants were contacted via mail with a cover letter, paper copy of the instrument, and a postage-paid reply envelope. The other half received an e-mail message and a link to the online instrument. The paper-based instrument and online instrument contained identical content.

In addition to manipulating the form of the instrument (i.e., paper or online), we manipulated whether participants responded to the hypothetical tax compliance scenario toward the beginning of the instrument or after responding to items measuring their general attitudes and beliefs about taxation. Thus, our study used four experimental conditions based on the form (paper or online) and presentation order (scenario early or later) to examine whether differences in the presentation form of the experimental instrument influenced the study's results. The final cell sizes for these experimental groups ranged from 50 to 64 individuals.³

In completing the instrument, all participants first responded to general demographic items (e.g., age, income, education, etc.). Next, one-half of the participants (in the "scenario early" manipulation) were presented with the hypothetical tax compliance scenario and answered items measuring what they and others would do if they were in a similar position. They then responded to items measuring their *social norms* of tax compliance (including personal, subjective, injunctive, and descriptive norms) for the *specific* scenario. These participants then responded to items measuring their general social norms of compliance for the four types of social norms, as well as items measuring their general *attitudes* and perceived *fairness* of the federal income tax system. Participants assigned to the "scenario later" manipulation completed the instrument in a different order; after responding to demographic items, these participants responded to the items measuring their general social norms and attitudes, and then received and responded to items regarding the specific tax compliance scenario. Both groups concluded the experiment by responding to an "overclaiming" scale to test for social desirability bias.

All content was identical between the online and paper versions, as each "page" within the paper instrument was reproduced as a "screen" within the online instrument. To obtain a professional appearance, the paper-based instrument was printed as a professional booklet, such that participants could view two pages at a time. In the online version, however, participants could toggle between screens, but could only view the items on one screen at a time.

The experimental instrument was pilot-tested several times. On the basis of these pilot tests, changes were made to the wording of some of the items, details were added to the tax compliance scenario, and control items were added. An initial pilot test was conducted with 91 undergraduate students who completed a paper-based instrument and answered debriefing questions; based on their feedback, some of the items were revised for clarity. Next, a subsequent pilot test of 122 experienced taxpayers completed the experiment

in paper-based form. On the basis of this second pre-test group, extensive changes were made to the items measuring social norms, additional details were added to the tax compliance scenario, and a scale was added to control for potential social desirability bias. In a final pilot test, 54 undergraduate taxation students completed the revised paper-based instrument and provided additional feedback on the items. On the basis of this final pilot test, minor changes were made to the wording of some of the items and their presentation. During pilot testing, we did not manipulate the form or ordering of the experimental tasks. We manipulated these factors in the main study to assess whether the results were sensitive to the presentation form and order of the experimental task.

Participants

Participants for the final study were obtained by having students, in exchange for extra credit, recruit subjects who were “non-accounting students at least 25 years old that are U.S. citizens or residents that have filed a federal income tax return.”⁴ This recruiting technique resulted in 470 potential participants. One-half of the potential participants (235 participants) were contacted via mail with a cover letter, paper copy of the instrument, and a postage-paid reply envelope. The other half received an e-mail message (with the cover letter information) and a link to the online instrument. Four of the e-mail addresses bounced as non-deliverable, whereas six of the paper-based packets that were mailed were returned as non-deliverable. Of the 460 contacted participants, 223 individuals responded and the responses of 218 individuals were included in the final analysis (115 who returned the instrument through the postal service and 103 who completed the instrument online). Of the five participants whose responses could not be included, one participant did not meet the study’s age requirement, three participants did not report their age (and could not be evaluated as to whether they met the study’s criteria), and one participant indicated that he or she had not previously filed an individual income tax return. This resulted in a usable response rate of 47%.

Table 3 reports the sample demographics and compares the sample to the national population. Our sample was older (71% between the ages of 45 and 64) and included more females (57%) than males. Participants also had higher income and were better educated than the national population. The majority of participants (68%) did *not* prepare their own tax returns.

Table 3. Sample Demographics.

Sample ($n = 218$) ^a	2007 National Population ^b
<i>Age</i>	<i>Age</i>
Under 25 = 0%	20–24 = 10%
25–44 = 21%	25–44 = 38.5%
45–64 = 71%	45–64 = 34.5%
65+ = 8%	65+ = 17%
<i>Family income</i>	<i>Family income</i>
Less than \$25,000 = 3%	Less than \$25,000 = 17%
\$25,000–\$49,999 = 16%	\$25,000–\$49,999 = 24%
\$50,001–\$74,999 = 17%	\$50,000–\$74,999 = 20.5%
\$75,000–\$99,999 = 21%	\$75,000–\$99,999 = 14.5%
Over \$100,000 = 42%	Over \$100,000 = 24%
<i>Gender</i>	<i>Gender</i>
Male = 43%	Male = 49%
Female = 57%	Female = 51%
<i>Education</i>	<i>Education</i>
Did not complete	Did not complete
High school = 0.5%	High school = 16%
High school graduate = 6.5%	High school graduate = 30%
Some college = 26.5%	Some college = 27%
College graduate = 41%	College graduate = 27%
Post-graduate study = 25.6%	Post-graduate study = 10%
<i>Prepare your own tax return?</i>	
Yes = 32%	
No = 68%	

^aNumbers are expressed as a percentage of the total sample of respondents providing information for each demographic question.

^bU.S. Census Bureau: *Age* represents percentage of adults over age 19 in each category; *Family income* numbers represent the percentage of families in each category; *Education* numbers represents percentage of persons over 25 in each category.

The data were collected in two “waves” during the spring and summer of 2008.⁵ No reminders were sent in the first wave, and the usable response rate was 39%. In the second wave, postcard reminders or e-mail reminders (based on the presentation form condition) were sent after two weeks; the usable response rate was 55%. The follow-up reminders thus resulted in a significantly higher response rate ($p < .001$). To assess for non-response bias, early and late responders from the second wave were compared. No statistically significant differences emerge, suggesting that few insights are lost from non-responders.

RESULTS

Social Norm Scales

As discussed previously, 36 items were included in the final experimental instrument to measure the four different social norm constructs. Sixteen items related to general social norms and 20 items related to scenario-specific social norms. As mentioned previously, there were at least eight items for each of the four social norm constructs (see [Table 2](#)). Cronbach's alpha was computed for each of the four scales, dropping items with low correlations or poor item-to-total correlations. Eliminating these items resulted in 29 items, with each scale demonstrating acceptable levels of Cronbach's alpha, ranging from 0.797 (injunctive norms) to 0.904 (personal norms). Next, principal components analysis with varimax rotation was used to produce orthogonal factors and to further purify the measures. Ten items were eliminated with low item loadings, while another item (measuring subjective norms) was eliminated due to loading on a factor composed of a different type of social norm. The final scale resulted in 18 items, four of which measure general social norms and 14 measure scenario-specific social norms. This implies that scenario-specific scale items may be better than general scale items at explaining a specific compliance behavior. This is consistent with attitude theory ([Azjen & Fishbein, 1980](#)) and suggests that researchers should take care to include scenario-specific items when constructing an experimental instrument.⁶ Refer again to [Table 2](#) for more details regarding the questions retained in the final scale development and in an initial pilot test.⁷

[Table 4](#) reports the results of the factor analysis, including the item loadings for each of the measures included in the final scales. We successfully achieved an important objective of this study as each distinct factor unambiguously measures one of the four types of social norms identified by [Cialdini and Trost \(1998\)](#). In total, the four factors explained 71% of the data's variance. Factor 1 contains items measuring personal norms and explained 45% of the variance. Factor 2 contains items measuring subjective norms and explained 10% of the variance. Factor 3 contains items measuring injunctive norms and explained 9% of the variance. Finally, Factor 4 contains items measuring descriptive norms and explained 7% of the variance. The Cronbach's alpha for the four scales comprised of the items retained by the final factor analysis all demonstrate acceptable levels of reliability, ranging from 0.766 (descriptive norms) to 0.927 (subjective norms). For the remainder of the analyses, these four rotated factor scores were retained as independent variables.⁸

Table 4. Factor Analysis Results – Overall.

Item Description ^a	Mean Response	Factor 1 Loading	Factor 2 Loading	Factor 3 Loading	Factor 4 Loading
How much of a moral obligation ... do you feel to be completely honest ...	5.88	.697			
I think it is acceptable to overstate tax deductions ... (R)	5.55	.711			
Would you feel guilty if you took the additional \$2,000 deduction?	5.61	.799			
Would you feel ashamed if you took the additional \$2,000 deduction?	5.32	.814			
Would you feel justified in taking the additional \$2,000 deduction? (R)	5.03	.697			
Would you feel pleased in taking the additional \$2,000 deduction? (R)	5.33	.733			
Most people you know (would disapprove)	4.50		.851		
Your co-workers (would disapprove)	4.32		.828		
Your family (would disapprove)	5.04		.832		
Your friends (would disapprove)	4.64		.852		
Your significant other (would disapprove)	5.26		.747		
Most people would feel guilty if they took the additional \$2,000 deduction	3.85			.788	
Most people would feel ashamed if they took the additional \$2,000 deduction	3.62			.816	
Most people would be afraid they would get caught ... (R)	4.48			.717	

Table 4. (Continued)

Item Description ^a	Mean Response	Factor 1 Loading	Factor 2 Loading	Factor 3 Loading	Factor 4 Loading
What percentage of taxpayers at your income level deliberately pay less ... (R)	7.32 (37%)				.664
What percentage of U.S. taxpayers at your income level carelessly, but unknowingly pay less ... (R)	8.13 (29%)				.740
What percentage of taxpayers at your income level ... would take the additional deduction ... (R)	6.24 (48%)				.746
What percentage of the additional \$2,000 ... is the closest to the amount the average U.S. taxpayer would deduct? (R)	6.40 (46%)				.720
Percentage of variance Explained		45%	10%	9%	7%
Cronbach's Alpha		.904	.927	.810	.766
Factor Description		Personal norms	Subjective norms	Injunctive norms	Descriptive norms

^aThe personal, subjective, and injunctive social norm items were assessed on a 1–7 Likert-type scale. Descriptive norm items were assessed on a 1–11 Likert-type scale. Items were coded so that higher values indicated more favorable social norms toward tax compliance. Factor loadings are varimax rotation scores.

To investigate whether the social norm scales help explain tax compliance behavior, we regress tax compliance intentions on the orthogonal scores for personal norms, subjective norms, injunctive norms, and descriptive norms. As shown in Table 5, all four factor scores have a statistically significant influence on tax compliance intentions ($p < .05$), while the influence of personal norms, subjective norms, and descriptive norms is particularly

Table 5. Social Norm Constructs^a.

	Parameter Estimate	Standard Error	Standardized Coefficient	p-value
Constant	5.113	.111		<.001
Personal norms factor	1.338	.111	.616	<.001
Subjective norms factor	.683	.111	.315	<.001
Injunctive norms factor	.237	.111	.109	.034
Descriptive norms factor	.424	.111	.196	<.001
<i>Model statistics</i>				
<i>F</i> statistic = 50.366				
<i>p</i> -value = <.001				
Adjusted <i>R</i> ² = .520				

Notes: Tax compliance intentions are measured on a 7-point Likert-type scale, with 7 = “very unlikely” and 1 = “very likely” participants would evade taxes by taking the additional deduction (thus, higher scores indicate more favorable tax compliance intentions). *Social Norms Factors* are the orthogonal factor scores obtained from the factor analysis described in Table 4, with higher values more favorable toward compliance.

^aRegression model tested: Tax compliance intentions = $\beta_0 + \beta_1$ personal norms + β_2 subjective norms + β_3 injunctive norms + β_4 descriptive norms.

strong ($p < .001$). The overall model is also statistically significant ($p < .001$) with an overall model adjusted R^2 of .520. Collectively, social norms exhibit a significant influence on tax compliance behavior.

Social Desirability Bias

We also investigate the potential effect of social desirability bias. We measured overclaiming using a 10-item scale (Randall & Fernandes, 1991), in which participants indicated their level of familiarity with items in five categories unrelated to taxes (music, movies, television shows, clothing designers, and consumer products); each category contained two fake items intermingled with three real items.⁹ Familiarity with each item was measured on a 5-point Likert-type scale with 1 = “Not at all” and 5 = “Very.” Only the responses to the 10 fake items are included in the scale; thus, the overclaiming scores ranged from 10 (minimum) to 50 (maximum). A score of 10 indicates the respondent did not overclaim at all (i.e., they responded “not at all” to all 10 fake items). A score above 10 suggests some level of overclaiming. The mean (standard deviation) on the overclaiming scale was 12.42 (3.40).¹⁰ Thus, participants exhibited low levels of overclaiming. As displayed in Table 6,

Table 6. Social Norm Constructs and Overclaiming.

Panel A: Regression ^a				
	Parameter Estimate	Standard Error	Standardized Coefficient	<i>p</i> -value
Constant	5.118	.437		<.001
Personal norms Factor	1.336	.116	.616	<.001
Subjective norms Factor	.682	.114	.315	<.001
Injunctive norms Factor	.235	.112	.109	.037
Descriptive norms Factor	.424	.113	.196	<.001
Overclaiming	.000	.033	-.001	.988
<i>Model statistics</i>				
	<i>F</i> statistic = 39.541			
	<i>p</i> -value = <.001			
	Adjusted <i>R</i> ² = .516			
Panel B: Correlations				
	Overclaiming Correlation Coefficient (<i>p</i> -value)			
Personal norms factor	-.242 (<.001)			
Subjective norms factor	.162 (.029)			
Injunctive norms factor	-.039 (.599)			
Descriptive norms factor	-.141 (.058)			

Notes: Tax compliance intentions are measured on a 7-point Likert-type scale, with 7 = “very unlikely” and 1 = “very likely” participants would evade taxes by taking the additional deduction (thus, higher scores indicate more favorable tax compliance intentions). *Social norms factors* are the orthogonal factor scores obtained from the factor analysis described in Table 4, with higher values more favorable toward compliance. *Overclaiming* is measured on a (minimum 10-point) maximum 50-point scale, with higher values indicating a greater tendency to overclaim.

^aRegression model tested: Tax compliance intentions = $\beta_0 + \beta_1$ personal norms + β_2 subjective norms + β_3 injunctive norms + β_4 descriptive norms + β_5 overclaiming.

Panel A, when we add participants’ overclaiming scale score to the regression analysis of tax compliance intentions, it does not approach statistical significance ($p > .75$) and none of the results for the coefficients or the overall models change (see Table 6, Panel A). We therefore conclude that the interpretation of our results is not affected by potential social desirability bias.

We further test whether overclaiming was related to any of the social norm scales. Table 6, Panel B, reports the Pearson correlation coefficient between the overclaiming scale score and each of the four social norm factors. Both the personal norms factor (correlation coefficient of $-.242$,

$p < .001$) and the descriptive norms factor (correlation coefficient of $-.141$, $p = .058$) are significantly negatively correlated with the overclaiming scale. These correlations indicate that individuals that claim knowledge of a fake item (i.e., overclaim) tend to have *less* favorable personal norms and descriptive norms regarding tax compliance. Interestingly, the correlation between subjective norms and overclaiming was positive and significant (correlation coefficient of $.162$, $p = .029$), meaning that individuals that overclaim are more likely to view their close others (family and friends) as having more favorable norms toward tax compliance.¹¹ Thus, while overclaiming was not related to tax compliance intentions in our study of experienced taxpayers, it did have a robust direct negative influence on personal norms and descriptive norms associated with compliance. Collectively, these results indicate that while overclaimers may have different social norms than non-overclaimers, these differences do not appear to directly affect reported tax compliance intentions.

Interestingly, in our pilot test of 54 undergraduate students, we did find that overclaiming influenced tax compliance decisions. Among those participants, overclaiming was significantly related to tax compliance intentions. Participants with higher overclaiming scores (claiming knowledge of a fake item) were *more* likely to report that they would comply with tax laws (and less likely to report they would engage in cheating behavior). Further analysis of the student responses split the data into overclaimers (those that reported knowledge of any of the fake items) and “non” overclaimers. While social norms influenced the tax compliance intentions of both groups, the models’ adjusted R^2 and standardized regression coefficients for social norms were lower for the overclaiming participants. One possible explanation for the difference between experienced taxpayers and college students is that college students may be more likely to overclaim than experienced taxpayers in the community at large. Undergraduate students in our pilot test exhibited statistically significantly higher ($p < .05$) degrees of overclaiming than experienced taxpayers; the mean on the overclaiming scale for our undergraduate sample was 13.80, similar to the mean of 13.50 reported by [Randall and Fernandes \(1991\)](#) in their undergraduate sample.

Presentation Order and Presentation Form Effects

We also test whether the presentation order (whether participants evaluated the tax scenario early or later) or the presentation form of the instrument

(online versus paper) affected the results. This 2×2 between-subjects manipulation resulted in four experimental conditions.

Results indicate that none of the experimental conditions differed in responses to demographic, attitudinal, or social norm items. However, a difference did emerge among tax compliance intentions. As shown in Table 7, Panel A, an ANOVA with order and form as fixed factors and tax compliance intentions as the dependent variable indicates that there was a marginally significant difference due to the presentation order ($p < .10$). Table 7, Panel B, displays descriptive statistics for the tax compliance intentions by

Table 7. Tax Compliance Intentions.

Panel A: ANOVA			
	<i>F</i> -Statistic	Significance Level	
Form	1.186	.277	
Order	2.913	.089	
Form \times order interaction	1.665	.198	
<i>Model statistics</i>			
<i>F</i> -statistic = 2.157			
Significance level = .094			
Adjusted R^2 = .016			
Panel B: Tax compliance intentions – by condition ^a			
	Paper-Based (Form)	Online (Form)	Row Means (Form)
Tax scenario early (order)	5.77** (1.974) $n = 64$	5.08 (2.212) $n = 50$	5.46* (2.100) $n = 114$
Tax scenario later (order)	4.90 (2.119) $n = 51$	4.96 (2.128) $n = 50$	4.93 (2.113) $n = 101$
Column means (order)	5.38 (2.076) $n = 115$	5.02 (2.160) $n = 100$	5.21 (2.118) $n = 215$

*Tax compliance intentions are significantly higher when the tax scenario is presented early (order), $p < .10$.

**The experimental condition with the paper-based form, tax scenario early has higher tax compliance intentions than all other experimental conditions, all $p < .10$.

^aMeans (standard deviations) of *tax compliance intentions*, which are measured on a 7-point Likert-type scale, with 7 = “very unlikely” and 1 = “very likely” participants would evade taxes by taking the additional deduction (thus, higher scores indicate more favorable tax compliance intentions). Three of the 218 participants (in the “tax scenario later” condition) did not respond to the tax compliance intentions question and hence are excluded from this analysis.

experimental conditions. In tandem, the ANOVA and descriptive statistics indicate that participants had somewhat higher tax compliance intentions when they evaluated the specific scenario early than when they answered items relating to general tax attitudes first and evaluated the tax scenario later (i.e., an *order effect*).

Further analysis investigates whether the order effect is stronger in one of the conditions. We split the sample in half to separately examine the paper-based and online instruments. When analyzing only the respondents that completed a paper-based instrument, a one-way ANOVA reveals that tax compliance intentions are significantly different based on the presentation order ($p < .03$); participants with the tax scenario early had significantly higher tax compliance intentions versus participants that had the tax scenario later. When analyzing only the respondents that completed the instrument online, no differences in presentation order are present ($p > .70$). Furthermore, *t*-tests show that the paper-based condition with the specific tax scenario early had higher tax compliance intentions than *all* other experimental conditions (all $p < .10$).

This difference may have been because participants in the online condition only viewed one screen at the time. However, the design of the experimental instrument was such that participants in the paper-based condition viewed two pages at once; thus, in the condition in which the tax scenario was presented first, participants saw the tax scenario when they first opened the instrument. Although we do find that tax compliance intentions are higher for the paper-based scenario early condition, the overall conclusions regarding the influence of social norms on tax compliance are robust to the effects of order and form.

DISCUSSION

Limitations

The results of this study should be interpreted in light of its potential limitations. First, it is possible that respondents did not report their level of tax compliance intentions honestly, particularly since this research dealt with illegal behavior. However, given that overclaiming did not affect the results for our experienced taxpayers, the interpretation of the results does not appear to suffer from extensive social desirability bias. Furthermore, as part of an initial pilot test, we asked participants if they felt they could “answer the questions in the questionnaire truthfully (e.g., did you feel like

your identity was sufficiently protected?).” All but one participant responded “yes.” We also asked if they thought under “real” survey conditions, “respondents will feel comfortable answering the questions in the questionnaire truthfully,” and 77% of the participants responded “yes.”

Second, generalization of results should be done with caution, as participants were not obtained through a random sampling method. Our taxpayer sample was also more likely to be female, slightly older, better educated, and had a higher income than the U.S. national population. However, because higher-income taxpayers may encounter more opportunities to cheat, understanding the factors influencing their tax compliance behavior is important.

Conclusions

The purpose of this study was to develop theoretically valid scales that measure social norms related to taxpayer compliance. Furthermore, we test the reliability of these scales and their influence on compliance intentions by considering the effects of social desirability and presentation order and presentation form. By drawing on [Cialdini and Trost’s \(1998\)](#) taxonomy of social norms, we provide more specificity to what is meant when researchers attribute an effect to “social norms.” Our factor analysis results identified the four distinct social norm constructs consistent with [Cialdini & Trost’s \(1998\)](#) taxonomy: personal norms, subjective norms, injunctive norms, and descriptive norms. All four of these social norm constructs were significantly related to tax compliance behavior. In addition, we find that items measuring scenario-specific norm constructs appear to be better than general items at explaining tax compliance intentions, consistent with attitude theory ([Ajzen & Fishbein, 1980](#)).

We also introduced an overclaiming scale to behavioral accounting researchers, which attempts to consider how social desirability may affect our findings. The use of an overclaiming scale is superior to many other methods of controlling for social desirability bias because socially desirable responses on the overclaiming scale are not confounded by true responses ([Randall & Fernandes, 1991](#)). We found that overclaiming was not related to experienced taxpayers’ compliance decisions, but it did influence results in our pilot test of undergraduate students. Future researchers should thus be particularly cautious about using undergraduate students in studies of taxpayer decision making.

Furthermore, we found an order effect that was driven by the paper-based condition. Specifically, participants in the paper-based condition who received the scenario early responded with higher levels of tax compliance intentions than participants that received the scenario later in the experiment; this effect was not present in the online condition. While this result did not affect our interpretation of our social norm scales or their relationship with taxpayer compliance, this does suggest that online administration of experiments may offer more stability and control than paper-based administration. Future researchers should use caution when using a non-controlled paper-based instrument to assess sensitive areas such as illegal tax evasion behavior. An additional advantage of online administration is experimenter control over participants' ability to go back to prior screens when completing the instrument. Finally, our study reinforces the significant improvement in survey-based response rates that can be obtained by sending out follow-up reminders.

In summary, we achieved the objective of our study, which was to create reliable and theoretically valid scales representing a comprehensive set of social norm constructs. Given the large degree of explanatory power provided by these scales, we believe future tax compliance researchers should include these scales when studying tax compliance behavior as they continue to develop a comprehensive understanding of the factors that influence tax compliance behavior. Furthermore, exploring ways to enhance taxpayers' social norms toward tax compliance is an important next step for researchers who are interested in studying ways to improve voluntary tax compliance.

NOTES

1. In the present study, we do not consider how each of these constructs is related. Instead, we concentrate on developing distinct measures of each. However, in a related study, Bobek et al. (2011), using the measures developed in the present study, further this research by considering, among other things, these interrelationships.

2. Cialdini and Trost (1998) discuss various motivations that individuals have to comply with social norms including building and maintaining social relationships, which they relate to injunctive and subjective norms, and making effective decisions, which they relate to descriptive norms. On the other hand, personal norms are a primarily *internal* influence on decision making that relate to our desire to live up to self-based standards or expectations (Schwartz, 1977).

3. All the experimental cells contained between 50 and 53 participants, except the paper-based group that received the scenario first contained 64 participants.

4. Two instructors of an introductory accounting class at a large southeastern university in the spring and summer of 2008 offered extra credit to their students for providing the names, mailing address, and email addresses of eligible participants that had expressed an interest in participating in the study. We emphasized that the responses were confidential, and that the participants would never be associated with the study or have individual responses reported. The extra-credit points awarded to students were tied to their *identification* of eligible participants, rather than to participants' actual completion of the study.

5. On the basis of correlation analysis, none of the demographic, control, independent, or dependent variables differed significantly (all $p > .10$) between the two groups of respondents.

6. In a separate (unreported) analysis, we construct two sets of scales for the four social norm constructs consisting of general and scenario-specific items separately. The scales with scenario-specific items only have far superior psychometric properties than the general items only scales, further confirming this recommendation.

7. While we performed three pilot tests in developing the scale, only the third pilot test included the same items that were included in the final scale developed for this study. Results from this study's factor analysis are similar to the results obtained during the third pilot test. We pilot-tested the items used in the four social norms scales with a sample of 54 undergraduate taxation students. As shown in the last two columns of Table 2, the majority of the items included in the final results of the factor analysis in the pilot study were also retained in the final scale for the experienced taxpayers.

8. We also performed factor analysis using promax rotation which allows for correlated factor scores. The items loading on each factor were identical to the varimax rotation method. The promax factor scores are significantly correlated with each other (Pearson correlation coefficients ranging from .362 to .570 all significant at $p < .001$). We use the orthogonal (varimax rotation) factor scores in all our reported regression analyses to illustrate how each type of social norm independently influences tax compliance.

9. For example, the television show category asked participants "How FAMILIAR are you with each of the following TV programs ..." The three real programs were: "The Greatest American Hero", "Grey's Anatomy", and "Everybody Loves Raymond". The two fake programs were: "The Adventures of Johnny" and "Chicago Heat". Participants responded on a 5-point scale, with 1 = not at all, and 5 = very. The responses to the real programs are ignored for purposes of determining the participants' overclaiming score.

10. Scores ranged from 10 to 31, and most scores were 12 or under.

11. To test the robustness of these correlation findings, we also reran the analysis using the correlated factor scores (i.e., obtained via promax rotation). The overclaiming scale is still significantly negatively correlated with the personal norms factor score and descriptive norms factor score (both $p < .05$), but is no longer significantly correlated with subjective norms ($p > .50$). Thus, the finding of the positive correlation between the subjective norms factor score and the overclaiming scale is not robust to alternative specifications of the social norm variables.

ACKNOWLEDGMENTS

We are very grateful for the helpful feedback from Robin Roberts and for the assistance of Jillian Phillips and Marcy Hampton in the collection of the experimental data.

REFERENCES

- Alexander, R. M., Blay, A. D., & Hurtt, R. K. (2006). An examination of convergent validity between in-lab and out-of-lab Internet-based experimental accounting research. *Behavioral Research in Accounting*, 18, 207–217.
- Azjen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice Hall, Inc.
- Alicke, M. D., Klotz, M. L., Breitenbecher, D., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Interpersonal Relations and Group Processes*, 68(5), 804–825.
- Alm, J., McClelland, G., & Schulze, W. (1999). Changing the social norm of tax compliance by voting. *KYKLOS*, 52(2), 141–171.
- Alm, J., & McKee, M. (1998). Extending the lessons of laboratory experiments on tax compliance to managerial and decision economics. *Managerial and Decision Economics*, 19, 259–275.
- Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature*, 36(2), 818–860.
- Asare, S. (1992). The auditor's going-concern decision: Interaction of task variables and the sequential processing of evidence. *The Accounting Review*, 67(2), 379–393.
- Barger, S. (2002). The Marlowe-Crowne affair: Short forms, psychometric structure and social desirability. *Journal of Personality Assessment*, 79(2), 286–305.
- Blanthorne, C., & Kaplan, S. (2008). An egocentric model of the relations among the opportunity to underreport, social norms, ethical beliefs, and underreporting behavior. *Accounting, Organizations and Society*, 33(7/8), 684–703.
- Bobek, D., Hageman, A., & Kelliher, C. (2011). *The social norms of tax compliance: An experimental investigation*. Working Paper. University of Central Florida.
- Bobek, D., & Hatfield, R. (2003). An investigation of the theory of planned behavior and the role of moral obligation in tax compliance. *Behavioral Research in Accounting*, 15, 13–38.
- Bobek, D., Roberts, R. W., & Sweeney, J. T. (2007). The social norms of tax compliance: Evidence from Australia, Singapore, and the United States. *Journal of Business Ethics*, 74(1), 49–64.
- Bryant, S. M., Hunton, J. E., & Stone, D. N. (2004). Internet-based experiments: Prospects and possibilities for behavioral accounting research. *Behavioral Research in Accounting*, 16, 107–129.
- Cialdini, R., & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In: D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology*. Boston, MA: Oxford University Press.

- Davis, J., Hecht, G., & Perkins, J. (2003). Social behaviors, enforcement, and tax compliance dynamics. *The Accounting Review*, 78(1), 39–69.
- Hanno, D. M., & Violette, G. R. (1996). An analysis of moral and social influences on taxpayer behavior. *Behavioral Research in Accounting*, 8, 57–75.
- Kaplan, S., Newberry, K., & Reckers, P. M. (1997). The effect of moral reasoning and educational communications on tax evasion intentions. *Journal of the American Taxation Association*, 19(2), 38–54.
- Kaplan, S. E., Reckers, P. M. J., & Roark, S. J. (1988). An attribution theory analysis of tax evasion related judgments. *Accounting, Organizations and Society*, 13(4), 371–379.
- Kennedy, J. (1993). Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research*, 31(2), 231–245.
- Leite, W., & Beretvas, N. (2005). Validation of scores on the Marlowe-Crowne social desirability scale and the balances inventory of desirable responding. *Educational and Psychological Measurement*, 65(1), 140–154.
- Mason, J. D., & Levy, L. G. (2001). The use of the latent construct method in behavioral accounting research: The measurement of client advocacy. *Advances in Taxation*, 13, 123–139.
- Murphy, K. (2004). *Procedural justice, shame, and tax compliance*. Working Paper. Centre for Tax System Integrity.
- Pei, B., Reckers, P., & Wyndelts, R. (1992). Tax professionals belief revision: The effects of information presentation sequence, client preference and domain experience. *Decision Sciences*, 23(1), 175–199.
- Pommerehne, W., Hart, A., & Frey, B. (1994). Tax morale, tax evasion, and the choice of tax policy instruments in different political systems. *Public Finance*, 49(Suppl.), 52–69.
- Randall, D. M., & Fernandes, M. (1991). The social desirability response bias in ethics research. *Journal of Business Ethics*, 10(11), 805–817.
- Scholz, J., & Lubell, M. (1998). Adaptive political attitudes: Duty, trust and fear as monitors of tax policy. *American Journal of Political Science*, 42, 398–417.
- Scholz, J., & Pinney, N. (1995). Duty, fear, and tax compliance: The heuristic basis of citizenship behavior. *American Journal of Political Science*, 39(2), 490–512.
- Schwartz, S. H. (1977). Normative influence on altruism. In: L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 221–279). New York, NY: Academic Press.
- Stalans, L. J., Kinsey, K. A., & Smith, K. W. (1991). Listening to different voices: Formation of sanction beliefs and taxpaying norms. *Journal of Applied Social Psychology*, 21(2), 119–138.
- U.S. Department of Treasury. (2009). Update on reducing the tax gap and improving voluntary compliance (July 8). Retrieved from http://www.irs.gov/pub/newsroom/tax_gap_report_final_version.pdf
- Weisbach, D. A., & Plesko, G. A. (2007). A legal perspective on unanswered questions in tax research. *The Journal of the American Taxation Association*, 29(2), 107–113.
- Wenzel, M. (2004). An analysis of norm processes in tax compliance. *Journal of Economic Psychology*, 25, 213–228.
- Wenzel, M. (2005). Motivation or rationalization? Causal relations between ethics, norms, and tax compliance. *Journal of Economic Psychology*, 26, 491–508.

QUANTIFYING INTUITIONS ABOUT RISK: COMPARING PUBLIC ACCOUNTING FIRM PARTNERS PERCEIVED AS “RISKY” AND “NON-RISKY”

Anne Norris, Deborah Saber, David Morrison,
Daven Morrison and Greg Trompeter

ABSTRACT

The purpose of this study is to identify a psychological profile for public accounting firm partners who are likely to place the partnership and client shareholder at risk. Proprietary data from an executive counseling firm provided a unique opportunity to compare two groups of partners: those identified by their senior partners as placing the firm at risk (n = 31) and those not so identified (n = 64). The groups were compared using psychological measures, lifestyle measures, personal measures, and work history variables. Results found no significant measurable difference between the audit partners who were identified as posing a risk and those not so identified. This suggests that specific factors cannot lead a partner to engage in risky behaviors, but rather several, in combination, may be necessary. Implications for research include learning more about concepts such as resistance to temptation, motivation, and rationalization.

Advances in Accounting Behavioral Research, Volume 14, 67–89
Copyright © 2011 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1475-1488/doi:10.1108/S1475-1488(2011)0000014006

Implications for practice are to focus on structuring business practices to provide early warning signs and minimize opportunities to engage in risky behavior. Continued and increased diligence in the client screening and client continuation and review process remain essential for best practices.

INTRODUCTION

The external auditor is a central player in detecting material fraud. The Association of Certified Fraud Examiners (ACFE) reports that external auditors detect nearly 16% of frauds over \$1 million (ACFE, 2009). External auditors are inherently vulnerable to entanglement in fraudulent situations because their outsider status makes them reliant on client management for access and information. Consequently, partners must maintain strong client management roles while running an effective and efficient audit engagement. They must maintain strong technical accounting and auditing skills, as well as strong interpersonal skills, lest place the partnership and client shareholders at risk.

Managing this risk requires certain business skills and personality characteristics. Individual partners may possess most *but not all* of these needed skills and characteristics. Alternatively, a partner may be in an emotional state (e.g., depression) or engaged in lifestyle behaviors (e.g., substance use) that impair their performance. Regardless, this creates an internal risk management problem for the audit firm.

To address this problem, experienced and perceptive partners charged with risk management responsibilities may intuitively identify skill and personality characteristic deficits, emotional states, or lifestyle behaviors that could put the firm at risk. They seek to manage this risk on a personal level through support and supervision. However, such individualized approaches to internal risk management can fail when partners providing risk management are promoted, retire, take on additional responsibilities inside their organization, become distracted by personal or professional events, or in any way have their judgment impaired.

In this study, we attempt to identify a psychological profile for partners who are likely to place the partnership and client shareholders at risk. Such a profile would allow firms to move away from the individualized approach that is so heavily dependent on the perceptions of individual risk management partners

and (a) allow systematic management of internal risk by targeting support and supervision to those in need and (b) create more effective risk management and fraud prevention. Specifically, in this study, we use proprietary data provided by an executive counseling firm to differentiate between two subsamples of public accounting partners – one subsample consists of partners who have been involved with litigation and identified by their firm as a partner who placed their partnership at risk, the second subsample consists of partners who have not been involved with litigation and have not been identified by their risk management group as having placed their partnership at risk.

Our proprietary database, which contains confidential psychological and other data on individuals, was drawn from a population of over 500 partners in public accounting who had engaged in career counseling with an executive consultation firm.¹ The consulting firm, which consisted of industrial psychiatrists, psychologists, and executive-level strategy advisors, spent a full day interviewing and testing these partners. This generated a rich set of qualitative and quantitative data that was made available for analysis. This large set of personal data should maximize the likelihood of discovering factors that will allow researchers and firms to identify partners who could well place a firm at risk.

Specifically, we discriminate between these two groups of partners (i.e., the 31 who were noted as placing the firm at risk vs. the 64 who were not so identified) by using such scores as psychological measures (e.g., of social intelligence, field independence, executive functioning, ego functioning, feelings of hopelessness, and depression), life style measures (alcohol use, sleep), personal measures (e.g., age, number of children, marital status), and work history variables (e.g., years experience, international experience). The ability to identify high-risk partners a priori could greatly help accounting firms identify individuals who may need additional support or supervision, make better hiring and promotion decisions, and develop more effective training programs. It could also help firms match partners to specific types of clients (e.g., more vulnerable partners to less risky or less demanding clients).

Despite having access to a unique dataset, containing many psychological measures, statistical analyses could not identify factors to discriminate between the two groups. Essentially our findings support the conclusion that identifying individual factors that could make a partner more or less likely to place its firm at risk is very difficult, even with extensive psychological and personal data.

HYPOTHESIS DEVELOPMENT AND RESEARCH QUESTIONS

Successful audit partners require a combination of personality characteristics, emotional states, lifestyle behaviors, and work experiences. Personality characteristics that enable partners to accurately interpret and respond to a client's behavior include (a) social intelligence that allows partners to accurately read social situations, and (b) field independence that allows them to stand firm in their opinion in light of disagreement from others (i.e., resistance to conformity). Other personality characteristics, emotional states, and lifestyle behaviors not only impact work with clients but also affect how partners make sense of the world more generally and evaluate the quality of their own work. These include (1) ego functioning, (2) executive functioning, (3) depression, and (4) substance use and sleep. The rest of this section discusses these personality characteristics, emotional states, and lifestyle behaviors in more detail and present the research questions related to them.

Social Intelligence

Social intelligence/or knowledge about one's social world underlies the ability to anticipate an individual's response across a broad range of circumstances (Kihlstrom & Cantor, 2000). This type of intelligence enables an individual to recognize motivations, empathize with individuals, and anticipate their actions (Galinsky, Maddux, Gilin, & White, 2008; Kosmitzki & John, 1993; Wright, 2002). Partners with low social intelligence should be less able to (1) recognize when they are being manipulated and (2) anticipate when a member of a client's workforce might be tempted to act dishonestly. Accounting firms work to avoid association with manipulative/dishonest clients through rigorous client acceptance and continuing client review practices (Asare, Hackenbrack, & Knechel, 1994; Asare, Cohen, & Trompeter, 2005). However, firms often fail to detect such problems in the early stages of the auditor-client relationship. In such cases, the audit engagement partner represents the "last line of defense" for the firm in its effort to detect, deter, and/or report inappropriate activity. Unfortunately, this may be difficult for partners with low social intelligence. Thus, following from prior research in personality, we hypothesize:

H₁. Partners who are identified by their firm as posing a risk will score lower on a measure of social intelligence than those who are not so identified.

Field Independence

Field independence has been identified as personality characteristic that conveys a resistance to conformity (e.g., Kasl, Sampson, & French, 1964; Witkin, 1974). Field-independent individuals perceive objects or details as discrete from their backgrounds (Witkin, 1974). Hence, unlike field-dependent individuals, their perception is less influenced by the prevailing field or context.

Both vulnerability and resistance to conformity have been the subjects of much research in social psychology (e.g., Asch, 1952). Individuals who yield to conformity pressure, revise the meaning of the stimulus with which they are presented, assuming their prior interpretation was wrong (Allen & Wilder, 1980). This suggests that individuals whose perception is more dependent on context (i.e., field dependent) are more likely to change their opinions, when confronted by a group disagreeing with them. Such individuals should be more likely to agree that a decision or interpretation offered by the client is acceptable when it is not acceptable. Hence, research regarding conformity supports the following hypothesis:

H₂. Partners who are identified by their firm as posing a risk will score lower on a measure of field independence than those who are not so identified.

Executive Functioning

Executive functioning generally describes a group of cognitive abilities that help individuals organize, evaluate, and modify their behavior in the course of achieving future goals (Cottone, Drucker, & Javier, 2007).² Executive functioning allows one to “think outside the box” coming up with creative and new approaches to situations (Benson & Sabbagh, 2010) and helps to predict professional success (Han, Delis, & Holdnack, 2008). Of particular relevance to the present study, executive functioning has been identified both theoretically (Goldberg, 2001) and empirically as a significant predictor of greater moral reasoning ability (Cottone et al., 2007). As such, it is a factor that we examine to assess whether partners who have been identified as posing a risk have lower levels of executive functioning.

H₃. Partners who are identified by their firm as posing a risk will score lower on a measure of executive functioning than those who are not so identified.

Ego Functioning

Psychiatry research and theory argues that ego functioning should critically impact a partner's judgment which in turn would cause a partner to be viewed as posing a risk to the firm. Poor ego functioning is characterized by impulsivity and a reliance on primitive defense mechanisms such as denial (e.g., "There is no problem.") and projection (e.g., blaming someone else for one's actions instead of owning the problem – "I just used what the client told me for revenue recognition") (Freud, 1961; Vaillant, 1992). Thus, research and theory in psychiatry support the following hypothesis:

H₄. Partners who are identified by their firm as posing a risk will score lower on measures of ego functioning than those who are not so identified.

Emotional (Depression, Hopelessness) and Lifestyle (Alcohol, Sleep) Factors

Research in social, clinical, and neuropsychology argues for emotional and life style factors undermining a partner's judgment, thereby increasing the likelihood that he or she would be perceived as risky (Hartlage, Alloy, Vazquez & Dykman, 1993; Streufert, Pogash, Roache, & Severs, 1994; Wyatt & Bootzin, 1994). These factors negatively impact information processing, increasing the likelihood that a partner could miss problematic cues in the client's behavior or records. Emotional factors such as depression and hopelessness create errors in thinking which impact judgment (e.g., Abramson, Metalsky, & Alloy, 1989). Depression also limits the cognitive resources available to process needed data to solve problems and make decisions (Conway & Giannopoulos, 1993; Hartlage et al., 1993; Roškar, Zorko, Bucik, & Maruši2007). Moreover, depression contributes to feelings of self-doubt that could undermine a partner's ability to follow through on "hunches" that something could be wrong (Schwarzer, 1996).

Alcohol and sleep greatly affect work performance (Streufert et al., 1994; Wyatt & Bootzin, 1994). Arnedt, Owens, Crouch, Stahl, and Carskadon (2005) argue that missing or having seriously disrupted sleep every fourth or fifth night is as cognitively disabling as having 0.04 to 0.05 g% blood alcohol concentration (BAC). Alcohol at 0.05 g% BAC (3–4 standard drinks) increases self-confidence and diminishes attention, judgment, and control (Council on Scientific Affairs, 1986). It also negatively affects management strategy and planning (Streufert et al., 1994).

Sleeping less than one’s optimal nightly amount (8–9 hours a night on average, depending on the individual) can cause decreased ability to think or concentrate as well as daytime sleepiness (Bonnet & Arand, 2003). Daytime sleepiness relates to an inability to recall information presented verbally in the three minutes before falling asleep (Wyatt & Bootzin, 1994). Research indicates that the effect of days of insufficient sleep builds over time, eventually producing the neurocognitive deficits seen in severe or full sleep deprivation (Banks & Dinges, 2007). Moreover, this cumulated effect of insufficient sleep occurs without full awareness (Goel, Rao, Durmer, & Dinges, 2009), creating a situation in which partners do not recognize that their cognitive abilities are impaired. Thus research regarding emotional and lifestyle factors supports the following hypothesis:

H₅. Partners who are identified by their firm as posing a risk will score higher on measures of depression and hopelessness and report more alcohol use and less sleep than those who are not so identified.

METHODS

Sample, Design, and Overview of Procedures Used to Generate Data Set

This secondary data analysis uses de-identified data for a sample of certified public accountants working as partners in accounting firms. The accounting firms contracted with a business consulting firm that specializes in advising and counseling executives. These counseling services were made available to the partners of the accounting firms as an employee benefit. Approximately 500 partners took advantage of this benefit. Of this group, 31 partners were identified by their accounting firm’s senior level risk management group as partners who had been involved in litigation and had placed the firm at risk. Our study uses a cohort design to compare these 31 partners with a random sample of 64 partners drawn from the larger subgroup of 470 partners who had never been identified by senior level risk management as placing the firm at risk.

All participating partners met with a business consulting firm for a day long evaluation and consultation meeting. The evaluation involved both observational and interview-based measures. Partners could be self-referred or referred by a senior partner. Both the audit firm and the business consulting firm framed the evaluation and consultation meeting as helpful and informative, rather than punitive.

The meeting began with an orientation to the day provided by the head of the consulting firm or his associate, both of whom are licensed, board certified, industrial psychiatrists. The purpose of the orientation was to explain policies and procedures, including confidentiality, and to put the partner at ease. Our review of anecdotal notes in the interview data indicate that partners were sometimes anxious during this orientation, but became relaxed and reported by the end of the day that they found the experience helpful and non-threatening.

Measures

Participating partners completed a battery of standardized, self-administered assessment instruments and were interviewed by psychiatric and counseling professionals who complete assessment instruments and documented their observations. All measures were selected for use by the head of the consulting firm and his associate, psychiatrists with more than 50 years experience combined in industrial psychiatry, including a fellowship at the Mayo Clinic. Only those measures used in the analyses presented here are discussed below.³

Demographic Characteristics

A series of demographic items were available to provide work and personal history information, including age, gender, marital history, and number of children.

Social Intelligence

Two measures of social intelligence were used – the Wechsler Adult Intelligence Scale (WAIS) Picture Arrangement subtest and Comprehension subtests (Lezak, 1995; Rapaport, Gill, & Schafer, 1968; Sippes, Berry, & Lynch, 1987).

The WAIS Picture Arrangement subtest assesses social and nonverbal reasoning abilities (Rapaport et al., 1968). This test involves arranging pictures in a particular order to tell a story. To arrange the pictures, respondents must find a logical sequence of 10 events (i.e., a series of 10 pictures), evaluate social situations, distinguish between relevant and irrelevant details, anticipate consequences of actions, and plan. Scores range from 0 to 11, the higher the score, the greater the ability to accurately interpret social situations. Internal consistency reliability for the Picture Arrangement subtest is .74 with test re-test reliability of .69 (The Psychological Corporation, 1997).

The WAIS Comprehension subtest measures common sense, reasoning, and social judgment in practical situations and is influenced by social experience. It contains 21 items, representing three types of questions: (1) what should be done; (2) why it should be done (i.e., provide explanation for some phenomenon); (3) what does it mean (i.e., define a proverb). Emotional difficulties frequently result in a lowering respondent’s overall score by interfering with judgment. Internal consistency reliability for the WAIS Comprehension subtest is .84 with test re-test reliability of .81 (The Psychological Corporation, 1997).

Field Independence

The Group Embedded Figure Test (GEFT; Oltman, Raskin, & Witkin, 2003) is a standardized measure of field Independence/dependence. The test requires finding simple forms which are embedded in larger figures. It contains 25 figures or items, and respondents receive 5 minutes per section of the test. The test contains three sections. The first section determines whether the respondent understands the test directions. The second and third sections are used to determine field independence/dependence. The score is the total number of simple forms traced in the second and third sections of the test in the time period allowed. Higher scores indicate more field independence: The longer respondents take to find the simple forms, the fewer forms identified. Hence, the more field dependent is his perception. GEFT reliability and validity is supported in university students (Witkin, 1960; Witkin et al., 1954; Witkin, Oltman, Raskin, & Karp, 1971) and the test has been extensively used for psychological assessment purposes for over 20 years.

Executive Functioning

The Halstead Booklet Category Test (McC Campbell & DeFilippis, 1979) and the WAIS Similarities subtest were used to assess executive functioning. The Halstead Booklet Category Test (McC Campbell & DeFilippis, 1979) assesses the following components of executive functioning: concept formation, reasoning, abstraction, and cognitive flexibility. The test is an adaption of Jarvis and Barth’s (1984) Halstead Category Test that was originally created with visual slides. The test contains seven subtests, each representing a unique idea or principle illustrated by a set of geometric patterns or figures for the respondent to identify to respond correctly. Respondents receive feedback from the test administrator’s statement that their response was either correct or incorrect. The test takes approximately 30 minutes to complete and is scored by calculating the number of errors

a respondent makes. Reliability of the Halstead Booklet Category Test is supported by test re-test values ranging from .59 to .87 (Franzen, 1989).

The 14-item WAIS Similarities subtest assesses the abstract verbal reasoning component of executive functioning by asking respondents to indicate similarities between pairs of concepts that vary in abstractness (e.g., horse and cow; east and south). The easier (more concrete) items measure previously learned associations. Difficult (more abstract) items measure abstract thinking. Scores range from 0 to 14. A high score indicates high abstract verbal reasoning. Internal consistency reliability for the similarities subtest is good (.86 with test re-test reliability of .83; The Psychological Corporation, 1997).

Ego Functioning

We used two measures of ego functioning, a simple, single-item measure of impulsivity and Bellak's Ego Function Assessment Survey (EFAS, Bellak & Goldsmith, 1984). The measure of impulsivity was the partner's response to the question: "What would you do if you were the first person to see smoke and fire in a movie theatre?" Individuals who yelled "fire" in response to this question were classified as impulsive.

The EFAS (Bellak & Goldsmith, 1984) assesses different ego functions. Each function (i.e., reality testing, judgment, sense of reality, regulation, object relations, thought processes, adaptive regression, defensive functioning, stimulus barrier, autonomous functioning, synthesis and integration, and mastery and competence) contains 10-items. Response options range from 0 (never) to 3 (always). In addition, there is a N/A response option. Higher scores indicate more adaptive (more mature, better) ego functioning (Juni, Stack, & Burton, 2000). The EFAS measure was formulated in a five-year project based on extensive literature review, clinical interviews, clinical psychological tests, and experimental procedures from psychological laboratories involving schizophrenics, neurotics, and normal subjects (Bellak, Chassan, Gediman, & Hurvich, 1973; Bellak & Hurvich, 1969; Bellak, Hurvich, Silvan, & Jacobs, 1968). Internal consistency reliability ($\alpha = .61-.89$) and validity is supported in a non-clinical urban population. Stability is supported by moderate to high test-retest correlations ($r = .61-.92$; Juni & Straehle, 2002). Coefficient alpha for the EFAS (.83) in the present sample indicates good internal consistency reliability.

*Emotional and Lifestyle Factors: Depression, Hopelessness,
Alcohol Use, and Sleep*

Depression

Beck’s 21-item depression inventory (BDI, Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) measures depression based on Beck’s cognitive theory of depression. This theory asserts that a negative triad of schemas (self, world, and future) in combination with stressors leads to cognitive distortions which in turn cause depression.

Each item in the BDI contains four statements arranged in order of severity from 0 (least) to 3 (most). Respondents choose the “true” statement. Responses are summed, with scores ranging from 0 to 60. A score above 16 indicates moderate to severe depressive symptomatology. Internal consistency reliability is excellent in both psychiatric (alpha = .86; Beck, Steer, & Carbin, 1988) and non-psychiatric populations (alpha = .81–.90; Beck et al., 1988; Magalhaes, Pinheiro, Horta, Pinheiro, & Da Silva, 2008),⁴ and validity is supported in a postpartum study (Magalhaes et al., 2008). Stability is supported in U.S. ($r = .60$ –.90; Beck et al., 1988) and in non-U.S. populations (e.g., Korean elderly population; $r = .60$; Jo, Park, Jo, Ryu, & Han, 2007). Validity is supported for psychiatric and non-psychiatric populations (Beck et al., 1988).

Hopelessness

The 20-item Beck’s Hopelessness Scale (Beck, 1974) measures the respondent’s degree of pessimism, which is one of Beck’s negative triad components in depression. Each item is answered as true or false (0 or 1). Nine items are keyed as false and 11 as true, with the 9 items being reverse scored (Beck, Brown, Berchick, Stewart, & Steer, 2006). Responses are summed, and a score exceeding 14 indicates psychological distress (Beck et al., 1961). Internal consistency reliability was excellent in a patient population (alpha = .93; Beck, Weissman, Lester, & Trexler, 1974). Internal consistency reliability (alpha = .80)⁵ and stability were excellent in college student populations with test-retest correlation ranges from $r = .87$ to .89 in men (Holden & Fekken, 1988). Validity was supported by high hopelessness scores associated with both suicide intent (Beck, Kovacs, & Weissman, 1975; Beck, Steer, Kovacs, & Garrison, 1985) and completed suicides (Lester, Beck, & Mitchell, 1979).

Alcohol Use and Sleep

Partners were asked about daily and weekly alcohol use (number of drinks) and average hours of sleep per night. They were also asked whether they felt rested (yes, no) as a measure of sleep quality, and to estimate the number of minutes to fall asleep as an indicator of sleep problems.

RESULTS

All observations represent audit partners with a single international public accounting firm. Audit partners in the “risky” group had been involved in litigation against their firm and had been identified by senior-level risk management partners in their firm as a partner who had placed the partnership at risk. Audit partners were defined as “non-risky” if they were not so identified. SPSS was used to randomly select cases from the large subgroup of audit partners ($n = 500$) in the data set that were defined as “non-risky”.

The two groups of partners (risky, non-risky) were compared on several personal and psychological characteristics and measures with a series of chi-square and *t*-tests. Fisher’s Exact was also used for chi-square analyses where expected frequencies were below 5. No effort was made to control or limit the type 1 error rate given the exploratory nature of these analyses. A type 2 error was deemed more serious given the limited state of the science.

Missing data were substantive for some of the measures in the data set, particularly for drinking behavior and the Beck’s Hopelessness Scale. However, these data were not missing due to partners refusing to answer the question, but to the measure not consistently being asked as part of the interview. Also, missing data did appear random, not systematic. It was unrelated to partner characteristics, but may have been related to interviewer characteristics. (Such data are unavailable.)

Demographic Characteristics

Data in [Table 1](#) pertain to demographic characteristics. Panels A through C present results for personal history, work history, and medication/alcohol use, respectively. Overall, these data suggest that these two groups of partners are quite similar. Statistics shown in panel A indicate that partners were very similar in terms of gender (predominately male), marital status (predominately married), number of children (2 on average), and history of divorce (less than 20%). Work history variables presented in panel B are

Table 1. Demographic Characteristics.

	Non-Risky Mean (SD)	Risky Mean (SD)	Significance Level
Panel A: Personal history			
Age	<i>n</i> = 64 41.02 (5.16)	<i>n</i> = 31 42.19 (5.24)	<i>p</i> = .30
Gender	<i>n</i> = 64	<i>n</i> = 31	
Male	93.8% (60)	100.0% (31)	<i>p</i> = .30 (Fisher's exact)
Female	6.2% (4)	0.0% (0)	
Marital status	<i>n</i> = 64	<i>n</i> = 31	<i>p</i> = .75 (Fisher's exact; married vs. other)
Married	85.9% (55)	90.3% (28)	
Divorced	6.2% (4)	6.5% (2)	
Unmarried	4.7% (3)	0.0% (0)	
Separated	3.1% (2)	3.2% (1)	
History of marital separation	<i>n</i> = 58 1.7% (1)	<i>n</i> = 30 10.0% (3)	<i>p</i> = .11 (Fisher's exact)
History of divorce	<i>n</i> = 58 19.0% (11)	<i>n</i> = 30 16.7% (5)	<i>p</i> = .79
Children	<i>n</i> = 64 2.11 (1.25)	<i>n</i> = 31 2.32 (1.42)	<i>p</i> = .46
Panel B: Work history			
International work	<i>n</i> = 55 5.5% (3)	<i>n</i> = 31 12.9% (4)	<i>p</i> = .26 (Fisher's exact)
Number of job transfers	<i>n</i> = 47 1.66 (1.81)	<i>n</i> = 31 1.65 (1.96)	<i>p</i> = .99 (Mann-Whitney U d/t severe skew)
Years with the company**	<i>n</i> = 57 16.14 (5.79)	<i>n</i> = 31 19.91 (5.47)	<i>p</i> = .004
Panel C: Medication and smoking			
Allergy medications***	<i>n</i> = 58 0% (0)	<i>n</i> = 31 19.4% (6)	<i>p</i> = .001 (Fisher's exact)
Hypertensive and cardiovascular medications	<i>n</i> = 58 8.6% (5)	<i>n</i> = 31 9.7% (3)	<i>p</i> = 1.00 (Fisher's exact)
Psychologically related medications	<i>n</i> = 58 8.6% (5)	<i>n</i> = 31 3.2% (1)	<i>p</i> = .66 (Fisher's exact)
Other medications	<i>n</i> = 58 8.6% (5)	<i>n</i> = 31 9.7% (3)	<i>p</i> = 1.00 (Fisher's exact)
Smoking	<i>n</i> = 45	<i>n</i> = 27	
Currently smokes	20.0% (9)	22.2% (6)	<i>p</i> = .82
Packs per day	.32 (.72)	.44 (.88)	<i>p</i> = .52

p* < .05; *p* < .01; ****p* < .001.

also quite similarity – the only difference is that risky partners had been with the firm for, on average, more years than had non-risky partners (19.91 vs. 16.14 years, $p < .01$). Finally, panel C also suggests little difference between the two groups, with the only significant difference being with respect to the use of allergy medications (19.4% of the risky partners vs. 0% of the non-risky partners; $p < .01$). Overall, these findings portray that the two groups are quite comparable.

*Social Intelligence, Field Independence, Executive Functioning,
Ego Functioning*

Per the results in Table 2, no significant differences are found between the two audit partner groups in social intelligence, field independence, or executive functioning ($p \geq .17$). Thus, there is insufficient support for H_1 – H_3 . Furthermore, results presented in Table 3 do not reveal any significant differences between the two groups of partners regarding ego functioning ($p \geq .33$). Thus, the analyses do not provide support for H_4 .

In general, the scores on these measures indicate high levels of social intelligence, field independence, executive functioning, and ego functioning. They speak to these partners' cognitive abilities and ability to lead as well as be a part of a team. These traits/abilities are all essential to be a successful audit partner. Individuals who are exceptionally low on these measures likely left the firm before becoming partners or failed to be promoted to partner. These results suggest that these measures may lack the sensitivity

Table 2. Social Intelligence, Field Independence, and Executive Functioning.

Variables	Non-Risky Mean (SD)	Risky Mean (SD)	Significance Level
Social intelligence	$n = 64$	$n = 31$	
WAIS comprehension	13.16 (2.98)	12.97 (2.68)	$p = .77$
WAIS picture arrangement	5.09 (5.71)	6.84 (5.95)	$p = .17$
Field independence (GEFT score)	$n = 64$ 1.58 (4.30)	$n = 31$ 2.19 (5.08)	$p = .54$
Executive functioning	$n = 64$	$n = 31$	
Halstead booklet score	12.56 (16.42)	13.58 (14.39)	$p = .77$
WAIS similarities	13.16 (2.73)	12.77 (1.98)	$p = .49$

Table 3. Ego Functioning.

Variables	Non-Risky Mean (SD)	Risky Mean (SD)	Significance Level
Impulsivity	<i>n</i> = 64 9.4% (6)	<i>n</i> = 31 12.9% (4)	<i>p</i> = .72 (Fisher 's exact)
Bellack EFAS			
Mastery competence	<i>n</i> = 57 2.51 (.31)	<i>n</i> = 30 2.54 (.28)	<i>p</i> = .61
Object relations	<i>n</i> = 47 1.56 (.20)	<i>n</i> = 26 1.58 (.21)	<i>p</i> = .73
Thought processes	<i>n</i> = 46 1.48 (.25)	<i>n</i> = 28 1.47 (.33)	<i>p</i> = .90
Autonomous functioning	<i>n</i> = 47 1.60 (.25)	<i>n</i> = 29 1.57 (.25)	<i>p</i> = .65
Stimulus barrier	<i>n</i> = 48 1.47 (.31)	<i>n</i> = 27 1.46 (.26)	<i>p</i> = .96
Regulation and control	<i>n</i> = 33 1.66 (.26)	<i>n</i> = 21 1.73 (.26)	<i>p</i> = .33
Defensive functioning	<i>n</i> = 41 1.44 (.28)	<i>n</i> = 26 1.40 (.26)	<i>p</i> = .51
Adaptive regression	<i>n</i> = 45 2.14 (.36)	<i>n</i> = 26 2.14 (.35)	<i>p</i> = .99
Sense of reality	<i>n</i> = 36 1.22 (.19)	<i>n</i> = 21 1.17 (.20)	<i>p</i> = .35
Reality testing	<i>n</i> = 36 1.26 (.20)	<i>n</i> = 19 1.22 (.26)	<i>p</i> = .50
Synthesis and integration	<i>n</i> = 56 2.49 (.24)	<i>n</i> = 30 2.52 (.28)	<i>p</i> = .53
Judgment	<i>n</i> = 46 1.54 (.27)	<i>n</i> = 25 1.58 (.30)	<i>p</i> = .60

needed to differentiate between partners who put their firm at risk and those who do not.

Emotional and Lifestyle Factors

As shown in Table 4, except for hopelessness little or no difference is observed for emotional and lifestyle factors, suggesting little support for H₅. Those in the risky group had higher scores on Beck’s Hopelessness Scale

Table 4. Emotional and Lifestyle Factors.

Variables	Non-Risky Mean (SD)	Risky Mean (SD)	Significance Level
	<i>n</i> = 64	<i>n</i> = 20	
Hopelessness scale	1.61 (2.15)	3.60 (3.89)	<i>p</i> = .004
Score >14	0% (0)	3.2% (1)	<i>p</i> = .33
	<i>n</i> = 64	<i>n</i> = 31	
Beck's depression inventory	8.41 (6.67)	8.42 (7.89)	<i>p</i> = .99
Score >16	12.5% (8)	6.5% (2)	<i>p</i> = .49
Drinking			
Number of drinks per day	<i>n</i> = 19	<i>n</i> = 9	
Median	1.92 (.71)	1.78 (1.48)	<i>p</i> = .73
(Range)	2.00 (0 to 3.00)	2.00 (0 to 4.00)	
Number of drinks per week	<i>n</i> = 26	<i>n</i> = 13	
Median	5.77 (6.81)	4.31 (6.18)	<i>p</i> = .52
(Range)	4.50 (0 to 36.00)	3.00 (0 to 24.00)	
Number of minutes it takes to fall asleep	<i>n</i> = 52	<i>n</i> = 26	
	17.26 (20.40)	12.83 (13.08)	<i>p</i> = .32
Feeling rested	<i>n</i> = 56	<i>n</i> = 30	
	32.14% (18)	36.67% (11)	<i>p</i> = .67
Numbers of hours of sleep	<i>n</i> = 56	<i>n</i> = 30	<i>p</i> = .88
Median	6.64 hrs. (.90)	6.67 hrs. (.84)	
(Range)	7 hours (3.5–8.0)	7 hours (5.0–8.0)	

(mean score = 3.60) than the non-risky group (mean score = 1.61; $p < .005$), albeit both scores are somewhat low. However, while the raw scores differ significantly, from a clinical perspective, the means from both groups are well under the score of 14, which is the cutoff benchmark of psychological distress. Furthermore, the proportion of each group scoring above 14 does not differ significantly ($p = .33$). Thus, while statistically significant, these results argue against the difference in group means having clinical significance.

Risky partners did not drink more than non-risky partners. Neither group reported heavy drinking (males drinking 5 or more drinks a day; women drinking 4 or more drinks a day; Wechsler, Dowdall, Davenport, & Rimm, 1995). Both groups had an average of slightly less than 7 hours of sleep a night (see also Table 4).

DISCUSSION

This analysis investigated the personal and psychological profiles of audit partners who had been involved in litigation and identified by the senior-level risk management partners in their accounting firm as “high risk” partners. Uniquely, these partners were compared with a cohort of their peers who had not been so identified. The only difference was that the higher risk group of partners had been with the firm for a slightly longer period of time.

Our study results are particularly striking for the lack of differences between audit partners who were identified as posing a risk and those not so identified. We observed no measurable differences in social intelligence, executive functioning, ego functioning, mental health, etc. These audit partners also do not differ in terms of work history, personal history, smoking, or alcohol use. They do not appear more impulsive. In short, across a battery of measures these two groups of partners appear very similar. Thus, this chapter suggests that there is nothing unique within a large range of factors that is sufficient to lead a supervising partner to put the organization at risk by engaging with a client in a risky way. In retrospect, this fits the experience of clinicians (including two of the authors) who find that rarely is one specific factor sufficient, but rather several are necessary.

The lack of significant findings is not likely due to a lack of statistical power. A power analysis indicates that we had sufficient power to detect medium or larger effect sizes ($d > .50$) for two-tailed significant tests at an alpha of .05, assuming a power of .80 in a majority of our analyses. Additionally, an inspection of means and percents revealed small differences (at most) between the two groups. Such small differences are unlikely to impact greatly behavior, given the myriad influences impinging on an individual in a work setting.

Some psychological differences may also exist between these two groups of partners that were not captured in our analysis. Our analysis is also limited by the collected data, which measured a broad variety of psychological characteristics (e.g., ego functioning, executive functioning, social intelligence); but, the two groups could have differed in other unmeasured respects.

Some of our measures could have been insensitive to group differences present in the data. There is debate in the literature regarding the ability of the WAIS Picture Arrangement and Comprehension subtests to measure social intelligence (e.g., [Campbell & McCord, 1996](#)), and little agreement exists about which cognitive abilities should be measured as indicators for executive functioning ([Salthouse, 2005](#)). Hence, other measures might have been more sensitive to the differences between these two groups. However,

the measures in this study were selected by clinicians with extensive training and experience (e.g., over 50 years of experience combined) in assessing an individual's potential to be successful in a high-level, professional, work environment. They have used these measures extensively and with good success in their clinical practice.

Despite such limitations, these findings affect auditing research and practice. First, they suggest that factors such as executive functioning, resistance to conformity (field independence), social intelligence, and ego functioning may not be key predictors of fraud in a well-educated, relatively affluent sample, such as a sample of audit partners from a large accounting firm. This is particularly interesting because our measures of executive functioning focused on attributes such as common sense, reasoning, and social judgment in practical situations; and, we were able to include measures of social intelligence in this study. These are the very qualities that one would expect from an audit partner.

One might expect a lack of differences on these measures but still anticipate a difference in field independence (i.e., resistance to conformity). However, the field independence scores obtained here argue against any group differences in resistance to conformity. Thus, a clear implication for research is to learn more about concepts such as resistance to temptation, motivation, and rationalization. An implication for practice is to focus on structuring situations and business practices to minimize opportunities for fraud and provide early warning signs until the science of individual differences catches up to the needs of the practice environment. At this point, insufficient science exists to justify a focus by Human Resources or supervisors on the screening for individuals who are likely to place a firm at risk using the measures included in this study.

A final implication for practice is the subjects reported chronically insufficient sleep, similar to many other U.S. professionals (Banks & Dinges, 2007; Centers for Disease Control, 2009). However, sleep researchers argue that this is a false trade-off because sleep deprivation produces impaired cognitive performance across various different tests, indicating the effect of sleep deprivation is robust to measurement error. Sleep-deprived individuals may be unaware of such deficiencies that could impair their performance.

In conclusion, despite a rich and extensive data set of many partners from a single firm, distinguishing – even in an ex post condition – audit partners that are likely to be put a firm at risk is difficult. This argues for additional research into this area. Furthermore, at least for the present, these findings argue for continued and increased diligence in the client screening and client continuation and review process.

NOTES

1. Certain public accounting firms offer executive career counseling to their partners as a part of their compensation and benefits package. The executive consultation firm that provided the data for this study has provided these services to over 500 partners. Of this population, 31 partners were subsequently identified as having involved their firm in litigation and placing their partnership at risk. Litigation and risk identification occurred after the individuals took advantage of the career consultation. These individuals were not referred to career counseling as a result of litigation. The remaining (approximately) 470 individuals in the population had not been so identified. Owing to resource constraints, examining the entire population was not deemed reasonable. Thus, after discussion with principals from the consulting firm, 64 of these partners were randomly drawn from the sub-sample of 470 for analysis.

2. This is similar to the “Collect the data, Boil it down, and Act” model of executive judgment used by the consulting firm to explain results of their findings to the partners. (See also Ramamoorti, Morrison, & Koletar, 2009.)

3. Information about the complete process can be obtained by contacting the authors.

4. Cronbach’s alpha could not be calculated for the present sample because only the total score was retained in the data file (responses to individual items are needed to calculate Cronbach’s alpha). However, reliability is assumed to be high based on (a) previous use of the measure in clinical and non-clinical populations, and (b) consistency between the data collection procedures used to generate the data for the analyses and recommendations made by the scale developers. Consistency between data collection procedures and scale developer recommendations should minimize procedure related measurement error effects that would alter reliability.

5. *Ibid.*

ACKNOWLEDGMENTS

The authors thank the participants of the Institute for Fraud Prevention December 2009 Research Summit for their helpful comments and Alan Reinstein for his careful review of this manuscript. In addition, they acknowledge the research assistance of Sara Tischhauser and the financial support of the IFP.

REFERENCES

- Abramson, L., Metalsky, G., & Alloy, L. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, 96(2), 358–372.
- ACFE. (2009). *2008 Report to the nation on occupational fraud and abuse*. Austin, Texas: Association of Certified Fraud Examiners.

- Allen, V., & Wilder, D. (1980). Impact of group consensus and social support on stimulus meaning: Mediation of conformity by cognitive restructuring. *Journal of Personality and Social Psychology*, 39(6), 1116–1124.
- Arnedt, J., Owens, J., Crouch, M., Stahl, J., & Carskadon, M. (2005). Neurobehavioral performance of residents after heavy night call vs. after alcohol ingestion. *JAMA: The Journal of The American Medical Association*, 294(9), 1025–1033.
- Asare, S., Cohen, J., & Trompeter, G. (2005). The effect of management integrity and non-audit services on client screening and staffing decisions. *The Journal of Accounting and Public Policy*, 24(6), 489–520.
- Asare, S., Hackenbrack, K., & Knechel, W. R. (1994). Client acceptance and continuation decisions. In: R. P. Srivastava (Ed.), *Auditing symposium XII: Proceedings of the 1994 Deloitte and Touche/University of Kansas symposium on auditing problems* (pp. 163–178). Lawrence, KS: University of Kansas.
- Asch, S. (1952). *Social psychology*. Englewood Cliffs, NJ: Prentice Hall.
- Banks, S., & Dinges, D. (2007). Behavioral and physiological consequences of sleep restriction. *Journal of Clinical Sleep Medicine: JCSM: Official Publication of the American Academy of Sleep Medicine*, 3(5), 519–528.
- Beck, A. (1974). The development of depression: A cognitive model. In: R. Freidman & M. Katz (Eds.), *Psychology of depression contemporary theory and research* (pp. 3–27). Washington, DC: Winston.
- Beck, A., Brown, G., Berchick, R., Stewart, B., & Steer, R. (2006). Relationship between hopelessness and ultimate suicide: A replication with psychiatric outpatients. *FOCUS*, IV(2). Retrieved on July 5, 2009, from <http://focus.psychiatryonline.org/cgi/content/full/4/2/291>
- Beck, A., Kovacs, M., & Weissman, A. (1975). Hopelessness and suicidal behavior: An overview. *Journal of the American Medical Association*, 234, 1146–1149.
- Beck, A., Steer, R., & Carbin, M. (1988). Psychometric properties of the Beck depression inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77–100.
- Beck, A., Steer, R., Kovacs, M., & Garrison, B. (1985). Hopelessness and eventual suicide: A 10-year prospective study of patients with suicide ideation. *American Journal of Psychiatry*, 142, 559–563.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561–571.
- Beck, A., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism. *The Journal of Clinical & Consulting Psychology*, 42, 861–865.
- Bellak, L., Chassan, J., Gediman, H., & Hurvich, M. (1973). Ego function assessment of analytic psychotherapy combined with drug therapy. *The Journal of Nervous and Mental Disease*, 157(0), 465–469.
- Bellak, L., & Goldsmith, L. (1984). *The broad scope of ego function assessment*. New York: Wiley.
- Bellak, L., & Hurvich, M. (1969). A systematic study of ego function. *The Journal of Nervous and Mental Disease*, 148(6), 569–585.
- Bellak, L., Hurvich, M., Silvan, M., & Jacobs, D. (1968). Toward an ego psychological appraisal of drug effects. *American Journal of Psychiatry*, 125(5), 593–604.
- Benson, J., & Sabbagh, M. (2010). Theory of mind and executive functioning: A developmental neuropsychological approach. In: P. D. Zelazo, M. Chandler & E. Crone (Eds),

- Developmental social cognitive neuroscience* (pp. 63–80). New York, NY: Psychology Press.
- Bonnet, M., & Arand, D. (2003). Clinical effects of sleep fragmentation versus sleep deprivation. *Sleep Medicine Reviews*, 7(4), 297–310.
- Campbell, J., & McCord, D. (1996). The WAIS–R comprehension and picture arrangement subtests as measures of social intelligence: Testing traditional interpretations. *Journal of Psychoeducational Assessment*, 14(3), 240–249.
- Centers for Disease Control. (2009). Perceived insufficient rest or sleep among adults – United States, 2008. *Morbidity Mortality Weekly Report*, 58(42), 1175–1179. Retrieved from <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5842a2.htm>
- Conway, M., & Giannopoulos, C. (1993). Dysphoria and decision making: Limited information use for evaluations of multiattribute targets. *Journal of Personality and Social Psychology*, 64(4), 613–623.
- Cottone, J., Drucker, P., & Javier, R. (2007). Predictors of moral reasoning: Components of executive functioning and aspects of religiosity. *Journal for the Scientific Study of Religion*, 46(1), 37–53.
- Council on Scientific Affairs. (1986). Alcohol and the driver. *JAMA*, 255, 522–527.
- Franzen, M. D. (1989). *Reliability and validity in neuropsychological assessment*. New York: Plenum Press.
- Freud, A. (1961). *The ego and the mechanisms of defense*. London, UK: Hogarth.
- Galinsky, A. D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological Science*, 19, 378–384.
- Goel, N., Rao, H., Durmer, J., & Dinges, D. (2009). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology*, 29(4), 320–339.
- Goldberg, E. (2001). *The executive brain: Frontal lobes and the civilized mind*. New York: Oxford University Press.
- Han, S., Delis, D., & Holdnack, J. (2008). *Extending the WISC-IV: Executive functioning. WISC-IV clinical assessment and intervention* (2nd ed., pp. 497–515). San Diego, CA: Elsevier Academic Press.
- Hartlage, S., Alloy, L. B., Vazquez, C., & Dykman, B. (1993). Automatic and effortful processing in depression. *Psychological Bulletin*, 113, 247–278.
- Holden, R., & Fekken, C. (1988). Test-retest reliability of the hopelessness scale and its items in a university population. *Journal of Clinical Psychology*, 44(1), 40–43.
- Jarvis, P. E., & Barth, J. T. (1984). *Halstead-Reitan Battery: An interpretative guide*. Odessa, FL: PAR.
- Jo, S., Park, M., Jo, I., Ryu, S., & Han, C. (2007). Usefulness of Beck depression inventory (BDI) in the Korean elderly population. *International Journal of Geriatric Psychiatry*, 22, 218–223.
- Juni, S., Stack, J., & Burton, J. (2000). Ego function assessment of substance abusers: Standardization and reliability. *Psychological Reports*, 87, 1185–1195.
- Juni, S., & Straehle, M. (2002). Ego function assessment of nonclinical individuals. *Psychological Reports*, 91, 679–686.
- Kasl, S., Sampson, E. E., & French, J. R. P. (1964). The development of a projective measure of the need for independence: A theoretical statement and some preliminary evidence. *Journal of Personality*, 32(4), 566–586.

- Kihlstrom, J. F., & Cantor, N. (2000). Social intelligence. In: R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 359–379). Cambridge, England: Cambridge University Press.
- Kosmitzki, C., & John, O. P. (1993). The implicit use of explicit conceptions of social intelligence. *Personality and Individual Differences, 15*, 11–23.
- Lester, D., Beck, A., & Mitchell, B. (1979). Extrapolation from attempted suicides to completed suicide. *Journal of Abnormal Psychology, 88*, 78–80.
- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York, NY: Oxford University Press.
- Magalhaes, P. V., Pinheiro, R. T., Horta, B. L., Pinheiro, K. A., & Da Silva, R. A. (2008). Validity of the Beck depression inventory in the postpartum period. *International Journal of Psychiatry in Clinical Practice, 12*(1), 81–84.
- McCampbell, E., & DeFilippis, N. (1979). The development of a booklet form of the category test: A preliminary report. *Clinical Neuropsychology, 1*(2), 33–35.
- Oltman, P. A., Raskin, E., & Witkin, H. A. (2003). *GEFT test booklet*. Menlo Park, CA: Mind Garden, Inc.
- Rapaport, D., Gill, M. M., & Schafer, R. (1968). *Diagnostic psychological testing*. New York, NY: International Universities Press.
- Ramamoorti, S., Morrison, D., & Koletar, J. (2009). Bringing Freud to fraud: Understanding the state-of-mind of the C-suite level, white collar criminal through A-B-C analysis. Institute for Fraud Prevention white paper. The Institute for Fraud Prevention at West Virginia University. Retrieved from <http://theifp.org/research-grants/IFP-Whitepaper-3.pdf>
- Roškar, S., Zorko, M., Bucik, V., & Marušić, A. (2007). Problem solving for depressed suicide attempters and depressed individuals without suicide attempt. *Psychiatra Danubina, 19*(4), 296–302.
- Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19*, 532–545.
- Schwarzer, R. (1996). Thought control of action: Interfering self-doubts. *Cognitive interference: Theories, methods, and findings* (pp. 99–115). Hillsdale, NJ; England: Lawrence Erlbaum Associates, Inc.
- Sipps, G., Berry, G., & Lynch, E. (1987). WAIS-R and social intelligence: A test of established assumptions that uses the CPI. *Journal of Clinical Psychology, 43*(5), 499–504.
- Streufert, S., Pogash, R., Roache, J., & Severs, W. (1994). Alcohol and management performance. *Journal of Studies on Alcohol, 55*(2), 230–238.
- The Psychological Corporation. (1997). *WAIS-IIIWMS-III: Technical manual*. San Antonio, TX: Harcourt, Brace & Company.
- Vaillant, G. E. (1992). *Ego mechanisms of defense: A guide for clinicians and researchers*. Washington, DC: American Psychiatric Press.
- Wechsler, H., Dowdall, G. W., Davenport, A., & Rimm, E. B. (1995). A gender-specific measure of binge drinking among college students. *American Journal of Public Health, 85*, 982–985.
- Witkin, H. (1960). “Embedded figures and personality”: A reply. *Perceptual and Motor Skills, 11*, 15–20.
- Witkin, H. (1974). Social conformity and psychological differentiation. *International Journal of Psychology, 9*(1), 11–29.

- Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. B., & Wapner, S. (1954). *Personality through perception*. New York: Harper.
- Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. (1971). *A manual for the embedded figures tests*. Palo Alto, CA: Consulting Psychologists Press, Inc.
- Wright, P. (2002). Marketplace metacognition and social intelligence. *Journal of Consumer Research*, 28, 677–682.
- Wyatt, J., & Bootzin, R. (1994). Cognitive processing and sleep: Implications for enhancing job performance. *Human Performance*, 7(2), 119–139.

THE EFFECT OF CONGRUENCE BETWEEN CULTURAL TIME ORIENTATION AND BUDGET PLANNING HORIZON ON EMPLOYEES' SATISFACTION WITH PARTICIPATIVE BUDGETING

Zhihong Wang and James E. Hunton

ABSTRACT

The purpose of the current study is to examine how employees from different cultures respond to participative budgeting when the budget planning horizon is congruent or incongruent with their cultural time orientation. We conducted a 2 × 2 quasi-experiment in which cultural time orientation (short term or long term) was measured and budget planning horizon (short term or long term) was manipulated. A total of 164 employees participated in the experiment – 87 from China and 77 from the United States, representing long-term and short-term cultural time orientations, respectively. The results indicate that satisfaction with participative budgeting was greater when cultural time orientation and budget planning horizon were congruent, relative to incongruent. Also, the differential reaction between congruence and incongruence was less extreme for the Chinese participants than the U.S. participants, which is

Advances in Accounting Behavioral Research, Volume 14, 91–116

Copyright © 2011 by Emerald Group Publishing Limited

All rights of reproduction in any form reserved

ISSN: 1475-1488/doi:10.1108/S1475-1488(2011)0000014007

consistent with Confucian thought of "The Doctrine of the Mean." The results of this study contribute to participative budgeting literature and suggest that managers who operate in different countries should be cognizant of cultural differences when employing participative budgeting processes.

INTRODUCTION

Participative budgeting has been widely studied in the management accounting literature. It refers to a management control system that enables affected employees to provide input into the budgetary process. The seminal work of Argyris (1952, 1953) recommends participative budgeting as a method to reduce dysfunctional effects caused by budget-achieving pressure arousal. Many prior studies (see Birnberg, Shields, & Young, 1990; Brownell, 1982, for review) have investigated the effects of participative budgeting on individuals' cognitions, attitudes, motivations, and behaviors; however, these studies have not produced unified findings. The purpose of the current study is to experimentally investigate the impact of a national cultural dimension, time orientation, on employees' attitudes toward participative budgeting, as understanding cultural differences can provide insight into why participative budgeting might not yield consistent effects.

National culture has been referred to as a learned trait that influences individuals' thoughts and decisions, and time orientation is one facet of national culture (Hofstede, 1980, 1991). According to Hofstede (1994), nations that score high on long-term orientation hold values that are focused on the future, whereas nations that score low on long-term orientation possess values that are focused on the present. By applying the participation congruence model (Clinton & Hunton, 2001), this study hypothesizes that participative budgeting will be more effective when employees who possess a short-term (long-term) time orientation are involved with budgets with a short-term (long-term) planning horizon. To test the hypothesis, we conducted a quasi-experiment with 87 participants from China and 77 participants from the United States, wherein time orientation (long term or short term) was measured and budget planning horizon (long term or short term) was manipulated. The results suggest that when cultural time orientation and budget planning horizon are congruent (incongruent), the participants' satisfaction with the participative budgeting process was relatively positive (negative).

Study results also reveal that participants from China, who indicated a relatively long-term time orientation, reacted less extremely to budget horizon congruence and incongruence, compared to participants from the United States, who expressed a relatively short-term time orientation. This finding is consistent with the Confucian thought of “The Doctrine of the Mean,” which is taught throughout China and emphasizes moderation in expression of feelings and actions as a way to maintain internal equilibrium and foster a harmonious society.

This study extends the participative budgeting literature by helping to explain prior conflicting findings of participative budgeting from a cultural perspective and suggests that managers who work for global organizations should be sensitive to different cultural time orientations when designing participative budgeting processes in different nations. The remainder of this chapter is organized as follows. The next section reviews relevant background literature and develops the research hypothesis. The subsequent two sections present the research method and study results. The final section discusses the findings and offers suggestions for future research in this area.

BACKGROUND AND HYPOTHESIS

Participative Budgeting

Participative budgeting is one of the most widely studied areas in management accounting research. Many prior empirical studies have investigated the effects of participative budgeting on individuals’ mental states and performance and have argued that budget participation can enhance employees’ acceptance of and commitment to budget goals (Argyris, 1952, 1953; Hofstede, 1967; also see Birnberg et al., 1990; Brownell, 1982, for review). For example, Argyris’ (1952, 1953) seminal and influential studies recommended participative budgeting as a method to reduce the dysfunctional effects caused by budget-achieving pressure arousal. Following Argyris’ suggestions, many recent studies conducted surveys and experiments to investigate the function of participative budgeting. These studies show that participative budgeting is positively associated with budget goal commitment (Chong & Chong, 2002; Hofstede, 1967; Milani, 1975), organizational commitment (Milani, 1975), job satisfaction (Milani, 1975), as well as job performance (Schuler, 1980). Wentzel (2002) conducted a comprehensive field study regarding these effects of participative budgeting at a single large hospital. Wentzel (2002) found that increased participation during budgeting

increased managers' fairness perceptions, which in turn increased budget goal commitment and enhanced job performance.

Although many of the prior studies suggest positive relationships between participative budgeting and individual outcomes, the results of some studies have not been consistent with the "more participation is better" thesis (e.g., Brownell & McInnes, 1986; Cherrington, & Cherrington, 1973; Libby, 1999; Merchant, 1981), as some studies have found that certain contextual factors intervene in the relationship between participative budgeting and employee outcomes. These contextual factors include the design of reward systems (Cherrington & Cherrington, 1973), centralization of the organizational structure (Merchant, 1981), and employees' perceptions of their influence on the final budget decision (Libby, 1999).

As indicated, prior studies have not produced a coherent or unified conclusion about the individual effects of budget participation, as contextual factors such as incentives, organizational culture, and procedural justice perceptions can moderate or nullify the influence of participative budgeting on employees' cognitions, affects, motivations, and behaviors. The current study adds to the contextual stream of participative budgeting literature by examining how cultural time orientation differences can produce both positive and negative effects of budget participation, depending on whether the budget planning horizon is congruent or incongruent with the employees' time orientation.

Cultural Time Orientation

Long-term time orientation is one dimension of Hofstede's culture framework. This dimension is based on Chinese Confucian philosophy. According to Hofstede (1994), nations that score high on long-term orientation hold values that are oriented toward the future, whereas nations that score low on long-term orientation possess values that are oriented toward the present. Using a standard questionnaire about business goals, Hofstede (2004) surveyed MBA students from 15 countries or international territories around the world to study time orientation. Hofstede's long-term orientation index scores range from a high of 118 (China) to a low of 16 (Ghana, Nigeria, and Sierra Leone, all tied). The United States ranks fairly low at 29, indicating a relatively short-term time orientation, as compared to China.

The aforementioned cultural difference suggests that employees who were raised and work in the United States are likely to be relatively short-term orientated, thus place more emphasis on short-term over long-term results.

On the contrary, employees who were raised and work in China are likely to be more long-term oriented, hence be more concerned about long-term than short-term results. This difference in time orientation across cultures raises the question as to the attitudinal effect of mismatching employees' cultural time orientation with the planning horizon (long term or short term) of the budget on which they have been asked to participate.

Budget Participation Congruence

Contingency theory suggests that there is no best way to design a management control system. Instead, the applicability of a specific management control system depends on the "congruence" or "fit" between the system and its environment. For example, [Lau and Chong \(2002\)](#) investigated the three-way interaction effect of budget participation, budget emphasize, and managers' organizational commitment on managers' behavior. Their findings suggested that managers' reaction toward budget emphasis in performance evaluation depended on the budget participation environment. Specifically, highly committed managers who are likely to strive for organizational goals reacted more favorably toward high-budget emphasis in a high-budget participation environment; however, highly committed managers also reacted more favorably toward low-budget emphasis when the organizational environment is low-budget participation.

Previous congruence research suggests a "fit" or "matching" construct as a congruence factor in the design of an effective participative decision-making process. For example, [Tushman and Nadler \(1978\)](#) found that the most effective participative strategy is when the individuals' participation requirements are congruent with the degree of participation, they are allowed, resulting in an effectively designed participation strategy. At the individual level, the congruence between the perceived need for participation and the actual participation is positively associated with individual outcomes such as satisfaction with participation ([Doll & Torzadeh, 1991](#)). When individuals experience a lower level of participation than they expect, they likely will experience a state of deprivation. Furthermore, participation congruence is a critical success factor in designing an effective participative budgeting strategy ([Clinton & Hunton, 2001](#)). These previous studies suggest that the effectiveness of a participation strategy is likely influenced by the fit between the individuals' perceived need for participation and the actual level of participation they are allowed.

Consistent with Clinton and Hunton's (2001) congruence framework, we propose a disordinal interaction between cultural time orientation (long term or short term) and budget planning horizon (long term or short term), as next articulated. When employees who have been raised and work in a culture that emphasizes a long-term time orientation are assigned to participate on a long-term (short-term) budget team, they should perceive the task more (less) important and hold more (less) positive effects toward the participative process. On the contrary, when employees who have been raised and work in a culture that reinforces a short-term time orientation are assigned to participate on a long-term (short-term) budget team, they are expected to perceive the task as less (more) important and hold less (more) positive effects toward the participative process. The above discussions lead to the following interaction hypothesis:

H₁. Employees who hold a long-term (short-term) cultural time orientation will be more satisfied with and will attribute more importance to their assignment on a long-term (short-term) budget planning team, relative to a short-term (long-term) budget planning team.

The Doctrine of the Mean

Hofstede's cultural time orientation is rooted in the teaching of Confucian philosophy. According to Hofstede, most eastern countries (e.g., China, Japan, Hong Kong, Taiwan, and South Korea), which have been deeply influenced by Confucian philosophy, possess a relatively long-term time orientation, whereas cultures that have not been as influenced by Confucian philosophy (e.g., the United States and almost all European countries) possess a relatively short-term time orientation. One particular stream of Confucian thought, The Doctrine of the Mean, is closely related to cultural time orientation. The Doctrine of the Mean suggests that individuals should consider the long-term implications of their expressions and behaviors, as overreacting to an immediate situation can create internal stress and social disharmony. The Doctrine of the Mean is a very influential stream of Confucian thought – one that emphasizes conflict avoidance and harmonious balance through moderation (Du, 2008).

The Doctrine of the Mean focuses on the power of self-constraint. The goal is to maintain balance and harmony by directing the mind to a state of constant equilibrium (Legge, 1893). The first chapter of the Doctrine of the Mean, as translated by Du (2008, p. 2), states the following: "Before the

feelings of pleasure, anger, sorrow, and joy are aroused, it is called centrality. When the feelings are aroused and each and all attain due measure and degree, it is called harmony. Centrality is the great foundation of the world, and harmony is its universal path.” In essence, The Doctrine of the Mean teaches individuals to think about long-term repercussion when processing both positive and negative stimuli and to react with judicious restraint.

In contrast, the notion of individualism that appears to be prevalent in western cultures seems to encourage relatively unrestrained affective and behavioral reactions to immediate stimuli (Bookchin, 1995; Hofstede, 1980). For instance, as described by the American social philosopher Bookchin (1995), individualists tend to express opposition in uniquely personal forms, which can become manifest in fiery tracts and outrageous behavior. The Doctrine of the Mean suggests that people who possess a long-term cultural time orientation (which is based in Confucian philosophy), relative to a short-term orientation, will indicate more moderation in their responses to congruence or incongruence with the budget planning horizon, as stated in the following hypothesis:

H₂. Employees who hold a long-term cultural time orientation will react less extremely to congruence and incongruence with the budgetary planning horizon, relative to employees who hold a short-term cultural time orientation.

RESEARCH METHOD

Participants

Initially, there were 101 participants from China; however, 14 participants did not completely respond to the dependent variable items; thus, they were dropped from the study. The remaining sample is composed of 87 participants from China and 77 participants from the United States, for a total usable sample size of 164.

The average (standard deviation) age of the Chinese participants was 30.65 (5.35); their average (standard deviation) years of work experience was 8.60 (5.79); there were 58 (67%) female and 29 (33%) male participants; 53 (61%) held a bachelor’s degree, 18 (21%) held a master’s degree, and 16 (18%) held other degrees. Among the U.S. participants, their mean (standard deviation) age was 34.43 (5.99); their average (standard deviation) years of work experience was 9.36 (5.92); there were 33 (43%) female and 44 (57%)

male participants; 29 (37%) held a bachelor's degree, 44 (57%) held a master's degree, and 5 (6%) held a PhD degree.

Between the two countries, mean ages were significantly different ($t=4.26$, $p<0.01$); mean years work experience were not significantly different ($t=0.84$, $p=0.40$); gender distribution was significantly different ($X^2=9.34$, $p<0.01$); and level of degree held distribution was significantly different ($X^2=23.86$, $p<0.01$). As the two populations differed on several demographics, all of the demographic variables will be included as possible covariates in upcoming hypotheses tests, thereby statistically controlling for potential systematic effects on the dependent variables.

Experimental Design

We employed a two (cultural time orientation: long term, short term) by two (budget planning horizon: long term, short term) between participant, quasi-experimental design in which cultural time orientation was measured and budget planning horizon was randomized. Participants were instructed to assume that they were accounting managers who worked for a large national company. The company recently launched an enterprise-wide budgeting project, in which they were asked to participate in developing budgets relevant to their area of responsibility. They were told that there will be two budget teams – one that will develop a one-year financial budget and another that will develop a four-year financial budget. The budget planning horizon treatment manipulated the budget planning horizon by indicating that the participant was assigned to a one-year financial budget (short-term) team or a four-year financial budget (long-term) team. The description of the one-year budget was that the budget would be subdivided into four quarterly financial budgets for 2010, and the description of the four-year budget was that the budget would be subdivided into four annual financial budgets for 2010 through 2013. Time orientation of the participants was reflected by Chinese participants (long-term cultural time orientation) and U.S. participants (short-term cultural time orientation).

Procedure

Consent Form and Case Materials

All participants confirmed their voluntary participation in the experiment by reading and signing a voluntary consent form. The consent form explained

that the purpose of the study was to investigate decision-making processes. The form also explained that the participants needed to complete a reading task, after which they would be asked to respond to some questions. They were told that the survey was strictly anonymous and confidential.

The case materials, shown in the [appendix](#), were written in Chinese and English. We used the back translation procedure to translate the case material into Chinese. One of the authors translated the English version into Chinese, and another PhD student translated the Chinese version back into English. Both translators were bilingual native Chinese speakers. The two English versions were then compared, and the Chinese version was revised and agreed upon by the two translators.

Budget Planning Horizon Manipulation

Participants in the short-term budget planning horizon condition read the following statement:

You have been assigned to the one-year budget team, which is responsible for preparing four quarterly financial budgets for the year 2010. The budget team consists of 20 team members who come from different departments. The whole budgeting project will last for two months.

Participants in the long-term budget planning horizon condition read the following statement:

You have been assigned to the four-year budget team, which is responsible for preparing four annual financial budgets for the years 2010 through 2013. The budget team consists of 20 team members who come from different departments. The whole budgeting project will last for two months.

After reading the case materials, participants were asked to complete a questionnaire related to the dependent variables, manipulation checks, and demographics.

Response Items

Dependent variables were measured using an 11-point scale (see the [appendix](#)). The questionnaire consists of two parts. The first part is designed to test participants' overall satisfaction with and perceived importance of being assigned to the budget team. We adapted and used items from the [Mehrabian and Russell \(1974\)](#) PAD (Pleasure, Arousal, and Dominance) scales to design the dependent variables for the current study. The PAD scales have been widely used and reported high reliability in previous psychology and marketing research. For example, [Mehrabian and Russell \(1974\)](#) reported internal consistency reliability of 0.72, 0.69, and 0.77 for

pleasure, arousal, and dominance, respectively. Holbrook Chestnut, Oliva, and Greenleaf (1984) studied the impact of emotions on the enjoyment of games using PAD scales. They reported coefficient alpha estimates of 0.89, 0.89, and 0.88 for each of the three dimensions. We selected satisfaction and happiness item from the Pleasure dimension (dependent variable response items 1.1 and 1.2 in the appendix). We used three items from the Arousal dimension to test the participants' perceived commitment, motivation, and respect (dependent variable response items 1.3, 1.4, and 1.5 in the appendix). Finally, we chose the item of importance (dependent variable response item 2 in the appendix) from the Dominance dimension.

We designed a manipulation check questionnaire for both of the independent variables in the experiment. With regard to cultural time orientation, we assumed that the Chinese (U.S.) participants would be more oriented toward the long term (short term). To test this assumption, we measured the participants' cultural time orientation using scale that was created by Bearden, Money, and Nevins (2006). We used only four items of the original eight items, as a factor analysis conducted by Bearden et al. indicated that these four items comprise a "planning" factor, which is related to the budget project in the case (manipulation check response items 3.1 through 3.4 in the appendix).

Budget planning horizon was checked in two ways. First, to confirm the randomized budget participation status, we asked the participants to choose which budget team, one year or four years, they were assigned (manipulation check response item 4). Then, we asked the participants to rate their perceptions of the length of the budget to which they were assigned on a 7-point scale (1 = very short, 4 = medium length, and 7 = very long) (manipulation check response item 5).

We also designed two additional debriefing items for use as potential covariates. First, we asked subjects how much effort they think it will take for the team to develop the financial budget (1 = very little effort, 4 = medium effort, and 7 = very much effort) (manipulation check response item 6). Second, we asked them to rate how they felt about the length of the two-month time period that would be spent on the project (1 = too short, 4 = about right, and 7 = too long) (manipulation check response item 7). Finally, the participants were asked to provide some demographic information.

Administration

For the Chinese sample, we created two survey questionnaires, one for the long-term budget planning horizon and the other for the short-term budget planning horizon, using SurveyMonkey.com. We published the

SurveyMonkey links in a Chinese accounting forum (bbs.chinaacc.com) and solicited volunteers. For those who volunteered, the short-term and long-term budget links were randomly sequenced in e-mails that were sent to the forum participants. Furthermore, we e-mailed the links to personal contacts, some of whom also e-mailed the links to their personal contacts.

The U.S. participants completed the experimental materials through pencil and paper. The participants were solicited from professional training/education classes that were being taught by one of the authors. To test for possible response order effects, we developed two versions of the experimental materials for each treatment, wherein the order of the dependent variable responses was randomized. The treatments and versions were randomly sequenced and stacked together. The instructor handed out the materials from the top to the bottom of the stack. Three training classes are represented in the U.S. sample.

RESULTS

Manipulation Checks

For the cultural time orientation construct, theory suggests that the Chinese participants will have a longer time orientation than the U.S. participants; hence, we chose these two countries as surrogates for the long-term and the short-term orientations. To confirm the appropriate use of the surrogates, we asked the participants four items related to their time orientation. The four items, shown in the [appendix](#), were used in a long-term orientation questionnaire by [Bearden et al. \(2006\)](#), which was found to have a high reliability (the average coefficient α value for planning factor of LTO in seven different studies conducted by [Bearden et al., 2006](#), was 0.76).

In the current study, the standardized Cronbach's alpha statistic was relatively high (0.79); thus, we averaged the four items into a composite index for analysis purposes. The index is oriented such that a higher value suggests a longer-term orientation. The means (standard deviations) of the time orientation index for the Chinese and the U.S. participants, respectively, were as follows: 5.16 (1.15) and 3.91 (1.39). The means are significantly different ($t = 6.31, p < 0.01$). On the basis of the results of manipulation check testing, the selection of relatively long- and short-term orientations is deemed successful.

To check the budget planning horizon (long-term or short-term) manipulation, we asked participants to indicate the budget team to which they had

been assigned (one year or four years). Among the Chinese sample, three participants in the one-year budget planning horizon did not answer this question, one participant in the short-term condition incorrectly indicated that he/she was assigned to the long-term budget planning horizon, and two participants in the long-term condition incorrectly indicated that they were assigned to the short-term budget planning horizon. Among the U.S. sample, all participants correctly responded in accordance with their randomized condition.

To further examine the budget planning horizon manipulation, participants were asked to describe the length of the budget to which they had been assigned (1 = very short, 4 = medium length, and 7 = very long). The means (standard deviations) of the short-term and long-term conditions were as follows, respectively: 3.52 (1.09) and 5.11 (1.31). The means were significantly different ($t = 8.455$, $p < 0.01$). We checked individual responses to this item for the Chinese participants who did not answer or incorrectly answered the budget horizon manipulation check question. The responses were either on the expected side of the mid-point of the scale (4), consistent with their budget horizon treatment, or at the mid-point of the scale. A robustness test (not tabulated) indicated that excluding these participants from the upcoming results strengthened the power of the statistical tests. Hence, for the sake of conservatism, all participants were retained in the sample.

Potential Covariates

Aside from demographic factors, we measured two additional potential covariates. One covariate focused on the amount of effort participants thought it would take for the team to develop the financial budgets (1 = very little effort, 4 = medium effort, and 7 = very much effort). We asked this question because participants in the four-year budget condition might consider the task to be more effortful than the one-year team. Additionally, due to cultural differences, perceived effort might differ between the two countries. Using country as one independent variable and budget planning horizon as the other independent variable, we analyzed the effort item using analysis of variance (ANOVA). The country variable was not significant ($F = 0.45$, $p = 0.50$), the budget planning horizon factor was significantly different ($F = 8.96$, $p < 0.01$), and the interaction term was nonsignificant ($F = 0.71$, $p = 0.40$). The mean (standard deviations) for the long-term budget planning horizon was 6.15 (1.04), and for the short-term budget

planning horizon, it was 5.56 (1.43). The effort variable will be used as a covariate when testing the hypotheses.

The other potential covariate asked participants about the length of time their team was allotted to complete the budget, which was held constant at two months in both conditions (1 = too short, 4 = about right, and 7 = too long). This question was asked because participants in the four-year budget condition might consider the time period to be too short, relative to the one-year team. Cultural differences might also lead to differences between the two countries. Using country as one independent variable and budget planning horizon as the other independent variable, we analyzed the effort item using ANOVA. The country variable was not significant ($F=0.21$, $p=0.65$), the budget planning horizon factor was significantly different ($F=18.04$, $p<0.01$), and the interaction term was nonsignificant ($F=0.35$, $p=0.55$). The means (standard deviations) for the long-term budget planning horizon was 3.61 (1.30), and for the short-term budget planning horizon, it was 4.52 (1.43). The time variable will be also used as a covariate when testing the hypotheses.

Preliminary Analyses

We used factor analysis to determine the number of factors reflected by the six dependent variable responses. Descriptive statistics for the six items and the factor analysis results are summarized in Table 1. As indicated, all six items loaded on a single factor, with a Cronbach's alpha of 0.91. Thus, the items were averaged to form a single "participative budgeting satisfaction" index, with a mean (standard deviation) of 2.12 (2.14).

Table 1. Descriptive Statistics and Factor Loadings.

Response Item	Mean	Standard Deviation	Factor Loading ^a
Satisfied	1.44	2.39	0.88
Happy	1.64	2.49	0.83
Committed	2.74	2.54	0.80
Motivated	2.52	2.57	0.82
Respected	1.88	2.49	0.87
Important	2.15	2.56	0.81

^aVarimax rotation, only factors with eigenvalues ≥ 1.00 are retained. Only one factor obtained, with an eigenvalue of 4.18 and a percent of variance explained of 69.75%. Standardized Cronbach's alpha = 0.91.

Analyzing only the U.S. sample, an analysis of covariance (ANCOVA) was run using participative budgeting satisfaction as the dependent variable, budget planning horizon as the independent variable, and session number and version number as covariates. Budget planning horizon was significant ($F = 92.35, p < 0.01$), session number was nonsignificant ($F = 0.62, p = 0.44$), and version number was nonsignificant ($F = 1.44, p = 0.24$). Thus, responses from the U.S. sample did not differ based on the educational session they were attending or the instrument version.

Hypotheses Testing

The first hypothesis (H_1) posits that participants with a long-term cultural time orientation will be more satisfied with their participative budgeting assignment when they are assigned to a long-term budgeting project, relative to a short-term project, and participants with a short-term cultural time orientation will be more satisfied when they are assigned to a short-term, relative to a long-term, budgeting project. We test the hypothesis using ANCOVA, the results of which are summarized in Table 2.

Table 2, panel A, presents descriptive statistics by treatment condition. Table 2, panel B, shows the results of ANCOVA testing. As indicated in panel B, the only (marginally) significant covariate is gender ($p = 0.09$). The main effect of budget planning horizon is significant ($F = 16.59, p < 0.01$), the main effect of cultural time orientation is not significant ($F = 0.19, p = 0.67$), and the interaction term is significant ($F = 71.09, p < 0.01$). The least square means are plotted on Fig. 1 for visual representation of the results. As indicated in Table 2, panel C, all means are significantly different from one another.

As suggested by the significant interaction term and depicted in Fig. 1, when there is congruence between cultural time orientation and budget planning horizon, satisfaction is significantly higher than when there is incongruence. This effect is consistent in both cultures. Thus, H_1 is supported.

The second hypothesis (H_2) posits that participants with a long-term orientation will react less extremely when the budget planning horizon is congruent or incongruent with their long-term orientation, as compared to participants with a short-term orientation. Panel D (Table 2) shows the results of a planned contrast designed to test for a difference-in-differences between the congruent and the incongruent conditions between the United States and China. As indicated, the extremity of responses is significantly greater ($p < 0.01$) for the U.S. employees (short-term orientation), relative to

Table 2. Results of Participative Budgeting Satisfaction.

Panel A: Least Square Means (Standard Deviations) and {Sample Sizes}

Cultural time orientation	Budget planning horizon		Main effect: Culture
	Short-Term	Long-Term	
Short term	4.05 (0.94) {37}	0.48 (2.09) {40}	2.19 (2.43) {77}
Long term	1.57 (1.72) {47}	2.60 (1.88) {40}	2.05 (1.85) {87}
Main effect: Budget planning horizon	2.67 (1.88) {84}	1.54 (2.24) {80}	2.12 (2.14) {164}

Panel B: ANCOVA Model Results

	Sum of squares	df	Mean square	F-ratio	p-value
Intercept	6.77	1	6.77	2.27	0.13
Age	0.17	1	0.17	0.06	0.81
Experience	0.01	1	0.01	0.01	0.97
Gender	8.55	1	8.55	2.87	0.09
Degree held	2.47	1	2.47	0.83	0.37
Perceived effort	0.86	1	0.86	0.29	0.59
Time period	0.13	1	0.13	0.04	0.84
Budget planning horizon	49.49	1	49.49	16.59	<0.01
Cultural time Orientation	0.56	1	0.56	0.19	0.67
Horizon × orientation	212.02	1	212.02	71.09	<0.01
Error	459.28	154	2.98		
Total	1,479.00	164			

Panel C: Multiple Pairwise Comparisons

Short-term time orientation and short-term budget horizon	Long-term time orientation and long-term budget horizon	Long-term time orientation and short-term budget horizon	Short-term time orientation and long-term budget horizon
Congruence 4.05	Congruence 2.60	Incongruence 1.57	Incongruence 0.48

Panel D: Contrast Test for H₂

Short-term time orientation (U.S.)		Long-term time orientation(China)	
Short-term budget	Long-term budget	Long-term budget	Short-term budget
4.05	0.48	2.60	1.57

Notes: Panel B: Adjusted $R^2 = .347$;

Panel C: $4.05 > 2.60 > 1.75 > 0.48$; Bonferroni at $\alpha = 0.05$, F -ratio = 30.30, p -value < 0.01;

Panel D: $(4.05 - 0.48) > (2.60 - 1.57)$;

t -statistic = 4.73, p -value < 0.01.

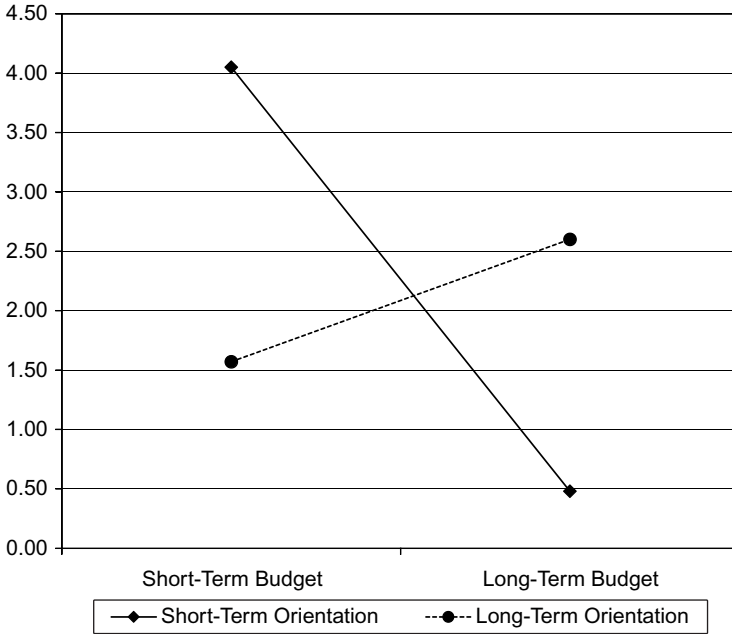


Fig. 1. Interaction Effect of Cultural Time Orientation by Budget Planning Horizon on Participative Budgeting Satisfaction (Least-Square Means are reflected on the Y-Axis).

the Chinese employees (long-term orientation), which is consistent with The Doctrine of the Mean and supportive of H₂.

Mediator–Moderator Analyses

As suggested by Baron and Kenny (1986), we conduct a mediator–moderator analysis to determine the influential nature of cultural time orientation on participative budgeting satisfaction, under conditions of congruence and incongruence with the budget planning horizon. To conduct the analyses, rather than using a dichotomous variable to represent long-term and short-term time orientation, as reflected by participants from China and the United States, we use the manipulation check results from the long-term orientation scale; also, rather than using a dichotomous variable to represent long-term

and short-term budget planning horizon, we use the manipulation check results from the scaled item that measured the participants' perceptions of the long-term or short-term nature of the budget team to which they were assigned.

First, we regress budget planning horizon on cultural time orientation and the results are nonsignificant ($\beta = -0.02$, $t = -0.023$, $p = 0.82$). Second, we regress budget planning horizon on participative budgeting satisfaction and the results are marginally significant ($\beta = -0.013$, $t = -1.68$, $p = 0.10$). Third, we regress both cultural time orientation and budget planning horizon on participative budgeting satisfaction, and the results are significant for cultural time orientation ($\beta = .248$, $t = 3.28$, $p < 0.01$) and marginally significant for budget planning horizon ($\beta = -0.013$, $t = -1.67$, $p = 0.10$). As the first regression is nonsignificant, and the effect of budget planning horizon is unchanged from the second and the third regressions, cultural time orientation does not appear to mediate the relationship between budget planning horizon and participative budgeting satisfaction. Instead, considering the significant interaction term in the ANCOVA model (Table 2, panel B), which is also significant in a regression model using the same scaled independent variables as employed in the three regressions above (interaction term: $\beta = 1.52$, $t = 7.68$, $p < 0.01$), cultural time orientation moderates the relationship between budget planning horizon and participative budgeting satisfaction. The nature of the moderation is depicted in Fig. 1.

SUMMARY AND CONCLUSION

The purpose of the current study is to examine how cultural differences can yield dissimilar responses to participative budgeting in organizations; specifically, we study how congruence between cultural time orientation and budget planning horizon affects employees' overall satisfaction with participative budgeting. This is an important topic because managers in global firms need to understand how and why cultural differences can yield unintentional negative consequences when employees are required to participate in the budgeting process. Through such understanding, managers can develop various interventions aimed at creating alignment between cultural orientations and job assignments, thereby enhancing individual and organizational outcomes.

A total of 164 participants took part in a quasi-experiment, where cultural time orientation (relatively long term and short term) was measured and

budget planning horizon (relatively long and short) was randomly manipulated. There were 87 participants from China and 77 from the United States, where the countries served as proxies for long-term and short-term cultural time orientation (Hofstede, 1991, 1994), as validated by a time orientation scale. Participants were assigned to short-term (one-year) and long-term (four-year) budget teams, which reflected the budget planning horizon treatment. The dependent variable reflected the participant's satisfaction with the participative budgeting team to which they were assigned. The findings revealed a significant interaction effect between budget planning horizon and time orientation, such that long-term (short-term) time orientation participants were more satisfied being assigned to a long-term (short-term) budget team, relative to a short-term (long-term) budget team. We also found that participants from the United States were more extreme in their positive reaction to congruence and negative reaction to incongruence than participants from China, which is consistent with Confucian philosophy, specifically The Doctrine of the Mean.

This study contributes to the budget participation literature by examining a particular contextual factor, cultural time orientation, which can negatively impact the presumed positive relationship between participative budgeting and individual outcomes. The results shed insight into the mixed findings of prior studies that were conducted in international regions with different cultural backgrounds; for instance, some of these studies reported that participative budgeting increased job satisfaction (Chong, & Chong, 2002; Milani, 1975; Wentzel, 2002), while other research reported negative correlation or mixed results (Cherrington & Cherrington, 1973; Kenis, 1979; Libby, 1999; Merchant, 1981). To our knowledge, this is the first study to explain previous inconsistent results through the lenses of cultural differences and contingency theory.

The current study is limited in several ways. Participants were not asked to actually participate in developing a budget; rather, they were asked how they felt, a priori, about being assigned to a long-term or short-term budget project team. Perhaps, their satisfaction would be different after participation. Future studies should develop behavioral tasks to test for the consistency and persistence of our results. Another limitation involves the web-based administration of the experiment in China and the paper-based administration in the United States. Whereas we do not know the extent to which these method differences might have affected the participants' responses in the current study, Bryant, Hunton, and Stone (2004) suggest that web-based and paper-based administrations of the same experiment or survey typically do not yield differential responses. Sample selection

limitations should also be noted. Some of the Chinese participants were selected from personal contacts, and the others were users of an Internet-based accounting forum. Participants in the U.S. sample were attending educational seminars held by an international training firm. Thus, we do not know how self-selection of the sample might have biased the results. Nevertheless, study results confirmed a cultural difference in time orientations between countries, which was the essence of one of the independent variables in our study. Accordingly, we suggest that from a theoretical perspective, the sample appears to be a reasonable proxy for cultural time orientation.

Furthermore, a concept known as “response style” could be confounding the results reported herein. As defined by Cronbach (1946, 1950), response style refers to the systematic tendency to choose certain portions of a rating scale. Previous psychological research has identified two predominate response styles – acquiescence and extreme. The acquiescence response style, which has also been referred as directional bias (Hui & Triandis, 1985) or positivity bias (Baumgartner & Steenkamp, 2001), refers to an individual’s preference to choose “true” or “yes” over “false” or “no” when answering a questionnaire. Some studies have indicated that the acquiescence response style is associated with impression management and self-presentation bias (Johnson et al., 1997; Landsberger & Saavedra, 1967; Lenski & Leggett, 1960). Other studies have suggested that the acquiescence response style is related more to cognitive limitations than impression management or self-presentation concerns (Knowles & Nathan, 1997; Zhou & McClendon, 1999). The fairly extensive literature on the acquiescence response style has not provided evidence of a cultural link to such acquiescence; however, cultural factors have been attributed to the extreme response style.

The extreme response style refers to the tendency to choose the farthest points of rating scales when answering a questionnaire (Berg & Collier, 1953). Previous studies have documented different levels of the extreme response style across ethnic and cultural groups. For example, Chen, Lee, and Stevenson (1995) examined cross-cultural differences in response styles with 5,162 high school students from Japan, Taiwan, Canada, and the United States. They found that the U.S. students were more likely to use the extreme values, whereas the Japanese and Taiwanese students were more likely to select the mid-point values. Chen et al. (1995) attribute the differences to Hofstede’s cultural dimension of individualism/collectivism. A collectivistic culture fosters restraint in the expression of self-feelings out of respect for others and to maintain a harmonious society; in contrast, an individualistic culture encourages individuals to maintain their independence from others

and express their unique emotional reactions. Prior literature reviews also support this view (e.g., Markus & Kitayama, 1991; Mesquita & Frijda, 1992).

There is also evidence suggesting that the extreme response style might be attributable to factors other than individualism/collectivism, as differences in extreme response styles have also been revealed across different social groups within an individualist culture. For instance, Berg and Collier (1953) found different extreme response styles between males and females as well as between black and white Americans. Hui and Triandis (1989) found that Hispanics had a stronger tendency for extreme responses than non-Hispanics. More recently, Herk, Poortinga, and Verhallen (2004) collected data from six countries (Greece, Italy, Spain, France, Germany, and the United Kingdom) in Europe. The results of their study show that people in Mediterranean Europe respond more extremely than people in North-western Europe. Hence, the extent to which response style can explain the study findings is unclear, as the extreme response style can be found within individualistic cultures, as well as across individualistic and collectivist cultures. Future research in this area is warranted to disentangle the individual and joint effects of response style and The Doctrine of the Mean on the more tempered responses from the Chinese participants, relative to the U.S. participants, indicated in the current study.

Future research on cultural differences in participative budgeting should also consider examining other cultural dimensions as well. Perhaps, employees who possess different perceptions of “power distance” between themselves and their superiors will be more or less willing to engage in participative budgeting (Lee, Pillutla, & Law, 2000). For instance, lower power distance employees might feel offended if they are not asked to participate in the budget setting process because they expect that their opinions should be impounded into their superiors’ decisions, whereas higher power distance employees might feel offended if they are asked to participate because they believe that it is their superiors’ job to make such decisions. Another potentially fruitful area of research in participative budgeting involves the cultural spectrum of “individualism/collectivism” (Hofstede, 1980) with regard to small groups. For example, in a team budgeting environment, individualists might withhold critical pieces their personal knowledge as a way to protect their private information advantage, whereas collectivists might be willing to share everything they know about the budget area as a way to enhance of overall organization. By further investigating the impact of cultural differences on participative budgeting outcomes, future research can help to improve the efficacy participative budgeting practices in multinational companies.

On the basis of the results of this study, managers should be cognizant of national differences when assigning employees from different cultures to participative budgeting teams. Naturally, managers cannot assign long-term orientation employees only to long-term budgeting projects and vice versa. We suggest, though, that if managers are about to assign, say, a long-term orientation employee to a short-term budgeting project, the manager should carefully explain to the employee the importance of participating on such a project and perhaps explain why the employee was chosen (e.g., based on the employees experience, skills, and education). Through such cultural sensitivity, there is a greater likelihood that the employee will feel valued, be motivated, and be pleased with the assignment.

REFERENCES

- Argyris, C. (1952). *The impact of people on budgets*. New York, NY: Controllershship Foundation.
- Argyris, C. (1953). Human problems with budgets. *Harvard Business Review*, 31(1), 97–110.
- Baron, R. M., & Kenny, D. A. (1986). The mediator-moderator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182.
- Baumgartner, H., & Steenkamp, J. B. E. M. (2001). Response styles in marketing research: Across-national investigation. *Journal of Marketing Research*, 38(2), 143–156.
- Bearden, W. O., Money, R. B., & Nevins, J. L. (2006). A measure of long-term orientation: Development and validation. *Journal of the Academy of Marketing Science*, 34(3), 456–467.
- Berg, I. A., & Collier, J. S. (1953). Personality and group differences in extreme response sets. *Educational and Psychological Measurement*, 13(2), 164–169.
- Birnberg, J., Shields, M., & Young, S. M. (1990). The case for multiple methods in empirical management accounting research (with an illustration from budget setting). *Journal of Management Accounting Research*, 2(Fall), 33–66.
- Bookchin, M. (1995). *Social anarchism or lifestyle anarchism: An unbridgeable chasm*. Stirling, Scotland: AK Press.
- Brownell, P. (1982). Participation in the budgeting process: When it works and when it doesn't. *Journal of Accounting Literature*, 20(Spring), 124–150.
- Brownell, P., & McInnes, M. (1986). Budgetary participation, motivation, and managerial performance. *The Accounting Review*, 61(4), 587–600.
- Bryant, S., Hunton, J. E., & Stone, D. (2004). Internet-based experiments: Prospects and possibilities for accounting behavioral researchers. *Behavioral Research in Accounting*, 16(1), 107–129.
- Chen, C., Lee, S. Y., & Stevenson, H. W. (1995). Response style and cross-cultural comparison of rating scales among East Asian and North American students. *Psychological Science*, 6(3), 170–175.
- Cherrington, D., & Cherrington, J. (1973). Appropriate reinforcement contingencies in the budgeting process. *Journal of Accounting Research Supplement*, 11, 225–253.

- Chong, V., & Chong, K. (2002). Budget goal commitment and informational effects of budget participation on performance: A structural equation modeling approach. *Behavioral Research in Accounting*, 14(1), 65–86.
- Clinton, B. D., & Hunton, J. E. (2001). Linking participative budgeting congruence to organization performance. *Behavioral Research in Accounting*, 13(1), 127–141.
- Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement*, 6, 475–494.
- Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educational and Psychological Measurement*, 10, 3–31.
- Doll, W. J., & Torkzadeh, G. (1991). A congruence construct of user involvement. *MIS Quarterly*, 22(2), 443–453.
- Du, W. M. (2008). *An insight of Chung-Yung*. Beijing, China: Renmin Press.
- Herk, H. V., Poortinga, Y. H., & Verhallen, T. M. M. (2004). Response styles in rating scales: Evidence of method bias in data from six EU countries. *Journal of Cross-Cultural Psychology*, 35(3), 346–360.
- Hofstede, G. (1967). *The game of budget control*. Assen, The Netherlands: Van Gorcum.
- Hofstede, G. (1980). *Culture's consequences*. Beverly Hills, CA: Sage.
- Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. London: McGraw-Hill.
- Hofstede, G. (1994). Management scientists are human. *Management Science*, 40(1), 4–13.
- Hofstede, G. (2004). Business goals and corporate governance. *Asia Pacific Business Review*, 10(3/4), 292–301.
- Holbrook, M. B., Chestnut, R. W., Oliva, T. A., & Greenleaf, E. A. (1984). Play as consumption experience: The roles of emotions, performance, and personality in the enjoyment of games. *Journal of Consumer Research*, 11(2), 728–739.
- Hui, C. H., & Triandis, H. C. (1985). The instability of response sets. *Public Opinion Quarterly*, 49(2), 253–260.
- Hui, C. H., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, 20(3), 296–309.
- Johnson, T. P., O'Rourke, D., Chavez, N., Sudman, S., Warnecke, R., & Lacey, L. (1997). Social cognition and responses to survey questions among culturally diverse populations. In: L. E. Lyberg, P. P. Biemer, M. Collins, E. De Leeuw, C. Dippo, N. Schwarz & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 87–113). New York, NY: Wiley.
- Kenis, I. (1979). Effects of budgetary goal characteristics on managerial attitudes and performance. *The Accounting Review*, 54(4), 707–721.
- Knowles, E. S., & Nathan, K. T. (1997). Acquiescent responding in self-reports: Cognitive style or social concern?. *Journal of Research in Personality*, 31(2), 293–301.
- Landsberger, H. A., & Saavedra, A. (1967). Response set in developing countries. *Public Opinion Quarterly*, 31, 214–229.
- Lau, C. M., & Chong, J. (2002). The effects of budget emphasis, participation and organizational commitment on job satisfaction: Evidence from the financial services sector. *Advances in Accounting Behavioral Research*, 5, 183–211.
- Lee, C., Pillutla, M., & Law, K. S. (2000). Power-distance, gender, and organizational justice. *Journal of Management*, 26(4), 685–704.
- Legge, J. (1893). *The great learning and the doctrine of the mean* (Translated, with Critical and Exegetical Notes, Prolegomena, Copious Indexes, and Dictionary of All Characters, 2nd revised). Oxford, UK: Clarendon Press.

- Lenski, G. E., & Leggett, J. C. (1960). Caste, class, and deference in the research interview. *The American Journal of Sociology*, 65, 463–467.
- Libby, T. (1999). The influence of voice and explanation on performance in a participative budgeting setting. *Accounting Organizations and Society*, 24(2), 125–137.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253.
- Mehrabian, A., & Russell, J. A. (1974). A verbal measure of information rate for studies in environmental psychology. *Environment and Behavior*, 6(4), 233–252.
- Merchant, K. (1981). The design of the corporate budgeting system: Influences on managerial behavior and performance. *The Accounting Review*, 56(4), 813–829.
- Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin*, 112(2), 179–204.
- Milani, K. (1975). The relation of participation in budget-setting to industrial supervisor performance and attitudes: A field study. *The Accounting Review*, 50(2), 274–284.
- Schuler, R. S. (1980). Definition and conceptualization of stress in organizations. *Organizational Behavior and Human Performance*, 25(2), 184–215.
- Tushman, M. L., & Nadler, D. A. (1978). Information processing as an integrating concept in organization design. *Academy of Management Review*, 3, 613–624.
- Wentzel, K. (2002). The influence of fairness perceptions and goal commitment on managers' performance in a budgetary setting. *Behavioral Research in Accounting*, 14(1), 247–271.
- Zhou, B., & McClendon, M. J. (1999). Cognitive ability and acquiescence. In: American Statistical Association (Ed.), *Proceedings of the joint statistical meeting, 54th annual conference of the American Association for Public Opinion Research*, Alexandria, VA, AMSTAT (pp. 1003–1012).

APPENDIX

Case Materials

Assume you work for a large nation-wide company that operates in most regions in your nation. The company's business is the manufacturing and selling of home appliances. The company is well operated and makes a very good profit.

Your position in the company is an accounting manager, and your main responsibilities are described below:

1. Prepare month-end balancing and journal entries
2. Prepare monthly balance sheet reconciliations
3. Participate in process design and process audits

The company has just launched an enterprise-wide budget management project. This project encourages employees to participate in the budget making process relevant to their main responsibilities. There are two budget

teams – one that will develop a one-year financial budget and another that will develop a four-year financial budget.

You have been assigned to the one-year (four-year) budget team, which is responsible for preparing four quarterly (annual) financial budgets for the year 2010 (years 2010 through 2013). The budget team consists of 20 team members who come from different departments. The whole budgeting project will last for two months.

Dependent Variable Response Items

(The numbering 1.1 through 1.5 did not appear in the experimental materials. They are included herein for clarity purposes)

1. How do you feel about being assigned to the one-year (four-year) budget team, rather than four-year (one-year) budget team? (circle one number on each of the scales below):

Item 1.1

5	4	3	2	1	0	1	2	3	4	5
Very Unsatisfied		Somewhat Unsatisfied		Neutral		Somewhat Satisfied				Very Satisfied

Item 1.2

5	4	3	2	1	0	1	2	3	4	5
Very Unhappy		Somewhat Unhappy		Neutral		Somewhat <i>Happy</i>				Very <i>Happy</i>

Item 1.3

5	4	3	2	1	0	1	2	3	4	5
Very Uncommitted		Somewhat Uncommitted		Neutral		Somewhat <i>Committed</i>				Very <i>Committed</i>

Item 1.4

5	4	3	2	1	0	1	2	3	4	5
Very Unmotivated		Somewhat Unmotivated		Neutral		Somewhat Motivated				Very Motivated

Item 1.5

5	4	3	2	1	0	1	2	3	4	5
Very		Somewhat		Neutral		Somewhat		Very		Respected
Disrespected		Disrespected				Respected				

2. The team to which I have been assigned is

5	4	3	2	1	0	1	2	3	4	5
Very		Somewhat		Neutral		Somewhat		Very		Important
Unimportant		Unimportant				Important				

Manipulation Check Response Items: Cultural Time Orientation

(The numbering 3.1 through 3.4 did not appear in the experimental materials. They are included herein for clarity purposes)

3. How would you describe yourself?

Item 3.1 I usually plan my life for the

1	2	3	4	5	6	7
Short			Intermediate			Long
Term			Term			Term

Item 3.2 I am working very hard now to gain success in the future

1	2	3	4	5	6	7
Strongly						Strongly
Disagree						Agree

Item 3.3 I don't mind giving up today's fun for success in the future

1	2	3	4	5	6	7
Strongly						Strongly
Disagree						Agree

Item 3.4 Persistence is very important to me

1	2	3	4	5	6	7
Strongly						Strongly
Disagree						Agree

Manipulation Check Response Items: Budget Planning Horizon

4. For the case scenario you just read, what is the budget period of the budget team to which you were assigned? (Check one)

One-year () Four-years ()

5. For the case scenario you just read, how would you describe the length of the budget period?

1	2	3	4	5	6	7
Very Short			Medium Length			Very Long

6. For the case scenario you just read, how much effort do you think it will take for the team to develop the financial budgets?

1	2	3	4	5	6	7
Very Little Effort			Medium Effort			Very Much Effort

7. For the case scenario you just read, I think that the two-month period during which my team will develop the financial budgets is

1	2	3	4	5	6	7
Too Short			About Right			Too Long

INTEGRATED INFORMATION SYSTEMS AND INTERORGANIZATIONAL PERFORMANCE: THE ROLE OF MANAGEMENT ACCOUNTING SYSTEMS DESIGN

Andreas I. Nicolaou

ABSTRACT

The interorganizational environment faced by business organizations presents unique challenges for management accounting and control. Past management accounting research has shown interest in such collaborations because despite their benefits, such relationships pose significant issues of coordination and control. As information and communication systems supplement management control systems in their support of decision facilitation and decision influencing, examining the design of management accounting systems (MASs) in the management of interorganizational relationships and assessing how it affects the attainment of interorganizational exchange partner performance objectives is important. In this chapter, I extend past accounting research to examine the complementary nature of decision-facilitation and decision-influencing objectives of MAS design as enabled by the use of integrated

Advances in Accounting Behavioral Research, Volume 14, 117–141

Copyright © 2011 by Emerald Group Publishing Limited

All rights of reproduction in any form reserved

ISSN: 1475-1488/doi:10.1108/S1475-1488(2011)0000014008

information systems in interorganizational settings. The economic theory of complementarity is employed to examine synergistic effects of complementary MAS objectives. A field survey is used to examine hypothesized relationships, and data were obtained from 116 organizations involved in strategic alliance activity. This chapter reports findings that support the view that the degree of complementarity in decision-facilitation and decision-influencing objectives assists in the development of capabilities that enhance performance in the interorganizational relationship. The study blends theory in the areas of strategy, information systems, and management accounting and extends management accounting research in the context of IT-enabled interorganizational relationships.

INTRODUCTION

In the recent past, business organizations have been heavily engaged in interorganizational exchanges, including both business alliances and joint ventures (Chalos & O'Connor, 2004; Das & Teng, 2000; Ireland, Hitt, & Vaidyanath, 2002). Despite this growth, an increased number of alliances fail or break up prematurely. Past research has posited various reasons for alliance failure, such as lack of partner cooperation (Harrigan, 1988; Zaheer & Venkatraman, 1995) and misfits in the adopted governance structure of these alliances (Gulati, 1995; Parkhe, 1993; Young-Ybarra & Wiersema, 1999; Zaheer & Venkatraman, 1995).

This chapter examines the effectiveness of the design of management accounting systems (MASs) that are used to govern interorganizational exchanges. Relationships at the interorganizational level present unique challenges for management accounting and control because of the potentially conflicting objectives of information that is being shared in such exchanges. For example, past management accounting research has examined aspects of control system design that are based on assumptions of opportunistic behavior between partners within the context of bounded rationality (Anthony & Govindarajan, 2004; Otley & Berry, 1980; Simons, 1995, 2000). This line of research emphasized agency costs associated with the separation of ownership and control (Baiman, 1982). Past literature has examined the decision-influencing uses of management accounting information to reduce tensions associated with increased monitoring of agent actions and willingness to share private information (for a review, see Sprinkle, 2003;

Sprinkle & Williasmon, 2007). In addition, while the decision-facilitating objective of management accounting information can be useful in revising beliefs about important dimensions of trading partner behavior (Baiman, 1982), research has only recently examined the joint impact of decision-influencing and decision-facilitation information use (Abernethy & Vangoni, 2004; Drake, Haka, & Ravenscroft, 1999; Grafton, Lillis, & Widener, 2010; Indjejikian & Matejka, 2006), while their interactions have not been examined in the interorganizational context.

In the modern business environment, interorganizational exchanges are supported by the implementation and use of integrated information systems. Integrated information systems are characterized by common data standards and business processes across partners and facilitate information flows and activity coordination (Barua, Konana, Whinston, & Yin, 2004; Johnston & Vitale (1988). Integrated information systems may also enable the development of organizational capabilities to support strategic goals in an interorganizational exchange (Sambamurthy, Bharadwaj, & Grover, 2003). As organizational capabilities are created by the effective design and development of management processes (Garvin, 1998; Grant, 1996), the design of a MAS could help an organization generate such capabilities. A MAS is conceptualized as the information and communication system (Davila & Foster, 2005) that supports the managerial processes of planning and control (Garvin, 1998). A MAS forms a subset of an organization's structural elements that make up its overall management control system (Simons, 1995, p. 5) and can serve both decision-influencing as well as decision-facilitating objectives (Baiman & Demski, 1980) in interorganizational information exchanges.

On the basis of economic complementarity theory, I predict that the decision-influencing and decision-facilitating objectives of MAS design represent complementary design choices made by management that are jointly affected by the availability of integrated information systems in interorganizational relationships. The economic theory of complementarity emphasizes the potential importance of interactions between different elements of organizational design (Athey & Stern, 1998) and provides a basis for understanding how various elements of organizational strategy and management process relate to one another (Milgrom & Roberts, 1990, 1995). On the basis of the complementarity logic, whereby firms are discovering greater avenues for competitive actions through their information value chains and through the functionalities of integrated information systems, there should be synergistic effects in the attainment of complementary interorganizational MAS objectives. These synergies represent

capability-building processes that enable the effective design of managerial processes and help attain desired goals (Grant, 1996; Teece, Pisano, & Shuen, 1997). The complementarity theory's tenets are thus important in explaining why firms develop synergies that enable the attainment of MAS objectives in an interorganizational environment and ensure the success of the interorganizational relationship.

This study therefore examines the synergistic effects of complementary decision-influencing and decision-facilitating objectives on the design of MAS and on the attainment of performance in interorganizational relationships. The study utilizes a field survey to examine these research questions. The use of a field survey allows for a richer, interactive history of responses that might be useful when examining relationships in an interorganizational exchange context (Coletti, Sedatole, & Towry, 2005; Kramer, 1999). The study offers the following contributions. First, it examines the interrelations between decision-facilitating and decision-influencing objectives of MAS design that constitute primary objectives of a MAS in an interorganizational context. Second, the study examines the effect of integrated information systems, which is a synergistic factor that affects the interrelations of MAS objectives and enables firms to enhance performance by making complementary choices in the design of the MAS. Third, the study introduces the strategic capabilities perspective in examining the effectiveness of MAS design and its performance effects in an interorganizational context. This study is one response to the "need for research on the extended enterprise that is linked to traditional management accounting research but which challenges these traditional boundaries using literatures that have begun to explore the contours of the new organizational landscape" (Anderson & Sedatole, 2003, pp. 38–39).

THEORETICAL BACKGROUND AND HYPOTHESIS DEVELOPMENT

In interorganizational alliances, integrated information systems are used to exchange information needed for the management of these relationships. The literature on integrated information systems in interorganizational exchanges has devoted considerable attention to the outcomes or benefits of such system use. Wang and Seidmann (1995) and Riggins, Kriebel, and Mukhopadhyay (1994) show that exchange benefits include efficient information sharing, improved coordination, minimized risk, and reduced

transaction costs. Similarly, Garicano and Kaplan (2001) suggest business-to-business (B2B) relationship success depends on the ability of technology to reduce transaction costs, including both coordination costs and motivation costs. B2B exchanges reduce coordination costs by providing high information quality that enables partners to transact efficiently. Lower coordination costs make markets more attractive than hierarchies (Malone, 1987; Malone, Yates, & Benjamin, 1987) and enable changes in the location of decision-making (Gurbaxani & Whang, 1991). Garicano and Kaplan (2001) find that B2B exchanges also reduce motivation costs, such as when imperfect commitment (Milgrom & Roberts, 1992) leads suppliers not to fulfill orders as promised, while other benefits include reduced errors, reduced inventory costs, and higher quality (Bakos & Brynjolfsson, 1993; Malone et al., 1987). Furthermore, information systems integration in interorganizational relationships enables firms to develop strategic capabilities that contribute to business performance (Saraf, Langdon, & Gosain, 2007).

In early research, the concept of strategy was used to denote actions or patterns of actions intended for the achievement of goals (Swamidass & Newell, 1987). Realized strategies, as defined by Mintzberg (1978), emerge through events and environmental interactions as they unfold over time (Dent, 1990). Distinctive competencies in specific strategies may be found within functional areas (Dent, 1990). Integrated information systems cannot by themselves create sustained performance gains for a firm (Chapman & Kihn, 2009). As a result, there is a clear distinction in the literature between the availability of integrated information systems (or IT resources) and the creation of IT-enabled managerial processes that lead to the development of performance-inducing organizational capabilities (Barua et al., 2004; Sambamurthy et al., 2003). Different organizations, therefore, may develop specific strategies that will allow the formation of distinctive competencies over time and lead to competitive advantage and superior performance.

The effect of specific strategies on the design of management accounting and control systems has also been proposed in past accounting research as a critical issue that has not received adequate research attention (Abernethy & Lillis, 1995; Langfield-Smith, 1997), notwithstanding the fact that interorganizational relationships may modify the types of organizational capabilities that are necessary for success. In interorganizational relationships, firms may develop organizational capabilities through the blending of information technology in organizational processes (Barua & Mukhopadhyay, 2000). The strategic management literature extends the resource-based view of the firm (Barney, 1991) to define higher-order organizational capabilities as the source of a firm's performance (Grant, 1996; Teece et al., 1997). According to this

perspective, a firm must develop dynamic capabilities to acquire, integrate, and use resources that are embedded in their social, structural, and cultural context (Eishnhardt & Martin, 2000). Recent IS research has developed theoretical models of such higher-order capabilities resulting from digital options (Sambamurthy et al., 2003) and empirically examined the IT-enabled processes that embed the availability of integrated information systems into organizational processes (Bharadwaj, 2000) and supply chain processes (Rai, Patnayakuni, & Seth, 2006).

The processes of interorganizational coordination and control are important managerial processes (Garvin, 1998) that dictate an organization's success in the alliance relationship. The development of performance-inducing organizational capabilities therefore is dependent on the IT-enabled processes of interorganizational MASs. This implies that research should examine the availability of highly integrated information systems as they enable these strategic processes and enhance the complementarity of specific interorganizational direction-setting and monitoring activities. In sum, the strategic literature in management accounting, information systems, and organizations suggests that integrated information systems are important to interorganizational performance, and they enable the design of MASs for the management of interorganizational relationships.

Integrated Information Systems, Complementarity in Interorganizational MAS Objectives, and Interorganizational Performance

In interorganizational alliances, the extent of use of integrated information systems helps construct the types of information exchange that occur in the relationship. The information exchange relationships related to a MAS could thus be formalized through the use of integrated information systems, to serve the two broad objectives of facilitating decision-making and decision-influencing actions to mitigate control problems (Indjejikian & Matejka, 2006).

The decision-influencing and decision-facilitating roles of management accounting information are not necessarily conflicting or disjoint (Sprinkle, 2003). Information that is provided by integrated information systems can be useful for both decision-influencing and decision-facilitating purposes. Consider, for example, a manager who makes a production capacity decision and is uncertain about sourcing availability of raw materials required for production. In this scenario, information about a critical alliance partner's sourcing capability has decision-facilitation or planning use as well as

decision-influencing or monitoring use. First, information about the alliance partner's sourcing capability may be made available by the integrated information systems, which enables the two partners to collaborate electronically and allows the manager to more accurately plan production capacity and reduce ex ante (pre-decision) uncertainty. Managers also need information to update their beliefs about the consequences of their own decisions and those made by trading partners; thus, interactive control systems facilitate learning (Simons, 1995). The manager could therefore use integrated information systems to obtain information about the alliance partner's future sourcing capability to revise plans about future production capacity and manufacturing budget. At the same time, ex post monitoring information about the alliance partner's past sourcing performance is also useful for future planning purposes. The likelihood of the manager using future decision-facilitating information supplied by the integrated information systems is thus affected by the manner in which the information is used for decision-influencing purposes. As a result, the two purposes are interdependent and both are affected by the extent of integration enabled by interorganizational information technology.

Recent findings in management accounting research support the complementary nature of decision-facilitating and decision-influencing objectives of MAS (e.g., Abernethy & Vangoni, 2004; Indjejikian & Matejka, 2006). In Indjejikian and Matejka, a MAS that emphasizes local decision support was found to exacerbate control problems at the corporate inter-unit level, whereas the use of management accounting practices that emphasize corporate control was found to undermine the effectiveness of local decision-making to the detriment of the firm as a whole. As a result, both decision-facilitating and decision-influencing objectives would need to be supported by MAS design. In a similar fashion, Abernethy and Vangoni (2004) report findings that the two roles of decision management and decision control are in fact complementary and not conflicting in the context of their study. In earlier studies, Drake et al. (1999) report experimental evidence where the benefit of providing detailed activity-based costing information is linked to the firm's incentive compensation system, thus supporting the complementary nature of the decision-facilitating and the decision-influencing objectives of MAS design, in that the use of information for control/monitoring enhances the use of information for planning/decision-making purposes. Tuttle and Burton (1999) also find that the presence of a modest financial incentive increased information cue usage, thus mitigating information overload and increasing task performance, suggesting a close interrelationship between the decision-facilitating and the decision-influencing objectives of MAS information.

As the preceding evidence suggests, recognizing the interrelationships between the decision-facilitating and the decision-influencing objectives of a MAS is important, especially as such interrelationships could have an impact on the realization of performance outcomes in an interorganizational alliance context.

In this study, I employ the theoretical tenets of the economic theory of complementarity (Milgrom & Roberts, 1990, 1992, 1995), which asserts that two factors are complementary when the changes in the level of one factor affects the marginal returns due to the other factor. The design of MAS to meet decision-facilitation and decision-influencing objectives is the result of adoption of economic rational firms of a coherent business strategy that exploits complementarity (Milgrom & Roberts, 1990) in the face of organizational capabilities that are enabled by the use of integrated information systems.

The adoption and use of an integrated information system in an interorganizational alliance is not a marginal decision but rather involves substantial and closely coordinated changes in organizational processes and a whole set of activities (Nicolaou, 2004a, 2004b; Nicolaou & Bhattacharya, 2006; Ross & Vitale, 2000; Scott & Vessey, 2000; Soh, Kien, & Tay-Yap, 2000; Stephanou, 2000). Prior studies examining the successful deployment of IT resources have emphasized the concomitant organizational changes associated with IT adoption (Brynjolfsson & Hitt, 2000; Brynjolfsson, Hitt, & Yang, 2002), which seem to be necessary for integrated information systems to have an effect on a firm's operational performance (Nicolaou, 2004b; Nicolaou & Bhattacharya, 2006). The use of integrated information systems could thus enable synergistic organizational processes, which could have a positive influence on firm performance (Kumar & Van Hilleegersberg, 2000). The presence of IT-enabled processes will result in organizational capabilities and enhanced performance when organizations make a series of linked strategic decisions so as to blend IT resources due to the use of integrated information systems with organizational processes and knowledge resources (Barua et al., 2004). If an organization, however, deploys IT resources that are not consistent with complementarity requirements in MAS design, it is not likely to realize returns to scale (Milgrom & Roberts, 1995) and likely to suffer a reduction in corresponding interorganizational performance. This is a similar effect as that observed in organizations that have not aligned their information technology and strategy (Davenport, 2000). As a result, the following research hypothesis is advanced:

H1. A high degree of complementarity between the decision-facilitating and the decision-influencing objectives of interorganizational MAS design will have a positive influence on a firm's interorganizational performance.

RESEARCH METHOD

To examine the study's research question, a combined archival and field survey methodology was employed on a target sample of US public companies. The mail survey provided data from the chief financial officers (CFOs) of organizations involved in interorganizational alliance activity. The archival method primarily assisted in the identification of organizations involved in strategic alliances to enhance the internal validity of the selected sample.

Sample and Respondent Selection

The sample for the study was extracted through a search of public companies that report alliance or joint venture activity in the *Mergent* database. This work has resulted in the identification of 1,896 separate alliances that were created by 893 different US public companies with third partners between the years 1982 and 2005. Of those alliances, 38% involved an international partner, while 70% were initiated in the 1990s, 8% occurred before that time, and the other 22% occurred between the years 2000 and 2005. The industry membership of alliance adopter companies included 47% in manufacturing (standard industrial classification – SIC code 2 or 3), 13% in hotels and other lodging places (SIC code 7), 10% in depository institutions (SIC code 6), and 9% in transportation (SIC code 4), among other industries with smaller participation. The CFO for each of those companies was selected as the appropriate target respondent for the study, as a CFO should have an understanding of the potential effects of the use of IT on the effective operation and control of an alliance.

Data Collection

The research instrument asks the potential respondent to choose one alliance their firm has had or currently has with another business entity. Such an alliance could be the result of a strategic agreement between two firms, and it might have involved the creation of a third entity (as in a joint venture) or not. The respondents have been instructed to focus on the relationship with this interorganizational alliance partner when responding to the various items included in the research instrument. Following [Dillman's \(1978\)](#) recommendation for conducting effective surveys, several steps were taken

during the entire data collection process. First, a preliminary draft of the research instrument was evaluated by expert panels, including faculty members and two individuals (CFOs) from the target population. The instrument was revised as a result of pretesting, ensuring the face validity of the constructs and items. Second, a preliminary letter explaining the study objectives was sent to each selected organization before mailing the first wave of surveys. Third, the first wave was mailed with a business-reply envelope and a letter requesting participation. The instrument was also coded on the author's web space, and potential respondents were given the choice of completing the paper or web-based version of it. Fourth, a postcard reminder was sent about a week after the initial mailing. Fifth, a second reminder packet (including a copy of the original questionnaire and web access instructions) was mailed to nonrespondents within eight weeks of the original mailing. Finally, an e-mail request was sent to nonrespondents with a direct link to the questionnaire web address. The response rate from all attempts is 17.13%, as a total number of 116 responses were received over an effective sample of 677 target respondent firms. Table 1 analyzes the response rate attained in the study.

Tests for nonresponse bias were performed to determine (a) whether the distribution of the effective sample of 677 organizations in the response or nonresponse categories was independent of available demographic characteristics (industrial classification, gross revenue, and number of employees) and (b) whether early and late respondents provided significantly different responses. Chi-square tests indicated no significant differences in the three demographic characteristics. The Hotelling's T^2 statistic also indicated no significant differences in the multivariate means of early versus late respondents.

Table 1. Effective Response Rate.

Original sample from Mergent database		893
Less		
Undelivered questionnaires with no forwarding information	67	
Declined response due to time pressures	49	
Declined response due to nonparticipation in surveys	65	
Declined response for unspecified reasons	35	(216)
Effective sample size		677
Number of completed questionnaires received		116
Effective response rate		17.13%

Measurement of Model Constructs and Control Variables

Distinguishing between the characteristics of integrated information systems and the objectives of MAS that are enabled by the availability of integrated systems is important for construct measurement. Firms that are involved in interorganizational alliances may be able to implement decision-influencing and decision-facilitating objectives through the adoption of integrated systems that support collaborative activities. The proper management of appropriate factors that contribute to the attainment of decision-influencing and decision-facilitating objectives relates to the complementary aspects of MAS design, whereas the mere adoption of integrated information systems and the availability of related technological capabilities relate to integrated information systems characteristics. The items used to measure all constructs are given in Table 2.

Table 2 summarizes the five items used to measure integrated information systems characteristics. The items are intended to measure the availability of characteristics such as web-based extranets for data sharing, web-based access over a partner's database, use of IT as a platform to build an organization's information infrastructure, use of web-based add-on modules, and collaborative capabilities.

As defined in the past literature (Zimmerman, 2006), decision-facilitating objectives are based on the set of those activities that take place in an interorganizational collaborative environment to initiate and implement business plans; decision-influencing objectives relate to those activities that take place in an interorganizational environment to ratify the adoption of business plans and monitor implementation. As a result, the constructs of decision-influencing and decision-facilitating objectives are each measured using new items (given in Table 2), which capture the extent to which integrated IT facilitates or enables the attainment of such objectives in an interorganizational environment. As no prior validated items exist for the measurement of these constructs, the items given in Table 2 have been originally developed in this study.

Two control measures were also used to eliminate influences on each of the *decision-influencing* and *decision-facilitating* objectives due to varied motivations to initiate an alliance. Interorganizational alliances may be formed in response to the need for either asymmetry or reciprocity of organizational objectives with those of the alliance partner (Oliver, 1990), and these may influence the types of objectives sought in a firm's MAS. The asymmetry contingency emerges from a desire for control over the other partner due to resource dependence or resource scarcity constraints

Table 2. Measurement Items of Model Constructs.**Integrated Information Technology/Systems: IIS (7-point scale, strongly agree to strongly disagree)**

1. The use of IT enables use of web-based extranets or other data sharing methods with my exchange partner.
2. My alliance partner allows me to have electronic web-based access over relevant portions of their internal database.
3. IT systems served as an essential platform to help build my firm's information infrastructure, including web enablement capabilities.
4. My firm's use of IT systems enables use of web-based add-on modules, including supply chain and customer relationship management.
5. My firm utilizes web-based collaborative capabilities enabled by its IT systems.

Complementary Objectives of MAS Design (7-point scale, strongly agree to strongly disagree)*Decision-influencing objective: DI*

Information provided by my firm's IT systems enables

1. Adequate control over outcomes or results of actions taken by my exchange partner in the alliance.
2. The assessment of alliance (exchange partner) performance over a number of operating metrics, including delivery on schedule, sharing of production plans, and minimization of production delays.
3. Use of web-based monitoring routines that provide information about partner performance over a number of specific metrics.
4. My firm to better monitor exchange partner performance.

Decision-facilitating objective: DF

1. The use of IT systems in general has allowed my firm to better coordinate decisions with the exchange partner in this alliance.
2. My firm's IT systems provide adequate information for me to make decisions that affect the relationship with my exchange partner in this alliance.
3. My firm's IT systems provide adequate information for me to plan in advance the potential outcomes of decisions that impact my relationship with this exchange partner.
4. The use of IT systems collaborative capabilities has allowed my firm to better coordinate decisions with the exchange partner in this alliance.
5. The use of IT systems increases transparency of my alliance partner's cost structure.

Interorganizational Performance: IOPRF (7-point; strongly disagree to strongly agree)

Please rate the performance of the strategic alliance

Financial dimension

- a. has been very profitable
- b. has generated a high volume of sales
- c. has achieved a high earnings growth

Strategic dimension

- d. has improved my firm's strategic competitiveness
- e. has strengthened my firm's strategic position
- f. has significantly increased my firm's market share

Table 2. (Continued)*Overall*

- g. has been very satisfactory
- h. has fully met my firm's expectations
- i. all in all, we expect that the strategic alliance with this exchange partner will continue in the long run

Exchange Partner Performance

- 1. Please rate your exchange partner's performance in following the terms of your agreement (measured as 1 = very poor; 4 = fair; 7 = excellent).

Prior Performance: PRIORPRF (Objective Measures)

- a. Please estimate your firm's revenue growth: (a) since the inception of the alliance; (b) during the alliance, if already dissolved (**PRF**).
- b. Please estimate total industry growth: (a) since the inception of the alliance; (b) during the alliance, if already dissolved (**PRI**).

Alliance Adoption Contingencies (strongly disagree to strongly agree, 1–7 rating scale)*Asymmetry Contingency (ASYMM)*

The primary motivation for my firm to enter into this strategic alliance is to

- a. Gain power over the exchange partner through control of resources and information supply.
- b. Exert dominant influence over the exchange partner.

Reciprocity Contingency (RECIP)

- c. Exchange equally important information and share plans useful to both my firm and the exchange partner.
- d. Coordinate business plans for mutual benefit.

(Pfeffer & Salancik, 1978). Such constraints encompass a need for power and control over external resources that may be critical to an organization's operational processing capabilities and could thus represent an exogenous influence on the *decision-influencing* objectives of MAS design. The reciprocity contingency assumes that two organizations that enter into an alliance do so because they anticipate reciprocal benefits, which far exceed costs related to loss of decision-making latitude and costs of managing the exchange. Interorganizational exchange theory (e.g., Levine & White, 1961) emphasizes cooperation, collaboration, and coordination of activities to achieve reciprocal benefits. As a result, it could represent an exogenous influence on the *decision-facilitating* objectives of MAS design. The asymmetry and reciprocity contingencies are each measured using two items (presented in Table 2) that were developed in this research based on the theoretical tenets each assumes.

Interorganizational performance is measured using a number of items capturing perceptions of alliance performance. Alliance performance was measured from the perspective of the focal firm using a set of items that capture the strategic benefits of the alliance. Similar measures of performance have been used in past interorganizational studies. Past research has suggested that alliance performance can be assessed by the extent to which the relationship is productive or worthwhile (Heide & Miner, 1992; Van de Ven & Walker, 1984). Others captured performance by measuring the extent to which the alliance contributes to profits, market share, or competitive advantage (Parkhe, 1993; Simonin, 1997), whereas Young-Ybarra and Wiersema (1999) measured alliance performance an overall expectation. As a result, six items were developed in this study to capture the financial and strategic dimensions of interorganizational alliance performance; in addition, three items are used to assess the perception of overall alliance performance.

In addition, one overall item of exchange partner performance is developed in this study to supplement the alliance performance measures. Exchange partner performance has been defined in past research as the extent to which the supplier has fulfilled the buyer's requirements in terms of price, timeliness of delivery, input quality, and supplier flexibility (Zaheer, McEvily, & Perrone, 1998). This study adapts this definition to develop the one item measuring exchange partner performance.

Furthermore, to control for a potential "halo" effect on performance that could confound the main hypothesized relationships (e.g., Brown & Perry, 1994), I include a measure of prior performance in the model. To the extent that the MAS objectives would maintain their effects on interorganizational performance even after controlling for prior performance, the reliability of the overall model would be enhanced. As mentioned in Table 2, I measured prior performance using the respondent's estimate of their firm's and corresponding industry's revenue growth observed during the time of the alliance.

Statistical Models

I examine the research hypothesis using a simultaneous equation model that is best suited to testing complementary relationships (Athey & Stern, 1998). The model includes the determinants of each of the endogenous variables (decision-influencing and decision-facilitating objectives) and their interrelation. To ensure identification, I also include the exogenous factors of asymmetry contingency (ASYMM) and reciprocity contingency (RECIP) in

each of the two models. H1 predicts that *decision-influencing* (DI) and *decision-facilitating* (DF) objectives of MAS design will be complementary choices, thus interrelated, and are also jointly determined by integrated information systems (IIS). Following past studies that tested similar complementary relations (e.g., Abernethy, Bouwens, & van Lent, 2004), I use the following system of equations to test this hypothesis:

$$DI_i = \alpha_0 + \alpha_1 DF_i + \alpha_2 IIS_i + \alpha_3 ASYMM + \varepsilon_i^{DI} \quad (1a)$$

$$DF_i = \beta_0 + \beta_1 DI_i + \beta_2 IIS_i + \beta_3 RECIP + \varepsilon_i^{DF} \quad (1b)$$

Ordinary least squares (OLS) estimates could be biased and inconsistent when endogenous variables appear as regressors in other equations in the system (Wooldridge, 2000), which might be a more frequent problem in models with complementary inputs (Athey & Stern, 1998). I use the Durbin–Wu–Hausman test (MacKinnon, 1992; Nakamura & Nakamura, 1981) to determine the presence of simultaneity equation bias that may be caused by potentially correlated error terms. In both models earlier, I find no evidence of simultaneity bias (Model 1a: $F=0.20$, $p<0.00$; Model 1b: $F=2.32$, $p<0.13$). For a robustness check, I have also estimated the preceding system of equations using two-stage least squares (2SLS) estimation. All results obtained using 2SLS reinforce the results obtained by OLS estimation. As a result, OLS will be used in the analysis as the 2SLS approach may be sensitive to weaknesses in instrumental variables (Nelson & Stratz, 1990).

Research hypothesis H1 predicts that the endogenously determined decision influencing and decision facilitating will each influence interorganizational performance (IOPRF). As a result, the following two equations are also examined using OLS:

$$IOPRF_i = \gamma_0 + \gamma_1 DI_i^{\wedge} + \gamma_2 \log(PRF)_i + \gamma_3 \log(PRI)_i + \varepsilon_i^{IOPRF} \quad (2)$$

$$IOPRF_i = \delta_0 + \delta_1 DF_i^{\wedge} + \delta_2 \log(PRF)_i + \delta_3 \log(PRI)_i + \varepsilon_i^{IOPRF} \quad (3)$$

where,

DI^{\wedge} and DF^{\wedge} are predicted values from Eqs. (1a) and (1b), respectively; PRF represents the control measure of prior performance of firm in the alliance;

PRI represents the control measure of prior performance of corresponding industry in which an alliance operates; and the logarithm of PRF and PRI is taken to correct for score range.

DATA ANALYSIS AND RESULTS

The measurement properties of the items have been examined and tested for convergent and discriminant validity (Boudreau, Gefen, & Straub, 2001). Convergent validity means how well each latent construct captures the variance in its measures. Convergent validity can be evaluated by examining the following measures: individual item reliability (standard is 0.5 or above); composite construct reliability and a measure similar to Cronbach's alpha (standard is 0.7 or above); and average variance extracted (AVE), which measures whether the variance captured by a construct is larger than the variance due to measurement error (standard is 0.5 or above) (Fornell & Larcker, 1981). Table 3 notes that all internal consistency reliability (ICR) coefficients met the 0.7 standard, whereas all constructs also met the 0.5 AVE criterion, supporting convergent validity. The Cronbach's alpha is also shown for each construct for comparative purposes and the same conclusions apply.

Discriminant validity means the extent to which measures of constructs are empirically distinct (Davis, 1989). I assessed discriminant validity by comparing the square roots of the AVE of two measured constructs (notes on the Table 3 diagonal) to the correlation between each pair of constructs. This test is satisfied by all construct pairs in the model. A stricter test of discriminant validity requires that the absolute value of the AVE of each construct is higher than its correlation to any other measured construct. As mentioned in Table 3, this stricter test of discriminant validity is met by all constructs in the model.

Testing Complementarity in Interorganizational MAS Objectives and Effects on Interorganizational Performance

The research hypothesis specifies the complementary nature of decision-influencing and decision-facilitating objectives of MAS design and predicts their effect on interorganizational alliance performance. Models 1a and 1b test the complementary nature of the two constructs and whether they are jointly determined by integrated information systems. These models correspond to the "adoption tests" for testing complementarity as specified by Athey and Stern (1998). Table 4 summarizes the OLS estimation results.

The results of Eqs. (1a) and (1b) (Table 4, panel A) provide strong evidence that decision-influencing objectives of MAS design are positively and significantly related to decision-facilitating objectives ($\alpha_1 = 0.64$; $t = 8.00$),

Table 3. Descriptives, Correlations, and Validity Statistics.

	Mean	Standard Deviation	1	2	3	4	5	6	7
1 IIS	4.68	1.66	0.92 [§]						
2 DI	4.48	1.47	0.531	0.910					
3 DF	4.60	1.36	0.491	0.731	0.862				
4 IOPRF	5.16	1.22	0.271	0.326	0.427	0.944			
5 PRIORPRF	18.09	24.57	0.226	0.288	0.178	0.330	0.926		
6 ASYMM	3.32	1.78	0.269	0.296	0.311	0.233	-0.082	0.935	
7 RECIP	5.44	1.51	0.282	0.229	0.317	0.258	0.060	-0.050	0.927
ICR*			0.943	0.907	0.896	0.971	0.923	0.933	0.924
Cronbach's alpha			0.716	0.706	0.701	0.709	0.730	0.856	0.837
AVE [§]			0.846	0.829	0.743	0.892	0.857	0.874	0.860

Notes: Correlations greater than |0.20| are significant at $p < 0.05$; correlations greater than |0.25| are significant at $p < 0.01$. ASYMM, asymmetry contingency; DF, decision-facilitating objective of MAS design; DI, decision-influencing objective of MAS design; IIS, integrated information systems; IOPRF, interorganizational performance; PRIORPRF, prior performance; RECIP, reciprocity contingency.

*ICR = Internal Consistency Reliability coefficient.

[§]AVE = Average Variance Extracted estimate (cf. Fornell & Larcker, 1981).

[§]Diagonal elements are the square root of the average variance extracted (AVE) estimate for each construct. Off-diagonal elements are the correlations between the different constructs.

while decision-facilitating objectives are in turn also positively and significantly related to the setting of decision-influencing objectives of MAS design ($\beta_1 = 0.57$; $t = 8.24$). The two objectives are thus interrelated and are jointly determined by integrated information systems (IIS: $\alpha_2 = 0.20$, $t = 3.11$; $\beta_2 = 0.10$, $t = 1.59$), while both models exhibit high explanatory power (model 1: adj. $R^2 = 54.38\%$; model 2: adj. $R^2 = 52.91\%$). Even though the effect of integrated information systems on decision-facilitating objective was very marginally significant ($p = 0.11$), the joint effect of integrated information systems on both decision-influencing and decision-facilitating objectives was not significantly different. A system test of equality in the effects of integrated information systems on decision influencing and decision facilitating in the 2SLS model was not rejected ($F = 1.81$; $p < 0.1796$), thus supporting the joint effects of integrated information systems on both decision-influencing and decision-facilitating objectives of MAS design.

Eqs. (2) and (3) use the predicted values of decision influencing and decision facilitating from Eqs. (1a) and (1b) to test the complementary

Table 4. Tests of Complementary Relations.

Panel A: Ordinary Least Squares Regressions to Test Complementarity between DI and DF Objectives				
Predictor	Coefficient	Std. error	<i>t</i> -Statistic	Probability (two-sided)
Model 1a: $DI_i = \alpha_0 + \alpha_1 DF_i + \alpha_2 IIS_i + \alpha_3 ASYMM + \varepsilon_i^{DI}$ (1a)				
Intercept	0.46	0.36	1.30	0.20
DF	0.64	0.08	8.00	0.00
IIS	0.20	0.06	3.11	0.00
ASYMM	0.04	0.06	0.81	0.42
$F = 46.70; p < 0.0001$; Adj. $R^2 = 54.38\%$				
Model 1b: $DF_i = \beta_0 + \beta_1 DI_i + \beta_2 IIS_i + \beta_3 RECIP + \varepsilon_i^{DF}$ (1b)				
Intercept	0.95	0.39	2.45	0.02
DI	0.57	0.07	8.24	0.00
IIS	0.10	0.06	1.59	0.11
RECIP	0.11	0.06	1.87	0.06
$F = 44.06; p < 0.0001$; Adj. $R^2 = 52.91\%$				
Panel B: Ordinary Least Squares Regressions to Test Effect of Complementary MAS Objectives on Interorganizational Performance				
Predictor	Coefficient	Std. error	<i>t</i> -Statistic	Probability (two-sided)
Model 2: $IOPRF_i = \gamma_0 + \gamma_1 DI_i^\wedge + \gamma_2 \log(PRF)_i + \gamma_3 \log(PRI)_i + \varepsilon_i^{IOPRF}$ (2)				
Intercept	2.26	0.46	4.96	0.00
DI^\wedge	0.53	0.10	5.35	0.00
Log(PRF)	0.49	0.31	1.60	0.11
Log(PRI)	0.12	0.30	0.41	0.68
$F = 16.08; p < 0.0001$; Adj. $R^2 = 33.46\%$				
Model 3: $IOPRF_i = \delta_0 + \delta_1 DF_i^\wedge + \delta_2 \log(PRF)_i + \delta_3 \log(PRI)_i + \varepsilon_i^{IOPRF}$ (3)				
Intercept	2.86	0.51	5.64	0.00
DF^\wedge	0.38	0.12	3.26	0.00
Log(PRF)	0.61	0.33	1.84	0.07
Log(PRI)	0.01	0.33	0.02	0.99
$F = 9.08; p < 0.0001$; Adj. $R^2 = 21.21\%$				

Notes: ASYMM, asymmetry contingency; DF, decision-facilitating objective of MAS design; DF^\wedge , predicted value of DF from Eq. (1b); DI, decision-influencing objective of MAS design; DI^\wedge , predicted value of DI from Eq. (1a); IIS, integrated information systems; IOPRF, interorganizational performance; PRF, control measure of prior performance of firm in the alliance; PRI, control measure of prior performance of corresponding industry in which alliance operates; PRIORPRF, prior performance; RECIP, reciprocity contingency.

effects of MAS design on interorganizational performance. A common method of testing for complementarities is the productivity approach, which involves measuring the effect that decision-influencing and decision-facilitating objectives will have on performance and examining whether interactive terms will have larger effects than the main effects alone (Athey & Stern, 1998). A model where the main and interactive effects of both decision influencing and decision facilitating are present, however, may suffer from a selection bias problem (Athey & Stern, 1998). This selection bias would occur if firms that adopt integrated information systems expect greater returns if they focus on both objectives simultaneously than when they only emphasize decision-influencing and decision-facilitating objectives alone. If these objectives are indeed complements, then the error term would be correlated to the regressors. To avoid this problem, I use the predicted levels of decision-influencing and decision-facilitating objectives from Eqs. (1a) and (1b), and test their individual significance on interorganizational performance. The predicted values incorporate the hypothesized complementarities between the two objectives, given the level of enablement facilitated by the use of integrated information systems. Models (2) and (3) in panel B of Table 4 therefore provide a test for observed complementarities in the two objectives of MAS design, after controlling for the exogenous factor of prior performance. Both models provide significant results with regard to the complementary impact of the decision-influencing and the decision-facilitating objectives on interorganizational performance (Eq. (2): $\gamma_1 = 0.53$; $t = 5.35$; Eq. (3): $\delta_1 = 0.38$; $t = 3.26$), while the explanatory power of both models is at a satisfactory level (model 2: adj. $R^2 = 33.46\%$; model 3: adj. $R^2 = 21.21\%$). These results provide support for research hypothesis H1.

DISCUSSION AND IMPLICATIONS

This study argues that an organization's adoption and use of integrated information systems in interorganizational alliances enables complementary strategies in the monitoring and facilitation of the interorganizational relationship. This may imply that to the extent a firm follows such complementary objectives and executes effective MAS design choices, it may develop organizational capabilities for the strategic management of the interorganizational relationship and enhance interorganizational performance. In an environment where integrated information systems facilitate the design of MAS and the setting of complementary decision-influencing and

decision-facilitating objectives, the effective employment of such complementary objectives should influence performance in inter-firm relationships.

The study's results support its theoretical arguments. The study argues that even though objectives of MAS design are significant factors of interorganizational performance, their performance influence is best explained when their effects are examined in combination. Economic theory argues that complementary effects are observed when strategies are pursued in combination. The complementarity analysis has shown that the two constructs are in fact interdependent and their influence on performance is best analyzed by examining the second-order effects they help generate. The significant pairwise correlation between the decision-influencing and the decision-facilitating objectives ($r = 0.731$), as given in Table 3, indicates that the two constructs are highly correlated as they are pursued simultaneously. The simultaneous equations approach used to test the research hypothesis, nevertheless, resulted in more powerful significant second-order effects (as specified in research models (2) and (3) by the predicted values of the decision-influencing and the decision-facilitating objectives). The interpretation of these second-order effects is not too dissimilar from results obtained in related studies that use second-order factor analytic techniques (e.g., Rai et al., 2006). These findings suggest that the objectives of interorganizational decision influencing and decision facilitation constitute important managerial processes that determine an organization's success in the alliance relationship. The effective design of a MAS to attain both decision-influencing and decision-facilitation objectives thus helps develop organizational capabilities that enhance performance.

Overall, the findings of this study show that the availability of highly integrated information systems enables these strategic processes and enhances the complementarity of specific interorganizational direction-setting and monitoring activities. These results are consistent with theoretical arguments presented in the strategic management (e.g., Teece et al., 1997) and information systems (e.g., Sambamurthy et al., 2003) literatures. They also extend past findings in information systems integration (Bharadwaj, 2000) and supply chain integration (Rai et al., 2006). This study also extends past management accounting research that advocates systemic approaches to examining the effectiveness of MAS design (e.g., Chenhall, 2003; Chenhall & Langfield-Smith, 1998), and its results help extend the boundaries of management accounting research in the interorganizational strategic context.

This study demonstrates that it is the design and organization of information that is the major explanatory variable of governance choices in IT-enabled business relationships. Although inter-firm relationships depend

on IT for the integration of information flows among networked firms, it is the effective use of IT through the design of MAS, which confers success to interorganizational arrangements. This study offers some novel insights in this area, and future research could build on its findings to further examine the role of information systems use, and effective design of management accounting and control systems, in the success of interorganizational relationships.

REFERENCES

- Abernethy, M. A., Bouwens, J., & van Lent, L. (2004). Determinants of control system design in divisionalized firms. *The Accounting Review*, 79(3), 545–570.
- Abernethy, M. A., & Lillis, A. M. (1995). The impact of manufacturing flexibility on management control system design. *Accounting, Organizations and Society*, 20(May), 241–258.
- Abernethy, M. A., & Vangoni, E. (2004). Power, organization design and managerial behaviour. *Accounting, Organizations and Society*, 29, 207–225.
- Anderson, S. W., & Sedatole, K. L. (2003). Management accounting for the extended enterprise: Performance management for strategic alliances and networked partners. In: A. Bhimani (Ed.), *Management accounting in the digital economy* (pp. 36–73). Oxford, UK: Oxford University Press.
- Anthony, R. N., & Govindarajan, V. (2004). *Management control systems* (11th ed.). Boston, MA: Irwin.
- Athey, S., & Stern, S. (1998). *An empirical framework for testing theories about complementarity in organizational design*. Working Paper no. 6600. National Bureau of Economic Research.
- Baiman, S. (1982). Agency research in management accounting: A survey. *Journal of Accounting Literature* (Spring), 154–213.
- Baiman, S., & Demski, J. (1980). Economically optimal performance evaluation and control systems. *Journal of Accounting Research*, 18, 184–220.
- Bakos, J. Y., & Brynjolfsson, E. (1993). Information technology, incentives, and optimal number of suppliers. *Journal of Management Information Systems*, 10, 37–53.
- Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.
- Barua, A., Konana, P., Whinston, A. B., & Yin, F. (2004). Assessing net-enabled business value: An exploratory analysis. *MIS Quarterly*, 28(4), 585–620.
- Barua, A., & Mukhopadhyay, T. (2000). Information technology and business performance. In: R. W. Zmud (Ed.), *Framing the domains of IT management: Projecting the future through the past*. Cincinnati, OH: Pinnaflex Press.
- Bharadwaj, A. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly*, 24(1), 169–196.
- Boudreau, M., Gefen, D., & Straub, D. W. (2001). Validation in information systems research: A state-of-the-art assessment. *MIS Quarterly*, 25(March), 1–16.
- Brown, B., & Perry, S. (1994). Removing the financial performance halo from *Fortune's* “Most Admired” companies. *Academy of Management Journal*, 37(5), 1347–1359.

- Brynjolfsson, E., & Hitt, L. (2000). Beyond computation, information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, 14(4), 23–48.
- Brynjolfsson, E., Hitt, L., & Yang, S. (2002). Intangible assets: Computers and organizational capital. *Brookings Papers on Economic Activity*, 1, 137–181.
- Chalos, P., & O'Connor, N. G. (2004). Determinants of the use of various control mechanisms in US-Chinese joint ventures. *Accounting, Organizations and Society*, 29, 591–608.
- Chapman, C. S., & Kihn, L. (2009). Information system integration, enabling control and performance. *Accounting, Organizations and Society*, 34(2), 151–169.
- Chenhall, R. (2003). Management accounting systems design within its organizational context: Findings from contingency-based research and directions for the future. *Accounting, Organizations and Society*, 28, 127–168.
- Chenhall, R., & Langfield-Smith, K. (1998). The relationship between strategic priorities, management techniques, & management accounting: An empirical investigation using a systems approach. *Accounting, Organizations and Society*, 23, 243–264.
- Coletti, A., Sedatole, K. L., & Towry, K. L. (2005). The effect of control systems on trust and cooperation in collaborative environments. *The Accounting Review*, 80(2), 477–500.
- Das, T. K., & Teng, B. (2000). Instabilities of strategic alliances: An internal tensions perspective. *Organization Science*, 11(January–February, 1), 77–101.
- Davenport, T. (2000). *Mission critical, realizing the promise of enterprise systems*. Boston, MA: Harvard Business School Press.
- Davila, A., & Foster, G. (2005). Management accounting systems adoption decisions: Evidence & performance implications from early-stage/startup companies. *The Accounting Review*, 80(4), 1039–1068.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Dent, J. F. (1990). Strategy, organization and control: Some possibilities for accounting research. *Accounting, Organizations and Society*, 15(February), 3–25.
- Dillman, D. A. (1978). *Mail and telephone surveys, the total design method*. New York, NY: Wiley.
- Drake, A. R., Haka, S. F., & Ravenscroft, S. P. (1999). Cost system and incentive structure effects on innovation, efficiency and profitability in teams. *The Accounting Review*, 74(3), 323–345.
- Eishnhardt, K., & Martin, J. (2000). Dynamic capabilities, What are they? *Strategic Management Journal*, 21, 1105–1121.
- Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39–50.
- Garicano, L., & Kaplan, S. N. (2001). The effects of business-to-business e-commerce on transaction costs. *The Journal of Industrial Economics*, 49(December), 463–485.
- Garvin, D. A. (1998). The processes of organization and management. *Sloan Management Review* (Summer), 33–49.
- Grafton, J., Lillis, A. M., & Widener, S. K. (2010). The role of performance measurement and evaluation in building organizational capabilities and performance. *Accounting, Organizations and Society*, 35(7 (October)), 689–706.
- Grant, R. M. (1996). Prospering in dynamically-competitive environments, Organizational capability as knowledge integration. *Organization Science*, 7(4), 375–387.

- Gulati, R. (1995). Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances. *Academy of Management Journal*, 38, 85–112.
- Gurbaxani, V., & Whang, S. (1991). The impact of information systems on organizations and markets. *Comm. ACM*, 34(1), 59–73.
- Harrigan, K. R. (1988). Strategic alliances and partner asymmetries. *Management International Review* (Special Issue), 53–71.
- Heide, J. B., & Miner, A. S. (1992). The shadow of the future, Effects of anticipated interaction and frequency of contact on buyer-supplier cooperation. *Academy of Management Journal*, 35, 265–291.
- Indjejikian, R. J., & Matejka, M. (2006). Organizational slack in decentralized firms, The role of business unit controllers. *The Accounting Review*, 81(4), 849–872.
- Ireland, R. D., Hitt, M. A., & Vaidyanath, D. (2002). Alliance management as a source of competitive advantage. *Journal of Management*, 28(3), 413–466.
- Johnston, H. R., & Vitale, M. R. (1988). Creating competitive advantage with interorganizational information systems. *MIS Quarterly*, 12(2), 153–165.
- Kramer, R. M. (1999). Trust and distrust in organizations, Emerging perspectives, enduring questions. *Annual Review of Psychology*, 50, 569–598.
- Kumar, K., & Van Hilleberg, J. (2000). ERP, Experiences and evolution. *Communications of the ACM*, 43(April), 23–26.
- Langfield-Smith, K. (1997). Management control systems and strategy, A critical review. *Accounting, Organizations and Society*, 22(February), 207–232.
- Levine, S., & White, P. E. (1961). Exchange as a conceptual framework for the study of interorganizational relationships. *Administrative Science Quarterly*, 5, 583–601.
- Malone, T. W. (1987). Modeling coordination in organizations and markets. *Management Science*, 33(10), 1317–1332.
- Malone, T. W., Yates, J., & Benjamin, R. I. (1987). Electronic markets and electronic hierarchies. *Communications of the ACM*, 30(June), 484–497.
- MacKinnon, J. G. (1992). Model specification tests and artificial regressions. *Journal of Economic Literature*, 30(March), 102–146.
- Milgrom, P., & Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy, and organization. *The American Economic Review*, 80(June, 3), 511–528.
- Milgrom, P., & Roberts, J. (1992). *Economics, organization and management*. Englewood Cliffs, NJ: Prentice-Hall.
- Milgrom, P., & Roberts, J. (1995). Complementarities and fit: Strategy, structure, and organizational change in manufacturing. *Journal of Accounting & Economics*, 19(March–May), 179–208.
- Mintzberg, H. (1978). Patterns in strategy formulation. *Management Science*, 24, 934–948.
- Nakamura, A., & Nakamura, M. (1981). On the relationships among several specification error tests presented by Durbin, Wu, and Hausman. *Econometrica*, 49(6), 1583–1588.
- Nelson, C. R., & Stratz, R. (1990). The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one. *Journal of Business*, 63(1), 125–140.
- Nicolaou, A. I. (2004a). Quality of post-implementation review for enterprise resource planning systems. *International Journal of Accounting Information Systems*, 5(May, 1), 25–49.
- Nicolaou, A. I. (2004b). Firm performance effects in relation to the implementation and use of enterprise resource planning systems. *Journal of Information Systems*, 18(Fall), 79–105.

- Nicolaou, A. I., & Bhattacharya, S. (2006). Organizational performance effects of ERP systems usage: The impact of post-implementation changes. *International Journal of Accounting Information Systems*, 7(Spring, 1), 18–35.
- Oliver, C. (1990). Determinants of interorganizational relationships: Integration and future directions. *Academy of Management Review*, 15(2), 241–265.
- Otley, D., & Berry, A. (1980). Control, organization, and accounting. *Accounting, Organizations and Society*, 5(2), 231–246.
- Parkhe, A. (1993). Strategic alliance structuring, a game theoretic and transaction cost examination of interfirm cooperation. *Academy of Management Journal*, 36(4), 794–829.
- Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations*. New York, NY: Harper & Row.
- Rai, A., Patnayakuni, R., & Seth, N. (2006). Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS Quarterly*, 30(2), 225–246.
- Riggins, F. J., Kriebel, C. H., & Mukhopadhyay, T. (1994). The growth of interorganizational systems in the presence of network externalities. *Management Science*, 40(8), 984–998.
- Ross, J. W., & Vitale, M. R. (2000). The ERP revolution: Surviving vs. thriving. *Information Systems Frontiers*, 2(2), 233–241.
- Sambamurthy, V., Bharadwaj, A., & Grover, V. (2003). Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms. *MIS Quarterly*, 27(June, 2), 237–263.
- Saraf, N., Langdon, C. S., & Gosain, S. (2007). IS application capabilities and relational value in interfirm partnerships. *Information Systems Research*, 18(September), 320–339.
- Scott, J. E., & Vessey, I. (2000). Implementing enterprise resource planning systems: The role of learning from failure. *Information System Frontiers*, 2(2), 213–232.
- Simonin, B. L. (1997). The importance of collaborative know-how: An empirical test of the learning organization. *Academy of Management Journal*, 40(5), 1150–1174.
- Simons, R. (1995). *Levers of control, how managers use innovative control systems to drive strategic renewal*. Boston, MA: Harvard Business School Press.
- Simons, R. (2000). *Performance measurement and control systems for implementing strategy*. Upper Saddle River, NJ: Prentice-Hall.
- Soh, C., Kien, S. S., & Tay-Yap, J. (2000). Cultural fits and misfits: Is ERP a universal solution? *Communications of the ACM*, 43(April), 47–51.
- Sprinkle, G. B. (2003). Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society*, 28, 287–318.
- Sprinkle, G. B., & Williasmon, M. G. (2007). Experimental research in managerial accounting. In: C. S. Chapman, A. G. Hopwood & M. D. Shields (Eds.), *Handbook of management accounting research* (pp. 415–444). Oxford, UK: Elsevier.
- Stephanou, C. J. (2000). The selection process of enterprise resource planning (ERP) systems. *Proceedings of the Americas Conference on Information Systems (AMCIS)*, 988–991.
- Swamidass, P. M., & Newell, W. T. (1987). Manufacturing strategy, environmental uncertainty and performance: A path analytic model. *Management Science*, 33(April), 509–524.
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509–533.
- Tuttle, B., & Burton, F. G. (1999). The effects of a modest incentive on information overload in an investment analysis task. *Accounting, Organizations and Society*, 24, 673–687.
- Van de Ven, A. H., & Walker, G. (1984). The dynamics of interorganizational coordination. *Administrative Science Quarterly*, 29, 598–621.

- Wang, E. T. G., & Seidmann, A. (1995). Electronic data interchange: Competitive externalities and strategic implementation policies. *Management Science*, *41*(3), 401–418.
- Wooldridge, J. M. (2000). *Introductory econometrics, a modern approach*. Cincinnati, OH: South Western College Publishing.
- Young-Ybarra, C., & Wiersema, M. (1999). Strategic flexibility in information technology alliances: The influence of transaction cost economics and social exchange theory. *Organization Science*, *10*(4), 439–459.
- Zaheer, A., McEvily, B., & Perrone, V. (1998). Does trust matter? Exploring the effects of I-O and interpersonal trust on performance. *Organization Science*, *9*(March–April), 141–159.
- Zaheer, A., & Venkatraman, N. (1995). Relational governance as an interorganizational strategy: An empirical test of the role of trust in economic exchange. *Strategic Management Journal*, *16*, 373–392.
- Zimmerman, J. L. (2006). *Accounting for decision making and control* (5th ed.). New York, NY: McGraw Hill/Irwin.

ARE ENGAGEMENT QUALITY REVIEWS REALLY OBJECTIVE?

Ambrose Jones III, Carolyn Strand Norman
and Jacob M. Rose

ABSTRACT

We investigate auditor objectivity as it relates to engagement quality reviews by examining whether engagement quality reviewers (EQRs) exhibit lower levels of objectivity when they have administrative, economic, or social ties with the audit engagement partner. Motivated reasoning theory suggests that EQRs with ties to the engagement partner will reach less conservative conclusions and be more willing to accept an engagement partner's decision relative to reviewers who have no connections with the engagement partner. We conduct an experiment where EQRs must review a decision by an engagement partner related to a contingent liability.

Results suggest that engagement quality reviews are an effective mechanism for reducing the effects of engagement partner biases to accept client-favored accounting choices. Participants with ties to the engagement partner (i.e., from the same office) and without ties (i.e., from the national office) both challenged the decision of the engagement partner and recommended disclosure of a contingent liability, which client management opposed. We also find an interaction of ties with the engagement partner and the probability of the contingent liability. National office EQRs were less likely to decide that disclosure was

Advances in Accounting Behavioral Research, Volume 14, 143–164

Copyright © 2011 by Emerald Group Publishing Limited

All rights of reproduction in any form reserved

ISSN: 1475-1488/doi:10.1108/S1475-1488(2011)0000014009

necessary than were local office partners when the probability of the contingent liability was low. With regard to the need to recognize a liability, EQRs with and without ties to the engagement partner concurred with the decision of the engagement partner.

INTRODUCTION

Section 103 of the Sarbanes-Oxley Act of 2002 (SOX) includes a requirement for registered public accounting firms to provide a concurring or second partner review and approval of each audit report they issue (PCAOB, 2009). This mandate places renewed emphasis on audit quality and the importance of an objective “second look” in the conduct of public company audits. Responding to SOX, the Public Company Accounting Oversight Board (PCAOB) issued Auditing Standard No. 7, *Engagement Quality Review (AS No. 7)*, which provides a framework for a second reviewer to objectively evaluate the significant judgments made by the audit engagement team. Engagement quality reviews represent one of the final quality control procedures performed by the firm before issuing its report.¹ These reviews are unlike those that take place under peer review programs or PCAOB inspections, which are done on a selective basis after completion of the engagement and issuance of a report and are performed by individuals independent of the firm conducting the audit.

AS No. 7 became effective in 2010 and was issued after extensive deliberations over comments received from two iterations of proposals. This standard includes new guidance for practitioners, and it applies to all audit engagements and engagements to review interim financial information conducted pursuant to the standards of the PCAOB. AS No. 7 also contains requirements regarding the qualifications of the engagement quality reviewer (EQR), which include competence, integrity, objectivity, and independence. Furthermore, the EQR must be a person associated with the firm and must be a partner.

The purpose of our research is to examine whether engagement quality reviews are conducted objectively. This issue is examined in the context of the evaluation of engagement partner’s judgments about the severity and disposition of potential adjustments for a contingent liability arising under ambiguous circumstances. We investigate whether EQRs’ judgments are influenced by directional goals held in common with the engagement partner. The EQR is commonly from the same office or economic unit

(cluster of offices) within the firm as the audit engagement partner. Thus, the engagement partner and the EQR may have administrative, economic, or social ties that could affect the objectivity of the EQR. That is, although the EQR might meet the stated requirements of AS No. 7, that partner might not meet the spirit or substance of the standard, and, in addition, the EQR may receive pressure to agree with the conclusions of the engagement partner.

To evaluate the objectivity of the EQR, we test the tension between accuracy and directional goals by conducting an experiment. Motivated reasoning theory (Kunda, 1990) suggests that individuals committed to directional goals will engage in reasoning that is biased to meet those goals. In other words, if one conclusion is preferred over others, the individual will process and interpret the information to support the desired conclusion. For example, auditors often have the directional goal of supporting client preferred accounting and reporting methods and could thus exploit ambiguous situations and standards to justify agreement with the client, as long as this is deemed reasonable (Hackenbrack & Nelson, 1996; Kadous, Kennedy, & Peecher, 2003).

The case used in our experiment contains an environmental remediation issue that requires participants to determine the degree to which a contingent liability should be disclosed or recognized as a liability in the financial statements. Our 2×2 between-participants design incorporates two manipulations: the resident office of the EQR (same as the engagement partner or from the firm's national office) and a high or low probability of the need for environmental remediation. Results indicate that the requirement to perform engagement quality reviews can decrease the likelihood that auditors favor client-preferred accounting choices whether or not the EQR has administrative, economic, or social ties to the audit engagement partner. However, when the probability of the contingency was low, the national office partners were less likely to decide that disclosure was necessary than were local office partners. Concerning recognition, both groups reached similar conclusions for high and low probability conditions and concurred with the engagement partner.

The same participants also completed a debriefing questionnaire. Responses indicate that participants tend to believe that the purpose of the engagement quality review is to form an independent positive conclusion about the financial statements (as opposed to negative assurance), that the reviewer should have somewhat limited contact with the audit client, that social interactions between partners have a positive effect on firm productivity, and that adequate guidance is provided by the firm to conduct

engagement quality reviews. Additionally, participants are not likely to believe that EQRs try to avoid conflict with the engagement partner.

The remainder of this chapter is organized as follows. The next section reviews prior literature and develops our hypothesis. The third section identifies our research methods, followed by the results section, which discusses our findings from the experiment and the survey questions. The final section draws conclusions and identifies implications of our results.

PRIOR RESEARCH AND HYPOTHESIS DEVELOPMENT

PCAOB Auditing Standard No. 7

Audit failures and related issues that led to the passage of SOX and the establishment of the PCAOB resulted from audits that had concurring partner reviews based on previous requirements for such reviews. The fact that SOX mandated the PCAOB to adopt or develop a standard for concurring or second partner reviews signaled the importance of these reviews and Congressional intent to reevaluate the existing requirements (PCAOB, 2004).

Following AS No. 7, we use the term “engagement quality review” in this chapter (even when referring to prior literature that used “concurring review” or “second partner review”) because it provides a more accurate description of the nature and objectives of the review than the terms “concurring partner review” or “second partner review.” As explained by the PCAOB Standing Advisory Group (SAG), the term “concurring partner review” may not be the most appropriate term because it implies that the objective of the review is to “concur” with the lead partner and engagement team rather than to provide an objective, critical assessment. Also, the term “second partner review” may not be the most appropriate term because large engagements could include two or more partners on the engagement team, making the EQR not necessarily the “second partner” (PCAOB, 2004).

Previous rules for engagement quality reviews were contained in requirements of the SEC Practice Section of the AICPA (SECPS). The thrust of AS No. 7 is largely to make explicit requirements that were previously implied by the SECPS. For example, the prior guidelines described the EQR’s responsibilities generally in terms such as discussing significant high-risk accounting, auditing, and reporting matters with the engagement team,

reviewing documentation for unadjusted differences, and confirming with the engagement team the absence of any significant unresolved matters (PCAOB, 2003). AS No. 7 sets forth very specific tasks required to be *performed and documented* by the EQR, such as evaluating the engagement team: significant judgments made in planning the audit based on the firm's acceptance and retention process, assessment of and response to significant internal control risks, materiality decisions regarding identified misstatements, independence evaluation, report on internal controls, review of other information in documents containing financial statements to be filed with the SEC, consultations on difficult or contentious matters, and communications with the audit committee (PCAOB, 2009).

AS No. 7 also requires that the person conducting the engagement quality review be a partner or individual in an equivalent position and that the reviewer must possess adequate technical competence related to auditing, accounting, and financial reporting. Additionally, the reviewer is required to be independent (of the client), to perform the review with integrity, and to maintain objectivity in performing the review (PCAOB, 2009). However, the standard is silent regarding the relationship of the EQR to the engagement partner.

Engagement Quality Review

Schneider, Church, and Ramsay (2003) cite two reasons to conduct an engagement quality review. First, the engagement team may develop a positive effect toward their client. Specifically, the engagement partner's ongoing association with the client may lead to a more favorable attitude, which in turn may affect the partner's ability to make objective decisions. Thus, the review would be performed to minimize the impact of this positive effect because the EQR generally has less interaction with the client than does the engagement partner.

Second, because the engagement partner participates in decisions made throughout the audit process, he or she may be reluctant to overturn those decisions in the final review stage (when the engagement quality review takes place). Earlier research in behavioral decision-making and auditing suggests that prior involvement in a task can affect subsequent judgments and decisions (e.g., Brockner, 1992; Brody & Kaplan, 1996; Church, 1991; Church & Schneider, 1993). Psychologists refer to this as "escalation," which is the tendency for a decision-maker to persist with a failing course of action (Brockner, 1992). The more an individual has invested in a course

of action (psychologically, materially, or both), the more unwilling the decision-maker will be to abandon a course of action that followed from his or her decisions (Festinger, 1957).

Epps and Messier (2007) investigated the engagement quality review practices of the six largest public accounting firms and found moderate consistency among firm policies. Regarding the qualifications of the EQR, firm policies are often patterned on prior guidance from the SECPS. The authors noted a few instances where firms went beyond SECPS guidance, but in no cases were any prohibitions noted about the EQR having administrative, economic, or social ties to the engagement partner. As a result, Epps and Messier (2007) call for future research to investigate this issue, suggesting that researchers study various attributes of the concurring partner. In particular, what characteristics make an EQR more effective? The present study responds to Epps and Messier's (2007) concerns regarding quality reviews.

Academic research that investigates engagement quality reviews is limited, and previous studies typically use relatively small sample sizes of auditors qualified to perform an engagement quality review or samples of individuals who are not qualified to perform engagement quality reviews (e.g., audit seniors or students). Some studies examine the responsibilities of the reviewer and the review process and procedures (Epps & Messier, 2006; Schneider et al., 2003). Other studies investigate whether the review process affects the behavior of the engagement partner. For example, Entwistle and Lindsay (2002) find that engagement quality reviews have no effect on engagement partners' decisions about client accounting treatments, whereas Tan (1995) finds that an impending review heightens engagement partner awareness. Matsumura and Tucker (1995) develop an analytic model of the review, arguing that it should improve independence, but experimental testing of the model results in only the reduction, not elimination, of engagement partner bias. Ayers and Kaplan (1998, 2003) study client acceptance and engagement risk assessments involving an EQR, finding that client acceptance decisions by the EQR are more conservative than the engagement partner.

As suggested by Messier, Kozloski, and Kochetova-Kozloski (2010), research is needed about the competence, independence, and objectivity of the person performing the quality review. Early studies found that reviewers with industry experience were more effective in detecting financial statement errors (Johnson, Jamal, & Berryman, 1991) and that reviewers, even though subject to incentive schemes designed to promote independent reporting, still exhibit a reporting bias (Tucker & Matsumura, 1997). Others claim that

objectivity may be compromised by the length of association between the EQR and the client. For example, Favere-Marchesi and Emby (2005) investigate the effects of prior involvement of the reviewer with a client and find that continuing reviewers are less likely than new reviewers to disagree with the engagement partner when the evidence does not support the engagement partner. Thus, the EQRs' continuity with clients seems to negatively affect their objectivity. Favere-Marchesi and Emby's results were consistent with Tan's (1995) findings that an auditor's prior involvement with a client will increase the auditor's attention to social pressures to remain consistent with initial conclusions.

Kraut and Davidson (1998) find that engagement partner tendencies to issue qualified opinions are influenced by the percentage of client fees to office fees as well as by the percentage of fees received from all clients in the same industry. These authors also find that review partners are less affected by the relative importance of client and industry fees. However, their experiment did not deal with the issue of whether the review partner was (or was not) linked by common administrative, economic, or social ties to the engagement partner.

Hypothesis Development

This study extends existing research by addressing the issue of possible conflict residing within the mind of the EQR arising from a goal to be accurate, when accompanied by directional goals (i.e., pressure to "concur" with the engagement partner's conclusions). Often, the EQR and the engagement partner are from the same office or economic unit within the firm. Because of this, the engagement partner and the EQR could have administrative, economic, or social ties that could create pressure for the EQR to agree with the engagement partner's conclusions. For example, one of the two partners may evaluate the performance of the other partner, the two partners could have common economic interests, and the two partners could have close social ties outside of the work environment.

Public accounting firms often base partner compensation on the profitability of the local office or economic cluster of offices, making it likely that the loss of a significantly profitable client would affect everyone in that office or economic unit (Trompeter, 1994). Client retention may also result in "bragging rights" for engagement partners, thereby increasing prestige and status. Furthermore, retaining a high-profile client may result in additional business for the firm (Matsumura & Tucker, 1995). Economic dependence

could dominate in local offices because local office partners are the primary beneficiaries of the client revenues they generate, whereas the costs of litigation and reputation loss are incurred by the entire firm (Reynolds & Francis, 2000). These factors increase the possibility that, if the engagement partner and the EQR have common economic incentives, they will also have the common directional goal of retaining profitable clients for the firm. Conversely, an EQR without these ties to the engagement partner may be less biased and more objective in reaching audit conclusions and more attentive to accuracy goals.

Citing cognitive dissonance theory, Kunda (1990) makes a case for “motivated reasoning,” which is the notion that motivation (i.e., the wish, desire, or preference for a desired outcome) may affect reasoning through a biased set of cognitive processes. According to cognitive dissonance theory (Festinger & Carlsmith, 1959), individuals have a tendency to seek agreement on beliefs and opinions. When inconsistency (dissonance) between attitude and behavior exists, attitude is likely to change to accommodate the behavior. Auditors, to achieve the goal of supporting client choices and maintaining the auditor/client relationship, have been shown to interpret ambiguous situations in a way that favors the client (Hackenbrack & Nelson, 1996).

An alternative to motivated reasoning (directional goals) would be reasoning driven by accuracy goals (Kunda, 1990). This suggests that when people are motivated to be accurate, they choose more complex and time-consuming strategies to arrive at the correct answer. Kunda (1990) cites significant research supporting the notion that when people process information carefully, having no reason to prefer one conclusion over another, their goal will be accuracy. Applied to an audit setting, this would indicate that when a client is taking an aggressive position to avoid disclosure or recording a liability, auditors with accuracy goals will be more conservative and will not support a client’s desire for aggressive accounting choices. This might also support the case for selecting an EQR who has no reason to advocate one particular conclusion over another, such as a partner from the firm’s national office.

Matsumura and Tucker (1995) concluded that EQRs have a greater incentive to make decisions that are not biased and are more conservative than engagement partners, but this assumes that the partners do not have common economic interests. In fact, an EQR without common economic incentives may be more driven to maintain a high reputation for their technical and reviewing expertise (accuracy). The preceding discussion suggests the following hypothesis:

H₁. EQRs who have no administrative, economic, or social ties with the engagement partner (national office) will reach more conservative conclusions relative to EQRs who have administrative, economic, or social ties with the engagement partner (local office).

RESEARCH METHOD

Development of Case

The materials used in our study were designed to present a realistic situation where auditors are faced with the challenge between making an accurate decision and the directional goal of keeping a satisfied client in an ambiguous setting. The case portrays a situation in which an important public company client of a national Certified Public Accountant (CPA) firm receives late-breaking news about a contingent liability. The client resists any disclosure or recognition for the contingency, and the audit engagement partner has agreed with the client's position. On the basis of previous research, auditors will have the directional goal of supporting client-preferred accounting and reporting methods and may exploit ambiguous situations and standards to justify agreement with the client, as long as this is deemed reasonable (Hackenbrack & Nelson, 1996; Kadous et al., 2003). Therefore, we establish two important criteria – directional versus accuracy goals and ambiguity.

The hypothetical CPA firm is described as a national firm with offices throughout the United States and with international affiliations. The firm is divided into 10 geographic regions, and partner compensation is based on the profitability of the local office and region along with the cross-evaluations of other partners in the local office and region. The office conducting the audit is described as one where the partners get along both professionally and socially. The partners emphasize teamwork in the local office/region and take great pride in achieving that.

The case describes a scenario where the client is faced with the potential of environmental remediation of a former manufacturing site that is being sold. Information about the potential remediation was discovered by the client and the audit engagement team during subsequent event procedures at the end of audit fieldwork. An environmental consulting firm determines that there is very little precedent for this type of remediation and that more testing is necessary to confirm the actual need for the remediation

along with a precise cost estimate. However, based on the information available at the time, a cost range is quoted, with the low end of the range bordering on the materiality amount used by the firm in conducting the audit.

Citing that the consultant cannot provide a more precise, reliable estimate of the need for the remediation and potential cost for several more months, management does not believe that it is necessary to disclose the contingency or to recognize a liability, and the engagement partner has agreed with the client's position. The applicable accounting standard is SFAS No. 5 *Accounting for Contingencies* (FASB, 1975). The accounting and disclosure requirements of SFAS No. 5 are determined based on classifying the contingent liability as remote (no disclosure), reasonably possible (disclosure), or probable (disclosure and recognition). The auditors are thus faced with an ambiguous situation, requiring a decision about whether a liability and/or disclosure are necessary in the financial statements.

Following the case, the participants responded to a series of questions that were included to gain additional insight regarding accountants' general beliefs and opinions regarding the conduct of a quality review (in general) and certain information regarding their firm.

Previous versions of the case and survey questionnaire were distributed to audit partners, directors, and managers of several national firms seeking input about the situation portrayed. To ensure a realistic scenario, a pilot study was also conducted with current or former partners of several national CPA firms. Pilot study participants indicated that they had no problems interpreting the case or the related survey questions and agreed that the hypothetical firm described in the case exhibited characteristics similar to their firm. Comments and suggestions from these participants were incorporated into the revised version of the case and survey questions, which were then used for our study.

Design

The experiment employed a 2×2 between-participants design. The independent variables were (1) office location of the EQR (same home office as the audit engagement partner or national office) and (2) the level of probability that the remediation would be required, as expressed by the environmental consulting company based on information available at the time (15% or 55% to represent the low and high end points of the "reasonably possible" classification in SFAS No. 5). We expect that

participants assigned to the “same home office as the audit engagement partner” group will have more incentive to concur with an engagement partner’s decision to not require disclosure of the contingency or recognition of a liability, while the “national office” group will be more attentive to achieving accuracy. The second manipulation is level of probability (15% or 55%) that the remediation would be required as expressed by the environmental consulting firm. These levels of probability represent the low and high end points of the “reasonably possible” range for contingent liability disclosures and accrual recognition. The manipulation allows us to determine whether EQRs become more likely to deviate from an engagement partner’s decision to not require disclosure/recognition when the probability of remediation increases. Previous research (see Amer, Hackenbrack, & Nelson, 1995; Harrison & Tomassini, 1989; Jiambalvo & Wilner, 1985; Raghunandan, Grimlund, & Schepanski, 1991; Reimers, 1992) supports the use of these probability expressions as the anchor points for the reasonably possible expressions, that is, flanked by the threshold between remote and reasonably possible (15%) and the threshold between reasonably possible and probable (55%).

After reading and analyzing the case materials, participants were asked to answer two questions, which served as the dependent variables for the study. Participants assumed the role of an EQR and indicated the degree to which they believed that the company should be required to *disclose* a contingent loss associated with the environmental remediation in the notes to the financial statements. The second question asked participants to indicate the degree to which they believed that the company should be required to *recognize* a liability associated with the environmental remediation in the financial statements. Both questions used an 11-point Likert-type scale (end points were definitely not required/definitely required). Part two of our study included questions that were used to obtain a better understanding of the participants’ beliefs and opinions regarding quality reviews, thus helping to inform our results.

Participants

The participants in our study were 159 partners, directors, and managers from various U.S. locations of a national public accounting firm. These participants have the level of expertise that is necessary to ensure understanding and familiarity with quality reviews and the process surrounding this task. On average, the participants had 17.2 years of experience (standard deviation

of 8.9) and had conducted an average of 6.1 engagement quality reviews. Additional demographic information is presented in [Table 1](#).

Procedure

Data for the study were collected online. A link to the study materials was included in an e-mail message sent to the managers, directors, and partners of the firm in various offices around the United States by the firm's National Director of SEC Services. Responses were collected anonymously using a web-based survey collection company, which was also used by the subject firm for other needs. Thus, the participants were familiar with the procedure that was used for the present study. Privacy was assured because of built-in controls by the web-based survey collection firm. Participants were first asked to make two assumptions (the independent variables) regarding the case that they were to read, and then they responded to the two questions following the case. After these materials were completed, the participants were not allowed to go back to an earlier screen. Finally, participants were instructed to go to part two of the study, which included the following: the manipulation checks, the debriefing questionnaire, and the demographic questions.

Table 1. Demographic Information for Participants.

	Sample	
	Number	Percent
Gender		
Male	100	62.9
Female	48	30.2
Did not answer	11	6.9
Total	159	100.0
Position in the firm		
Partner	64	40.3
Senior manager or director	46	28.9
Manager	41	25.8
Did not answer	8	5.0
Total	159	100.0
	Mean	Standard Deviation
Public accounting experience (years)	17.2	8.9
Tenure with current firm (years)	11.2	8.3
Second partner reviews completed	6.1	6.3

RESULTS

Preliminary Analyses

In conducting our analyses, we did not include nine participants who failed the manipulation checks. As a result of eliminating those observations, the final sample size for our analyses is 159 participants. Descriptive analyses of the first dependent variable (disclosure) are summarized in Table 2, which provides means, standard deviations, and sample sizes for all treatment conditions. The pattern of means in Table 2 suggests that the participants, regardless of location, recommend disclosure and are more likely to do so when the consultant estimates a higher likelihood of need for remediation. The descriptive statistics suggest that EQRs tend to be objective regardless of their affiliation with the engagement partner, because EQRs from all offices were

Table 2. Descriptive Statistics for the Dependent Variable of Disclosure Recommendation.

Office Location of EQR	Probability of Remediation (%)	Mean ^a	Standard Deviation	<i>n</i>
Local	15	9.10	2.47	31
	55	9.54	2.25	35
	Total	9.33	2.35	66
National	15	8.02	3.21	52
	55	9.78	1.47	41
	Total	8.80	2.73	93
Total	15	8.42	2.99	83
	55	9.67	1.86	76
	Total	9.02	2.58	159

^aMeans are based on the following dependent variable scale (scale is converted to have endpoints of 1 and 11, with a midpoint of 6, for analysis):

Please indicate the degree to which you believe that SRI should only be required to disclose a contingent loss associated with the environmental remediation in the notes to its 2008 financial statements.

5	4	3	2	1	0	1	2	3	4	5	
Disclosure in notes not required					I am undecided		Disclosure definitely required in notes				

willing to go against the recommendation of the engagement partner (i.e., to not disclose or recognize a liability) and against the desires of the client.

Next, we conducted an analysis of variance (ANOVA) with disclosure as the dependent variable (Table 3). A preliminary analysis of covariance (ANCOVA) that included demographic variables as covariates indicated that “position in firm” (i.e., manager versus partner) significantly affects decisions to require disclosure. Partners were more likely to require disclosure than were managers. As a result, we include position in firm as a covariate in the model used to test our hypothesis. Although there is no main effect of office location, there is a significant interaction ($p = 0.05$). As noted from the graph (Fig. 1, panel A), participant reviewers from both the local and the national offices make very similar judgments about the need to *disclose* when the consultant’s probability estimate is high. However, when the probability estimate is low, office location makes a difference. Contrary to our hypothesis, however, the national office EQRs are less likely to decide that disclosure is necessary than are local office EQRs.

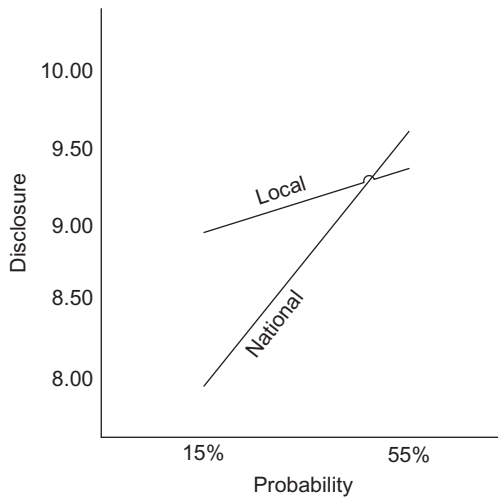
We also conducted a contrast test that examines the difference between local versus national office within the low probability condition. This is presented in the graph in Fig. 1 and, the table in, panel B, and it is statistically significant ($p = 0.059$). As one would expect from the graph, a similar contrast test that compares local to national EQRs in the high probability condition is not statistically significant ($p > 0.70$).

Accordingly, we conclude that office location only matters when the probability estimate is low, but the direction of the result is contrary to our

Table 3. ANOVA Results for Disclosure Recommendation.

Source	Type III Sum of Squares	df	Mean Square	F	Significance
Corrected model	105.118	4	26.280	4.108	0.003
Intercept	1,234.667	1	1,234.667	193.205	0.000
Office location of EQR	5.648	1	5.648	.883	0.349
Probability of remediation	44.046	1	44.046	6.886	0.010
Office location × probability	24.814	1	24.814	3.879	0.051
Position in firm	19.309	1	19.309	3.019	0.084
Error	933.875	146	6.396		
Total ^a	1,3252.000	151			
Corrected total	1,038.993	150			

^aEight participants failed to indicate whether they were partners or managers. Therefore, the final sample size for the ANOVA model with “position in firm” as a covariate is 151 participants.



Panel B: Contrast Tests of Local vs. National Office for Low Probability Condition

	Contrast	Value of Contrast	Standard Error	t-value	df	Significance (2-tailed)
Disclosure (Assume equal variances)	1	.539	.284	1.90	155	.059

Fig. 1. Interaction of Location and Probability for the Disclosure Recommendation.

hypothesized relationship whereby the EQRs from the national office will reach more conservative conclusions than the local office EQRs. Specifically, we hypothesized that local office EQRs will have closer relationships with the engagement partner and will therefore be more reluctant to go against the engagement partner’s decisions than will national EQRs. Although we cannot determine why national EQRs were less likely to recommend disclosure than were local EQRs, one question in the debriefing questionnaire suggests an avenue for future research. The debriefing question asked, “For the case that you just completed, how would disagreement with the engagement partner affect your relationships with your professional peers?” The scale response for this question is designed such that lower responses represent more negative perceptions of the effect of disagreement on professional relationships. The responses indicate that national EQRs (mean = 5.59) are more concerned ($p = 0.07$) that going against the original

partner's decision will adversely affect their relationships with peers than are local EQRs (mean = 5.90). Clearly, the differences in responses are relatively small, but additional research could examine whether concerns about relationships with other partners are important drivers of EQRs decisions.

Regarding our second dependent variable (recognition), descriptive analyses are reported in Table 4, including means, standard deviations, and sample sizes for all treatment conditions. Table 5 summarizes the results of an ANOVA for recognition. The participant EQRs, regardless of location, do not recommend recognition. There is no overall significant difference between local EQR's (mean = 3.89) and national EQR's (mean = 3.70) decisions to require recognition ($p = 0.80$). Furthermore, the pattern of means and the ANOVA indicate no significant differences between local and national EQRs' requirement of recognition, regardless of whether the consultant has provided a low or high probability estimate of the need to remediate. The findings suggest a high level of reluctance to go against

Table 4. Descriptive Statistics for the Dependent Variable of Recognition.

Office Location of EQR	Probability (%)	Mean ^a	Standard Deviation	<i>n</i>
Local	15	3.58	2.33	31
	55	4.17	2.83	35
	Total	3.89	2.61	66
National	15	3.21	2.27	52
	55	4.33	3.05	40
	Total	3.70	2.68	92
Total ^b	15	3.35	2.29	83
	55	4.25	2.93	75
	Total	3.78	2.64	158

^aMeans are based on the following dependent variable scale (scale is converted to have endpoints of 1 and 11, with a midpoint of 6, for analysis):

Please indicate the degree to which you believe that SRI should be required to recognize a liability associated with the environmental remediation in its 2008 financial statements?

5	4	3	2	1	0	1	2	3	4	5
Definitely not required					I am undecided					Definitely required

^bOne participant provided a response to the recognition dependent variable that was not interpretable. Therefore, the final sample size for the recognition analyses is 158 participants.

Table 5. ANOVA Results for Recognition as the Dependent Variable.

Source	Type III Sum of Squares	df	Mean Square	F	Significance
Corrected model	35.28	3	11.76	1.71	0.168
Intercept	2,224.83	1	2,224.83	322.63	0.000
Office location of EQR	0.44	1	0.44	0.06	0.800
Probability of remediation	27.65	1	27.65	4.01	0.047
Office location × probability	2.60	1	2.60	0.38	0.540
Error	1,061.97	154	6.90		
Total	3,353.00	158			
Corrected total	1,097.25	157			

another partner when recognition is involved. However, reflecting on SFAS No. 5 *Accounting for Contingencies* (FASB, 1975), a contingent liability is recognized if a loss is probable (the remediation is likely to occur) and the amount of loss can be at least reasonably estimated. Under these circumstances, it seems rational that EQRs (at either level) would agree with the engagement partner who is intimately familiar with the client and the specific circumstances surrounding the contingency. In other words, given the ambiguous nature of the case, not recognizing the liability may be the most appropriate decision.

Finally, we turn to our debriefing question analysis to gain more insight into public accountants' beliefs regarding engagement quality reviews. We report the findings of selected questions from the survey in Table 6. Results suggest that the purpose of the review is to form an independent positive conclusion about the financial statements, rather than give negative assurance; that the EQR should have somewhat limited contact with the audit client; that participants are more likely to believe that social interactions between partners have a positive effect on firm productivity; and that adequate guidance is provided by the firm to conduct engagement quality reviews. Surprisingly, participants are not likely to believe that EQRs try to avoid conflict with the engagement partner, regardless of the position level or experience of the EQR.

CONCLUSIONS

The purpose of this study was to investigate objectivity as it relates to engagement quality reviews. To explore this topic, we administered an

Table 6. Descriptive Statistics for Selected Debriefing Questions^a.

Question	<i>n</i>	Minimum	Maximum	Mean	Standard Deviation
Question #1 (positive conclusion)	148	1.00	11.00	7.60	3.16
Question #2 (avoid conflict)	152	1.00	11.00	4.51	2.87
Question #3 (client contact)	152	1.00	11.00	4.50	2.80
Question #4 (social interaction)	153	3.00	11.00	8.76	1.92
Question #5 (firm guidance)	152	1.00	11.00	7.66	2.93

^aAll scales below are converted to have endpoints of 1 and 11, with a midpoint of 6, for analysis.

Question #1. In general, do you believe the purpose of the second partner review is to form an independent positive conclusion about the financial statements or to provide negative assurance about the financial statements?

5	4	3	2	1	0	1	2	3	4	5	
Negative assurance					Neutral						Positive conclusion

Question #2. In general, do you believe that second partner reviewers try to avoid conflict with the engagement partner?

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
Rarely					I am undecided						Usually

Question #3. In general, do you believe that the second partner reviewer should have limited or extensive contact with the audit client?

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
Limited					I am undecided						Extensive

Question #4. In general, do you believe that positive social interactions between partners affect firm productivity?

5	4	3	2	1	0	1	2	3	4	5	
Very negative effect					No effect						Very positive effect

Question #5. In general, do you believe that adequate guidance is provided in your firm manuals for the conduct of a 2nd partner review?

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
Inadequate					I am undecided						Adequate

experiment and questionnaire to 159 partners, directors, and managers at a large national public accounting firm. We hypothesized that EQRs who have no administrative, economic, or social ties with the engagement partner will reach more conservative conclusions relative to those EQRs who have such connections with the engagement partner. We also analyzed debriefing responses to provide additional insights related to engagement quality reviews.

Our results suggest that engagement quality review as required by AS No. 7 can be effective in reducing the effects of engagement partner biases. Participants in both groups (national office or local office) made a more conservative decision than the engagement partner. However, when the probability of remediation was low, the national office EQRs were less likely to decide that *disclosure* is necessary than were local office EQRs. Pertaining to recognition, both groups reached similar conclusions for the high and low probability conditions.

Our analysis of the debriefing question responses suggests that participants are more likely to believe that (1) the purpose of the review is to form an independent positive conclusion about the financial statements, (2) the EQR should have somewhat limited contact with the audit client, (3) social interactions between partners has a positive effect on firm productivity, and (4) adequate guidance is provided by the firm to conduct engagement quality reviews. In addition, participants are not likely to believe that EQRs try to avoid conflict with the engagement partner.

These results have interesting implications for public accounting firms and regulators. For regulators, the results support the enhanced, rigorous engagement quality review requirements contained in AS No. 7, which was issued with the purpose of safeguarding against erroneous or insufficiently supported audit opinions so as to give investors assurance on the quality of audit engagements. Commenters to the PCAOB release containing AS No. 7 noted concerns that partners may experience internal pressures within the firm to provide concurrence with the conclusions of audit engagement partners (PCAOB, 2009). The results of our study help to respond to that concern. Specifically, the general requirements in AS No. 7 for the EQR to be independent of the client, while performing the review with integrity and objectivity, appear sufficient within the context of our experiment.

On the basis of the similarity of conclusions reached by auditors from the local office and the national office, we do not find support that EQRs make more aggressive reporting decisions when they have ties with the engagement partner. Perhaps, this is because a local office EQR believes that he or she has more to lose from the negative consequences of an audit failure than do national office EQRs. Their own wealth and reputation may be more directly

affected by adverse outcomes to the local office. Also, as our debriefing responses indicate, national partners may have greater concerns about how they are perceived by other partners than are local partners, and these concerns could counterbalance the effects of connections between engagement partners and EQRs. The mechanisms that potentially create differences between the judgments of local versus national EQRs offer opportunities for further investigation.

Our results may prove to be beneficial to practitioners faced with engagement quality review requirements for client entities. Having an EQR from the local office or practice unit may be more practical from a logistical and communicative perspective. A local office EQR will be available for consultation as audit issues arise and travel difficulties might be avoided. It may even improve the accountability of partners to each other during the audit process and in the administration of their practice. Additionally, unlike previous studies (e.g., [Tucker & Matsumura, 1997](#)) that used students as subjects, we used partners, directors, and managers from a national public accounting firm. Furthermore, our conclusions are conducted under the latest standard for engagement quality reviews (AS No. 7), as compared to previous requirements of the SECPS.

More research is warranted to understand the complex social interactions between audit partners in public accounting firms. Future research designs might benefit from including quality review partners from several unrelated public accounting firms to test the effects of firm culture and policies. Obviously, understanding the dynamics of engagement quality reviews and the value they add to the accuracy of financial statements is an important factor in maintaining public trust.

NOTE

1. Hereafter, we use the term engagement quality reviewer (EQR) consistent with terminology used in AS No. 7. These reviews have historically been referred to as “concurring reviews” or “second partner reviews.”

ACKNOWLEDGMENTS

We thank participants of the 2011 Auditing Midyear Conference, our discussant Jodi Gissel, and workshop participants at the University of North Carolina Greensboro for their valuable comments.

REFERENCES

- Amer, T., Hackenbrack, K., & Nelson, M. (1995). Context-dependence of auditors' interpretations of the SFAS No. 5 probability expressions. *Contemporary Accounting Research*, 12(1), 25–39.
- Ayers, S., & Kaplan, S. E. (1998). Potential differences between engagement and risk review partners and their effect on client acceptance judgments. *Accounting Horizons*, 12(2), 139–153.
- Ayers, S., & Kaplan, S. E. (2003). Review partners' reactions to contact partner risk judgments of prospective clients. *Auditing: A Journal of Practice and Theory*, 22(1), 29.
- Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review*, 17(1), 39–61.
- Brody, R. G., & Kaplan, S. E. (1996). Escalation of commitment among internal auditors. *Auditing: A Journal of Practice and Theory*, 15(1), 1–15.
- Church, B. K. (1991). An examination of the effect that commitment to a hypothesis has on auditors' evaluations of confirming and disconfirming evidence. *Contemporary Accounting Research*, 7(2), 513–534.
- Church, B. K., & Schneider, A. (1993). Auditor objectivity: The effect of prior involvement in audit programme design. *Accounting & Finance*, 33(2), 61–79.
- Entwistle, G. M., & Lindsay, D. (2002). Audit partner judgment: the effects of concurring review, compensation practices, and accounting flexibility. Presented at the *Accounting and Finance Association of Australia New Zealand Annual Meeting*.
- Epps, K. K., & Messier, W. F. (2006). *Improving the effectiveness of engagement quality review*. Working Paper. Georgia State University/Kennesaw State University, Atlanta, GA.
- Epps, K. K., & Messier, W. F. (2007). Engagement quality reviews: A comparison of audit firm practices. *Auditing: A Journal of Practice and Theory*, 26(2), 167–181.
- Favere-Marchesi, M., & Emby, C. (2005). The impact of continuity on concurring partner reviews: An exploratory study. *Accounting Horizons*, 19(1), 1–10.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, 58(2), 203–210.
- Financial Accounting Standards Board (FASB). (1975). Statement of Financial Accounting Standards No. 5: Accounting for Contingencies.
- Hackenbrack, K., & Nelson, M. W. (1996). Auditors' incentives and their application of financial accounting standards. *The Accounting Review*, 71(1), 43–59.
- Harrison, K. E., & Tomassini, L. A. (1989). Judging the probability of a contingent loss: An empirical study. *Contemporary Accounting Research*, 5(2), 642–648.
- Jiambalvo, J., & Wilner, N. (1985). Auditor evaluation of contingent claims. *Auditing: A Journal of Practice and Theory*, 5(1), 1.
- Johnson, P. E., Jamal, K., & Berryman, R. G. (1991). Effects of framing on auditor decisions. *Organizational Behavior and Human Decision Processes*, 50(1), 75–105.
- Kadous, K., Kennedy, J., & Peecher, M. E. (2003). The effect of quality assessment and directional goal commitment on auditors' acceptance of client-preferred accounting methods. *The Accounting Review*, 78(3), 759–778.
- Kraut, M., & Davidson, R. (1998). Effects of incentives and risks on decisions made by engagement and review partners. *Northwest Journal of Business and Economics*, 29–41.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480.

- Matsumura, E. M., & Tucker, R. R. (1995). Second partner review: An analytical model. *Journal of Accounting, Auditing & Finance*, 10(1), 173–200.
- Messier, W., Jr., Kozloski, T., & Kochetova-Kozloski, N. (2010). An analysis of SEC and PCAOB enforcement actions against engagement quality reviewers. *Auditing: A Journal of Practice and Theory*, 29(2), 233–252.
- Public Company Accounting Oversight Board. (2003). SECPS 1000.39 Appendix E. In Rule 3400T. Washington, D.C.
- Public Company Accounting Oversight Board. (2004). Potential standard-engagement quality reviews. PCAOB Standing Advisory Group Meeting, June 21–22, 2004, pp. 1–18.
- Public Company Accounting Oversight Board. (2009). Auditing Standard No. 7-Engagement Quality Review and Conforming Amendment to the Board's Interim Quality Control Standards. Release No. 2009-004. Washington, D.C.
- Raghunandan, K., Grimlund, R. A., & Schepanski, A. (1991). Auditor evaluation of loss contingencies. *Contemporary Accounting Research*, 7(2), 549–569.
- Reimers, J. L. (1992). Additional evidence on the need for disclosure reform. *Accounting Horizons*, 6(1), 36–41.
- Reynolds, J. K., & Francis, J. R. (2000). Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics*, 30(3), 375–400.
- Schneider, A., Church, B. K., & Ramsay, R. J. (2003). Concurring partner review: Does involvement in audit planning affect objectivity?. *Research in Accounting Regulation*, 16, 185–195.
- Tan, H.-T. (1995). Effects of expectations, prior involvement, and review awareness on memory for audit evidence and judgment. *Journal of Accounting Research*, 33(1), 113–135.
- Trompeter, G. (1994). The effect of partner compensation schemes and generally accepted accounting principles on audit partner judgment. *Auditing: A Journal of Practice and Theory*, 13(2), 56–68.
- Tucker, R. R., & Matsumura, E. M. (1997). Second-partner review: An experimental economics investigation. *Auditing: A Journal of Practice and Theory*, 16(1), 79–98.

DOES THE CHANGE TO PRINCIPLES-BASED ACCOUNTING INCREASE JUROR ASSESSMENTS OF AUDITOR LIABILITY?

John T. Sennetti, Charles P. Becker
and Howard J. Lawrence

ABSTRACT

This chapter investigates whether jurors, in their attribution of auditor responsibility, may be inappropriately influenced by the client use of a principles-based accounting standard, even if this standard is properly applied. Following prior research on questionable auditor conduct and its subsequent evaluation by juries, which is often subject to hindsight and outcome bias, this chapter examines whether an auditor's legal liability increases when its client uses principles-based accounting standards, by conducting a controlled experiment with 124 qualified jurors serving a county circuit court. Each juror is properly instructed and provided one of four different cases, obtained by manipulating two levels of an accounting standard, one principles-based and one rules-based, and by manipulating two subsequent client-loss outcomes, one moderately negative and one severely negative. This study finds jurors evaluate auditors more negatively if auditors have relied on a principles-based accounting standard. This attribution is influenced by hindsight bias and the perceived risk-taking

Advances in Accounting Behavioral Research, Volume 14, 165–189

Copyright © 2011 by Emerald Group Publishing Limited

All rights of reproduction in any form reserved

ISSN: 1475-1488/doi:10.1108/S1475-1488(2011)0000014010

responsibility of the investor, but independent of the client-loss outcome severity. These results contribute to the discussion of adopting or converting to the principles-based International Financial Reporting Standards (IFRS) by the United States.

INTRODUCTION

In one theory of juror decision-making, “jurors construct a narrative storyline out of the evidence presented during the trial” to explain their verdicts (Pennington & Hastie, 1981, 1986, 1988, 1993; Winter & Greene 2007). Jurors construct these stories not only with their prior knowledge and experience (Pennington & Hastie, 1988) but also with their intuitions (Hastie, 2008), “beliefs, attitudes, and cognitive capabilities” (Winter & Greene, 2007, p. 743.) This may lead jurors to inappropriately rely on ex post loss outcome information in rendering their verdicts, even in the presence of generally accepted auditing standards (GAAS) (Kadous, 2000).

This chapter extends this question to whether properly applied but principles-based accounting standards may unfairly bias juror’s verdicts. This issue is important given the anticipated conversion to the principles-based International Financial Reporting Standards (IFRS). The Securities and Exchange Commission (SEC) approved for release the proposed “Roadmap” for U.S. public companies to converge to some form of IFRS reporting as early as 2014, but, because of litigation, valuation, and cost-to-implement questions, the SEC subsequently delayed this action (Accounting Today Staff, 2010; Johnson, 2009; Leone, 2009; WebCPA, 2009). This delay occurred even though all but two other countries in the world “require or already have plans to permit IFRS” (Prophix Software, 2010, p. 3).

While having a single set of high-quality global accounting standards may provide a strong foundation for global capital markets, three of the major international auditing firms have requested litigation protection in defending their clients’ use of these more principles-based standards in the United States (Flynn, 2008; Heffes, 2008). These new standards increase the application of professional judgment and possibly increase the likelihood that any judgment will be challenged in court. The CEO of Grant Thornton admits accounting firms in other countries are not structured for US legal liability, and, while arguing for effective global regulation, he thinks any study of this is “a long way off” (Reilly, 2009, p. 35). Charles D. Niemeier, a board member of the Public Company Accounting Oversight Board,

“lambasted the SEC’s plan to move to IFRS, saying it would put the US regulatory system in jeopardy. All research shows that the US is unique in its regulation ... No [country] is as effective” (Reason, 2009, p. 36).

US accounting standards, which are based on “bright-line” rules such as historical cost, are easier for jurors to understand and audit clients to defend and differ from those based on principles. For example, Financial Accounting Standard No. 5 recognizes contingency losses of the client only if the losses are both estimable and probable, interpreted to be “known” and “likely,” making accounting choices easier to defend by the client (Amer, Hackenbrack, & Nelson, 1994, 1995; Nelson, 2003). But that may not be the case when the client chooses a principles-based standard, which has not-so-precise language the auditor must defend in light of subsequent negative outcomes: “When the auditor gets up on the witness stand and the plaintiff attorneys question him, he can say, ‘Well, it was a judgment call’” (Reason, 2009, p. 36). Such judgment calls are not considered favorable to the auditor and are believed to increase litigation risk.

Although rules-based accounting standards may limit litigation risk, they permit auditors more freedom to succumb to client pressure. Segovia, Arnold, and Sutton (2009) in a study of 114 experienced auditors find them “more willing to allow clients to manage earnings” under rules-based standards. Clients are more likely to attempt earnings management and auditors are less likely to require adjustment when clients use precise standards (Nelson, Elliott, & Tarpley, 2002).

Segovia et al. (2009) find regulation pressure may have a positive effect on less-experienced auditors who succumb to client pressure in a rules-based environment. Regulation advice from the FASB has been in favor of principles-based standards requiring more auditor judgment:

Preparers and auditors would need to apply professional judgment in more circumstances, while the SEC, investors, creditors, and other users of financial information must accept the consequences of applying professional judgment, including some divergence in practice. Concerns about SEC enforcement actions and related litigation matters are significant, potentially affecting the extent to which preparers and auditors would be willing to apply professional judgment in more circumstances. (Financial Accounting Standards Board [FASB], 2002, p. 9)

Unfortunately, even reasonable auditor judgment can be challenged unfairly by jurors who use additional, ex post victim-loss outcome knowledge (e.g., Latham & Linville, 1998). This is common in cases involving medical decisions, capital budgeting, and military situations (Hawkins & Hastie, 1990). Also, knowledge of larger damages may even increase the juror

assessment of liability. Kadous (2000) finds that an increase in the severity of the client-loss outcome increases the attribution of auditor responsibility.

To investigate these issues, an experiment is conducted to examine juror's decisions. Differences in juror attributions of auditor responsibility are measured for different case scenarios by manipulating the type of accounting standard (rules- vs. principles-based) and the client's subsequent investment-loss outcome (moderately negative vs. severely negative). This research finds jurors evaluate auditors more negatively if auditors have relied on a principles-based accounting standard, independent of the subsequent client-loss outcome, but dependent perhaps on perceived investor risk-taking responsibility. These results contribute to the discussion of the effects of the US adopting IFRS.

The rest of this chapter is organized as follows. The next section reviews the literature and develops the hypotheses. This is followed by the methods used to test the hypotheses, the results, and the conclusion that discusses the findings and limitations.

HYPOTHESES DEVELOPMENT

The Effect of the Accounting Standard

The standard audit opinion reports to users that the financial statements are presented in accordance with generally accepted accounting principles (GAAP) for the period observed. Subsequent to the date of this opinion, a finding of a material violation of GAAP for the audit period increases the likelihood of a lawsuit or a SEC action.

Auditors who permit the client to follow a rules-based standard of GAAP may be less liable. Rules provide actors (auditors) with ex ante guidance that encourages them to behave in a manner during the audit that avoids legal sanctions (Kaplow, 1999). Principles that use client ex ante judgment may provide plaintiffs with more opportunity to challenge auditor judgments after outcomes are known. Rules provide measures of protection, whereas principles create legal uncertainties (Kaplow, 1999). Research in medicine examining a similar type issue indicates that physicians are less negatively evaluated when they follow practice guidelines (i.e., rules) (Lawton & Parker, 2002).

When a plaintiff initiates a lawsuit against the auditors, the legal "story" (Hastie 2008; Pennington & Hastie, 1993) is usually one of negligence. Four facts are necessary to prove negligence: (1) the auditor has a duty of care,

(2) the auditor breached that duty, (3) the plaintiff suffered a loss, and (4) the auditor's breach of duty was the proximate cause of the plaintiff's loss (Causey & Causey, 1991). Plaintiff's attorneys will argue that the audit was substandard (breach of duty), and the substandard audit caused the plaintiff's losses (proximate cause). Defense attorneys will try to frame two arguments (Hastie, 2008): (1) the *auditor's activity* story, that the audit was conducted in accordance with GAAS, which permitted acceptable accounting standards, and, therefore, there was no breach of duty; and, (2) the *alternative explanation* story, the outcome was one of many possible alternative (and very specific) business outcomes possible given the specific client's actions and an audit conducted by GAAS.

When the claim of negligence centers on breach of duty, research suggests that auditors properly following GAAS may be following "moving targets" in the eyes of the jurors (Kadous, 2000). This adherence to auditing standards does not necessarily relieve the auditor of liability. When the thrust of the case centers on causality, research suggests that negligence cases may still move forward to trial even if causality is in doubt (Cloyd, Frederickson, & Hill, 1996). This creates legal uncertainty for auditors and other professionals facing negligence lawsuits (Kaplow, 1999). Auditor defenses against negligence, in addition to demonstrating the due care of following auditing standards, may include the use of decision aids (Lowe, Reckers, & Whitecotton, 2002) and may include other standards, such as those written by the government (Buckless & Peace, 1993).

Weiner's Framework

Weiner (1979) points to three influences in assigning attributions to the auditor: (a) locus of control (internal or external), (b) stability, and (c) control factors. Jurors may assign internal or external locus attributions to observed events. Internal attributions assign responsibility to the person, and external attributions assign responsibility to outside forces. Therefore, because auditors are forced to interpret principles-based accounting standards, this may result in more internal attributions to the auditor. Because the auditor is perceived as a professional, with specific skills and knowledge, and hence able to correctly interpret accounting principles, the auditor must be held more responsible. Contrarily, auditor reliance on "less-interpreted" rules-based accounting standards may result in more external than internal attributions.

Next, the audit is considered to have stability, because it is a well-planned and executed task. Auditors perform functions in which they are required to

be experienced and well trained and, under Sarbanes Oxley Act of 2002, are subject to federal oversight. Such stability of the environment results in more internal and less external attributions.

These locus and stability issues contribute to the third and more critical attribution, controllability. When control is possible, the observer assigns responsibility to the person, and when not, the observer assigns responsibility to external forces. Auditors are perceived to have more control over more choices when standards are principles-based, as there may exist many reasons for selecting any one of the choices that could be made. Hence, the juror's perceived internal locus, stability, and control attribution may result in causal attributions toward the auditor, applying [Weiner's \(1979\)](#) framework.

Kelley's Framework

[Kelley \(1973\)](#) suggests that the actor appearing in court is seen as more responsible for his/her actions. Accordingly, the auditor may receive attribution of blame from his/her actions that are not consistent, routine, or in consensus with others. For this issue, [Kelley \(1973\)](#) asks three general questions. Does the actor always respond in this manner (consistently)? Do other similar events result in the similar responses; are the responses routine or distinctive? Do other actors, given the same stimulus, react similarly, with a consensus? Tasks performed under rules-based rather than principles-based guidelines are more likely to be seen as consistent, routine, and follow consensus. Therefore, using either [Weiner's \(1979\)](#) or [Kelley's \(1973\)](#) heuristic models for causal attributions, the attribution of auditor responsibility should be less favorable when the auditor places reliance on a principles-based, rather than a rules-based, accounting standard.

The Effect of Language

One cause for these heuristic attributions may be found in the language of the standard. Principles-based standards can be understood by the juror as favoring the audit client (e.g., see [Amer et al., 1994, 1995](#); [Nelson, 2003](#).) Financial Accounting Standard No. 5 recognizes contingency losses of the client only if the losses are both estimable and probable, interpreted to be "known" and "likely," for example, almost certain ([Amer et al., 1994, 1995](#); [Nelson, 2003](#)). This language issue could occur in practice now if the bright-line rule for lease capitalization (lease term is "equal to 75 percent or more"

of expected life), which is consistent with Statement of Financial Accounting Standard (SFAS) No. 13 (FASB, 1976), is replaced by the *less precise* language of the principles-based criterion (lease term is “for the major part” of expected life) as found in International Accounting Standard (IAS) No. 17, the potential IFRS replacement (IASB, 1997). Next, this could also occur in the convergence process because the words such as “most likely life of the lease” are used in the suggested convergence of IFRS with GAAP (AICPA, 2011). This interpretation of principles-based language may not be clear to a juror who does not know that “likely” means, as with SFAS No. 5, almost certainty (and this also would be even more certain than of “most likely”). Jurors, when polled after their decisions were made on preponderance cases (where the weight is assumed to be greater than 51 percent), have been found to demand a weight of the evidence to be higher than 75 percent for the “more likely than not preponderance level” (Kaye, 1982). Even if the “most likely” language has not been tested, this language, too, should imply an even higher level of certainty than “more likely” (i.e., higher than 75 percent).

Given these heuristics found in Kelley (1973), Weiner (1979), and vague nature of the accounting language used in principles-based standards, the auditor may be seen by the juror as more in control of accepting principles-based accounting when favorable to the client.

H₁. Jurors will evaluate auditors’ performance less favorably when auditors rely on a principles-based, versus a rules-based, accounting standard.

Outcome Effects

The juror’s attribution of blame increases with hindsight bias (Tan & Lipe, 1997). Hindsight makes it easier to identify clear possible choices that could have made a difference in the audit outcome (Buchman, 1985, p. 269) and influences evaluative judgments (Hawkins & Hastie, 1990). Given the ex post chain of events, jurors assume that the subsequent events regarding the company being audited were more predictable and find that the decision-maker should have been able to foresee and avoid the unfavorable events (Lowe & Reckers, 1994). Outcome knowledge that is not legally relevant to the evaluations of decision-makers is shown to influence jurors (Paetzold & Huss, 1993). Jurors seem unable to separate the client’s loss outcome information from evaluations of auditor liability (Paetzold & Huss, 1993).

Even auditors, making attributions, can fall victim to hindsight bias. [Emby, Gelardi, and Lowe \(2002\)](#) had 123 audit partners from Big 5 firms evaluate the work of their peers. Client-loss outcome knowledge was manipulated at three levels: negative, positive, and no outcome knowledge. Results suggest that negative outcome information results in lower probability assessments and lower evaluations of the auditor ([Emby et al., 2002](#)). Even audit partners, who know that they should be evaluating their peers from an ex ante perspective, may fall victim to hindsight bias and its negative outcome effects.

Increasing the magnitude of outcome damage or losses influences judges' evaluations of auditor responsibility in the direction of the magnitude ([Anderson, Lowe, & Reckers, 1993](#)). Given a similar fact pattern, other nonauditors attribute more responsibility for client loss to the auditors than do certified public accountants ([Arrington, Hillison, & Williams, 1983](#)).

In an early example of attribution of responsibility, [Walster \(1966\)](#) finds that more severe outcomes result in greater responsibility assigned to the decision-maker. In Walster's experiment, the decision-maker takes reasonable precautions in parking a vehicle on a hill. Several scenarios are presented, with varying outcomes. In the more serious outcome scenarios, participants attribute greater responsibility to the decision-maker for those severe outcomes ([Walster, 1966](#)).

Prior research also indicates that the severity of the possible outcomes of an illegal police search influences jurors in their awarding of compensatory and punitive damages ([Casper, Benedict, & Perry, 1989](#)). When the outcome of this illegal search results in more damaging evidence of police misconduct, the compensatory and punitive damages are greater ([Casper et al., 1989](#)). Similarly, [Winter and Greene \(2007\)](#) find increasing the severity of outcomes for personal injury cases juror's damage awards. Even when the plaintiff is "less honorable," the known damage outcome may increase the amount awarded to the plaintiff.

In a mock personal injury trial, [Chapman and Bornstein \(1996\)](#) varied the ad damnum. In one condition, the hypothetical plaintiff claimed that her birth control pills led to her ovarian cancer and requested the large sum of \$1 billion in compensation. Although she was perceived as more selfish and less honorable than a hypothetical plaintiff who asked for a more modest \$5 million, still jurors awarded more to her than to the more reasonable plaintiff ([Winter & Greene, 2007, p. 756](#)).

In certain auditor liability cases, no outcome effect is found when outcomes are just severe ([Jennings, Kneer, & Reckers, 1993](#)) or for different conditions that are only moderately severe and under a higher standard of care ([Kadous, 2000](#)). Still, when outcomes are very severe, juror evaluations

increase auditor responsibility even under the higher standard of care (Kadous, 2000). Hence, given the anticipated negative hindsight bias for jurors:

H₂. Jurors will evaluate auditors' performance less favorably when the ex post audit, client-loss outcome knowledge is more negative.

RESEARCH METHOD

Design of the Experiment

Following a design similar to Kadous (2000), this study tests these hypotheses by manipulating the two levels of accounting standard and two levels of client-loss outcome knowledge in a between-subjects 2 × 2 factorial design. Unlike Kadous (2000) who uses 23 jurors and 84 nonjurors, and unlike many legal scholars (e.g., Hastie, 2008) who use students as “mock” jurors, this study uses 124 juror-qualified individuals as participants.

Participants

The venue for the research study is a circuit court in West Virginia. The court selects jurors from a geographical area encompassing the city and the surrounding county. Jurors are randomly selected to serve in the jury pool for a one-month period. Jury pools of 80 to 90 qualified participants are called to orientation on the second Monday of each month. Participants in this study are drawn from the pool of remaining jurors, after the orientation and the day's actual jury, if any, is seated. The remaining jurors, in groups of 25 to 50, are asked to remain in the courtroom, where the experiment is conducted.

To achieve maximum participation, participants were offered nominal remuneration of \$5 for their time to participate in the experiment. One hundred and ninety-two jurors volunteered to participate and completed the experimental instrument. The instrument included three manipulation check questions. Participants were asked first if they understood the case materials and instructions. Five participants answered this question in the negative and were dropped from the analysis. For the outcome knowledge manipulation, participants were asked if they remembered the outcome of their case and were given a choice as follows: (a) the company experienced

some difficulty, but was expected to recover or (b) the company filed for bankruptcy and employees lost their jobs. These responses correspond to the actual case outcome. These manipulations and other checks reduced the number of participants from 192 to a reliable set of 124.

The experiment took place over five months, with groups of 25 to 50 participants on orientation day, the second Monday of each month. This data was then compared with selected information provided to the court from juror qualification reports. As mentioned in [Table 1](#), the average age of qualified jurors is 49.19 years. This compares favorably with the average age of participants (44.16 years) in this study.

The gender breakdown, from the qualification report, indicates that 47.3 percent were male and 52.7 percent were female. This compares somewhat

Table 1. Demographic Information.

Panel A: Continuous Measures					
Variable description	<i>N</i>	Minimum	Maximum	Mean	Standard deviation
Age in years	124	20	68	44.19	11.44
Income per year	114	0	\$250,000	\$39,436	\$34,237
Education in years	124	6	19	14.32	2.42
Panel B: Discrete Measures					
Variable Description	Level	<i>N</i>	Percent		
Gender	Male	40	32		
	Female	84	68		
Marital status	Married	84	68		
	Single	40	32		
Employment	Employed	94	76		
	Unemployed	14	11		
	Retired	16	13		
CPA	Yes	1	1		
	No	123	99		
Auditor	Yes	2	2		
	No	122	98		
Stockholder	Yes	50	40		
	No	74	60		
Attorney	Yes	1	1		
	No	123	99		
Property owner	Yes	92	74		
	No	32	26		

favorably with 37 percent male and 63 percent female for this study. As the juror qualification report represents the results of the total county-wide pool of potential jurors who are subject to jury service for the year, the gender breakdown of any individual juror orientation pool may not reflect the qualification report averages. Gender data was not collected for the orientation pool before asking for volunteers for the study; therefore, no further comparisons were undertaken. On average, 68 percent of the participants were married, whereas 32 percent were single. Seventy-six percent were employed, 11 percent were unemployed, and 13 percent were retired. Additional demographic information regarding whether the participant was an auditor, a CPA, stockholder, attorney, or property owner is presented in [Table 1](#).

Task

Participants were told that the experiment would require approximately 20 minutes. Each jury pool was then read a scripted statement that described the case materials and the task to be performed. Each participant was then given a packet containing an instruction letter, which included the instructions on how to answer the case, the information the judge would normally provide on jury conduct, the case materials, and a questionnaire that provides evaluations and checks on these evaluations.

The Case

The case materials distributed depict a transaction involving leased equipment. The company, PEP, manufactures parts for the automotive industry, and its management believes that leasing rather than buying new equipment, needed for manufacturing, is necessary to return the company to profitability. To maintain favorable financial ratios, the company feels it is important that the leases, valued approximately as a \$10 million obligation, should be structured as operating (not capital) leases, and hence, the \$10 million is not listed as liability on the balance sheet.

During the audit of the financial statements, Smith & Co, the auditors, evaluated the lease transaction and concurred with management's accounting. The auditors then issued a standard audit opinion (a favorable audit report) as they had in all previous years. During the next year, the PEP sales from the lease operations continued to decline, and because the lease obligation expenses remained, earnings dropped. As a result, stock price declined greatly.

The stockholders, who either bought PEP stock or owned stock after the audit report, suffered significant losses with the decline in the stock price. These stockholders then filed a suit against PEP and Smith & Co. alleging that the true financial picture was not consistent with the audited (approved) financial statements.

Independent Variables

Participants received one of the four possible case scenarios, each one from the 2×2 factorial design of the possible combinations. The manipulations are summarized in the [appendix](#). The type of accounting standard is manipulated at two levels: a principles-based accounting standard and a rules-based accounting standard. The severity of the consequence of the client loss is manipulated at two levels: moderately negative and severely negative. In the moderately negative condition, the client is currently experiencing financial distress, and in the severely negative condition, the client has declared bankruptcy.

Type of Standard

Each case scenario involves the question as to whether management correctly accounted for the leasing arrangement by not capitalizing the leased assets. For the rules-based accounting manipulation, the “bright-line” rules of Standard No.13 (FASB, 1976) are provided as the guidelines that the management and the auditors followed in the preparation of the financial statements. Standard No. 13 states that a lease must be capitalized if the lease term is equal to 75 percent or more of the economic life of the leased asset *or* the present value of the minimum lease payments equal to 90 percent or more of the fair market value of the leased asset (equipment). For the principles-based manipulation, capitalization must occur if the lease term includes “most” of the economic life of the asset *or* if the present value of the minimum lease payments equal to “most” of the fair market value of the leased asset (equipment). The word “most” used here corresponds to similar use of the word “most” used in the suggested convergence (as opposed to adoption) of IFRS with GAAP (AICPA, 2011).

For both cases where the accounting standard changed, the juror must decide if the auditor’s decision to concur with the client’s decision to classify the lease as an operating lease was correct. In both cases, management structured the lease agreements such that the lease term equals 70 percent of the economic life of the equipment and the present value of the minimum

lease payments equals 88 percent of the fair market value of the leased equipment. If the participant received the rules-based standard manipulation, and the standard seems to clearly support the auditor's decision, the auditor would not seem to be liable because 70 percent is less than 75 percent and 88 percent is less than 90 percent as set forth in the standard. In the principles-based manipulation, the juror had to decide whether the word "most" corresponded to a percentage that exceeds both the 75 percent and the 88 percent and decide whether the auditor should be held liable. Previous research on how jurors are expected to value the word "most" suggests they use a number greater than 75 percent but not necessarily one that is greater than 88 percent. (See the previous discussion on language and Kaye, 1982).

Type of Outcome

Following the design similar to Kadous (2000), outcome was manipulated using two levels of consequence from the client-loss, moderately negative and severely negative. In the moderately negative outcome, the company suffered large losses and price declines, but the company is expected to recover and continue. In the severely negative outcome, the company suffered large losses that led to bankruptcy and the subsequent loss of jobs.

Dependent Variable and Covariate

In addition to receiving the instruction letter and one of four case scenarios, the participants also received a questionnaire in which they were asked to render a decision regarding the case. (These questions are also shown in the appendix, part C). A dependent variable was constructed from the sum of answers to the first two questions for each participant. Each question is similar to one of two questions asked by Lowe et al. (2002): whether the auditors (1) were correct in agreeing with the client and (2) were competent? Lowe et al. (2002) presents a third question as part of the dependent variable, but in this study, this question score is represented as a covariate, one to measure investor risk-taking responsibility: (3) To what degree were the investors responsible for their own loss? Each response to these questions was properly coded and scored as investor responsibility, from 0 to 9 (similar to Lowe et al., 2002).

The Design Model

The following general linear (analysis of covariance) model¹ permits analysis of the variables of the 2×2 factorial design:

$$E(Y) = C + (P_{pb} \times X_{pb}) + (P_{so} \times X_{so}) + (P_{ir} \times X_{ir}) + (P_i \times X_{pb}X_{so})$$

where,

$E(Y)$ = Expected value of the dependent variable Y , the total score of correctness and competence (min = 0, max = 18)

C = Constant score for rules-based, moderate outcome

P_{pb} = Parameter for principles-based (pb) ($X_{pb} = 1, 0$)

P_{so} = Parameter for severe outcome (so) ($X_{so} = 1, 0$)

P_{ir} = Parameter for investor risk-taking responsibility (ir) ($X_{ir} = 1, 0$)

P_i = Parameter for interaction, i , between pb and so ($X_{pb}X_{so} = 1$, if $X_{pb} = 1$ and $X_{so} = 1$)

RESULTS

Table 2, panel A, presents the cell means for this 2×2 factorial design for each question, the combined means, and the sums of both questions. The statistical null hypotheses (which equate to “not H_1 ” and “not H_2 ”) for the (two) main effects are tested after first testing the interaction effects in Table 3. The interaction of type of standard and outcome is not significant ($p = 0.9095$). Given that this interaction is not significant, the first main effect test to consider is the investor responsibility covariate effect, which is significant $p = 0.0106$. The next test on the effect of the principles-based standard is also significant ($p = 0.0001$), providing support for H_1 . The principles-based standard matters even in the face of the covariate, investor risk-taking responsibility. This is also mentioned in Table 2, panel B, where the reduced means of the evaluations for the standards are compared (12.10 vs. 7.21) and found to be statistically different, ignoring the outcome effect.

Table 3 provides insufficient evidence to support the outcome effect predicted in H_2 ($p = 0.6227$), the claim that increasing the severity of the outcome damages will increase the negative evaluation of the auditor. Although H_2 is not supported, the sign of the differences in the means, 10.09 versus 9.23, is in the positive direction, consistent with expectations, as mentioned in Table 2, panel C.

Table 2. Means and Test of Differences for Dependent Variable.

Panel A: Means of the Dependent Variables: Correctness and Competence			
Type of standard	Dependent variable	Client-loss outcome	
		Moderate	Severe
Rules-based standard	Correctness	3.17	3.88
	Competence	3.50	3.84
	Combined	6.67 (<i>n</i> = 30)	7.72 (<i>n</i> = 32)
Principles-based standard	Correctness	6.22	6.10
	Competence	5.41	6.00
	Combined	11.63 (<i>n</i> = 32)	12.61 (<i>n</i> = 30)

Panel B: Test of Means: Accounting Standard Ignoring Client-Loss Outcome		
Type of standard	Means (unadjusted)	LSD test (5%)
Rules-based standard	7.21	Significant
Principles-based standard	12.10	

Panel C: Tests of Means: Client-Loss Outcomes Ignoring Accounting Standard		
Client-loss outcome	Means (unadjusted)	LSD test (5%)
Severe	10.09	Not significant
Moderate	9.23	

Note: Lower values indicate greater agreement with auditors' decision and higher levels of auditor competence.

Table 3. ANCOVA for Accounting Standard, Outcomes, and Investor Responsibility on Juror Evaluations.

Source	df	SS	MS	<i>F</i>	Significance
Type of standard (H_1)	1	468.94	468.94	17.06	0.0001*
Client-loss outcome (H_2)	1	6.69	6.68	0.24	0.6227
Standard \times client-loss outcome	1	0.36	0.36	0.01	0.9095
Investor responsibility, covariate	1	185.27	185.27	6.74	0.0106*
Error	119	3270.07	27.48		
Total (adjusted)	123	3931.33			

*Significant at 0.05.

Hindsight Bias

Under the rules-based decision, neither the economic life (70 percent <75 percent) nor the value of the lease payments (88 percent <90 percent) require capitalization; hence, there should be no blame to assign to the auditor or the client. Therefore, the average scores on all questions for the rules-based decision should not be above 0.0, but they are above 3.0 for every answer to all the rules-based questions in Table 2, panel A. This demonstrates evidence of (hindsight) bias against the auditor.

Next, this bias continues for the covariate, the perceived investor risk-taking responsibility in Table 4, where both means (2.33 vs. 3.31) are still greater than zero as shown in panel A. In Table 4, panel B, the differences in means (2.84 vs. 4.14) of attribution to the investor is found not statistically significant if the accounting standards change and the outcomes are ignored. In panel C, the difference in means (2.64 and 4.03) of attribution to the

Table 4. Means and Test of Differences for Covariate, Investor Responsibility.

Panel A: Means of the Covariate			
Type of standard		Client-loss outcome	
		Moderate	Severe
Rules-based standard	Mean	2.33	3.31
	(Standard deviation)	(2.64)	(2.55)
	<i>n</i>	30	32
Principles-based standard	Mean	4.03	4.26
	(Standard deviation)	(2.47)	(2.55)
	<i>n</i>	32	30

Panel B: Test of Means: Given Type of Accounting Standard Ignoring Client-Loss Outcome		
Type of standard	Means (unadjusted)	LSD test (5%)
Rules-based standard	4.14	Not significant
Principles-based standard	2.84	

Panel C: Test of Means: Given Client-Loss Outcome Ignoring Type of Accounting Standard		
Client-loss outcome	Means (unadjusted)	LSD test (5%)
Severe	4.03	Significant
Moderate	2.64	

investor is statistically significant for the outcome effect and the accounting standard is ignored. Given these tests and that the means of [Tables 4](#) appear smaller than those of [Table 2](#), it appears that the juror's perception of the responsibility that investors assume when purchasing stock may mitigate the blame to the auditor in the rules versus the principles debate and not the client-loss outcome effect as would be anticipated.

CONCLUSIONS

Findings

The Use of a Principle-Based Standard

This study investigates whether the use of principles-based accounting standard, as compared to a rules-based standard, results in a more negative evaluation of the auditors by jurors. The findings support this investigation and the increased liability-risk concerns, first set forth in the [FASB \(2002\)](#) proposal for a principles-based approach to standard setting and also found in the proposed adoption of the IFRS ([Flynn, 2008](#); [Heffes, 2008](#); [Johnson, 2009](#); [Reason, 2009](#); [Reilly, 2009](#); [WebCPA, 2009](#)). Although this study does not directly address exactly the same language used in the IFRS, it does find more blame assessed to the auditor when the standard is principles-based. Furthermore, this blame may be not appear if the juror thinks the investor assumes the risk when investing in the stock ([Table 4](#)). The client-loss outcome severity effect may not increase the blame to the auditor but may change depending on the perception of investor responsibility. Finally, in every situation under the rules-based scenario, some unwarranted responsibility was assigned to the auditor, no matter what the outcome, demonstrating hindsight bias in professional liability cases continues ([Rachlinski 1998](#)).

Increasing the Severity of the Negative Outcome

These results do not find that the severity of the negative outcome increases the attribution of auditor responsibility. This finding differs from prior research on auditor attribution that finds it dependent on the severity of negative client outcomes ([Kadous, 2000](#)) and studies not involving auditors ([Chapman & Bornstein, 1996](#); [Winter & Greene, 2007](#)).

This lack of finding that increasing outcome severity increases attribution may be explained in many ways. First, this perception may be context-dependent. Job loss is more familiar to jurors in West Virginia, which has

higher than average unemployment. If so, these jurors are more likely to blame the employer-client for job loss than the auditor.

Next, the conceptual distance between levels of outcome information presented in the cases may not have been perceived great enough. In these cases, jurors were not asked to establish damage awards, and no information was provided to give participants a sense of the dollar losses associated with the outcomes. Differences in damages measured in dollars would heighten the perception of differences in the outcomes. Prior studies have increased the impact of the severe outcome effect with still an additional variable, the type of GAAS violation, such as a fraud and type of fraud, as discussed by Kadous (2000). This outcome effect difference may also be changed by the use of reliable decision aids (Lowe et al., 2002), also not considered here.

Yet, this finding is consistent with expected juror behavior that should follow the instructions given by the judge, found in the letter provided to the jurors as part of the orientation: "Your judgment of the auditor's performance should be based upon the information available during the audit, before the lawsuit was filed."

Limitations of the Study

In actual litigation, the jurors would have access to much more information than was provided in this experimental case. The jurors would hear attorneys' opening and closing statements, witness testimony, and the judge's instructions that the event outcomes should not bias judgment. The materials provided to participants in this experiment were presented as factual and pertinent to the case. Had the participants had access to the additional information typically found in discovery and other evidentiary procedures, their decisions might have been different.

Another limitation is that responses were sought from individual participants in what would normally be a group decision-making process. Although first ballots by individuals capture the majority of final jury decisions, it is not known how a case, such as the one described in this research, would ultimately be decided by a jury. However, because participants who did not understand the case were removed because they did not pass the manipulation checks, only those participants who fully understood the case were included in the hypotheses testing. Had those participants who were not familiar with the outcomes been included in an actual jury, they would have become so during group discussions and jury

deliberations. Therefore, the findings may be more conservative than those of jurors not removed but who would have served in an actual jury setting.

Limiting the position of the outcome information may change the case scenarios. Participants were informed of the outcome at the end of the case scenario. This was the last piece of information before completing the questionnaire. In an actual court setting, the jurors might expect to hear about outcomes during voir dire or from judges' and attorneys' opening statements. Unlike Kadous (2000), this study does not ask for a verdict, but instead (similar to Lowe et al., 2002), this study asks about questions of responsibility.

Finally, this research investigated one rules-based accounting standard and made assumptions about the wording of "most" in the more principles-based case. Although research shows a juror's evaluation of "most" could be a percentage greater than 75 percent, the question of whether "most" exceeds 88 percent is open, but the word "most" can be found in the language of the suggested conversion to IFRS/GAAP (AICPA, 2011). If adoption of the IFRS is considered instead, then in place of "most," the language likely to be used is "major part" and "substantially all."

Future Research

Future research could include a similar study of jurors, as this study has the advantage of actual jurors as subjects, as opposed to previous studies that use combinations of some jurors and mock jurors or students (Hastie, 2008; Kadous, 2000; Winter & Greene, 2007). It should be conducted under more controlled conditions for measuring the perception of investor responsibility and the effects of different jurisdictions and use different case materials, those closer to the current research on accounting standards subject to IFRS conversions (Whitehouse, 2011). Participants should receive more training designed to enhance juror cognition, such as court-approved videos for jurors.

NOTE

1. The analysis could have been presented as a multivariate analysis of covariance (MANCOVA), but this analysis (a) would not follow Lowe et al. (2002), (b) is not as informative because the responses to the questions on

correctness and competence were so similar, and (c) would unnecessarily increase the complexity of the analysis.

ACKNOWLEDGMENTS

We appreciate the helpful comments that we received from the Editor, Vicky Arnold, and two reviewers of this journal, from Kathryn Kadous, and those comments made by reviewers for and participants attending the 2008 Midyear Auditing and 2009 national meetings of the American Accounting Association.

REFERENCES

- Accounting Today Staff. (2010). IFRS – Convergence or adoption? *Accounting Today*, September 14. Retrieved from <http://www.webcpa.com/news/IFRS-Convergence-Adoption-55554-1.html>
- AICPA. (2011). The FASB/IASB Lease Accounting Project: Advocacy Standards Guidance. Available at http://www.aicpa.org/InterestAreas/AccountingAndAuditing/Resources/AcctgFinRptg/AcctgFinRptgGuidance/DownloadableDocuments/Leases_Brief.pdf
- Amer, T., Hackenbrack, K., & Nelson, M. (1994). Between-auditor differences in the interpretation of probability phrases. *Auditing: A Journal of Practice & Theory*, 13(1), 126–137.
- Amer, T., Hackenbrack, K., & Nelson, M. (1995). Context-dependence of auditors' interpretations of SFAS No. 5 probability expressions. *Contemporary Accounting Research*, 12(1), 25–39.
- Anderson, J. C., Lowe, D. J., & Reckers, P. M. J. (1993). Evaluation of auditor decisions: Hindsight bias effects and the expectation gap. *Journal of Economic Psychology*, 14, 711–737.
- Arrington, C. E., Hillison, W. A., & Williams, P. F. (1983). The psychology of expectations gaps: Why are there so many disputes about auditor liability? *Accounting and Business Research*, 13(52), 243–250.
- Buchman, T. (1985). An effect of hindsight on predicting bankruptcy with accounting information. *Accounting, Organizations and Society*, 10(3), 267–285.
- Buckless, F. A., & Peace, R. L. (1993). The influence of the source of professional standards on juror decision making. *The Accounting Review*, 68(1), 164–175.
- Casper, J. D., Benedict, K., & Perry, J. L. (1989). Juror decision making, attitudes, and the hindsight bias. *Law and Human Behavior*, 13(3), 291–310.
- Causey, D. Y., Jr., & Causey, S. A. (1991). *Duties and responsibilities of public accountants*. Mississippi State, MS: Accountant's Press.
- Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. *Applied Cognitive Psychology*, 10, 519–540.

- Cloyd, C. B., Frederickson, J. R., & Hill, J. W. (1996). Motivating factors in lawsuits against independent auditors: Experimental evidence on the importance of causality. *Journal of Accounting and Public Policy*, 15, 185–218.
- Emby, C., Gelardi, A. M., & Lowe, D. J. (2002). A research note on the influence of outcome knowledge on audit partners' judgments. *Behavioral Research in Accounting*, 14, 87–103.
- Financial Accounting Standards Board. (1976). *Statement of financial accounting standards no. 13*. Norwalk, CT: FASB.
- Financial Accounting Standards Board. (2002). Proposal for a principles-based approach to U.S. standard setting. Retrieved from <http://www.fasb.org>. Accessed on October 19, 2002.
- Flynn, T. (2008). US warming to IFRS as it moves on from GAAP. Retrieved from <http://www.ft.com/cms/s/0/e944709e-7a19-11dd-bb93-000077b07658.html>. Accessed on September 4, 2008.
- Hastie, R. (2008). What is the story? Explanation and narrative in civil jury decisions. In: B. H. Bornstein, R. L. Wiener, R. Schopp & S. Willborn (Eds.), *Civil juries and civil justice: Psychological and legal perspectives* (pp. 23–43). New York, NY: Springer. Retrieved from <http://springer.com/978-0-387-74488-9>.
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after outcomes are known. *Psychological Bulletin*, 107(3), 311–327.
- Heffes, E. (2008). Global accounting firm CEOs on challenges-transitioning from GAAP to IFRS, and more. *Financial Executive* (May), 14–16.
- International Accounting Standards Committee. (1997). *Leases. International accounting standard 17*. London: IASC.
- Jennings, M., Kneer, D., & Reckers, P. M. J. (1993). The significance of audit decision aids and precase jurists' attitudes on perceptions of audit firm culpability and liability. *Contemporary Accounting Research*, 9(Spring), 489–507.
- Johnson, S. (2009). Simplified reporting: Forgotten in the crisis? *CFO.com*. Retrieved from http://www.cfo.com/article.cfm/14209682/1/c_2984368?f=search. Accessed on August 12.
- Kadous, K. (2000). The effects of audit quality and consequence severity on juror evaluations of auditor responsibility for plaintiff losses. *The Accounting Review*, 75(3), 327–341.
- Kaplow, L. (1999). General characteristics of rules. In: B. Bouckaert & G. DeGeest (Eds.), *Encyclopedia of law and economics* (pp. 502–528). Cheltenham, UK: Edward Elgar.
- Kaye, D. H. (1982). The limits of the preponderance of the evidence standard: Justifiable naked statistical evidence and multiple causation. *American Bar Foundation Research Journal*, 2(Spring), 487–516.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist* (February), 107–128.
- Latham, C. K., & Linville, M. (1998). A review of the literature in audit litigation. *Journal of Accounting Literature*, 17, 175–213.
- Lawton, R., & Parker, D. (2002). Judgments of the rule-related behavior of health care professionals: An experimental study. *British Journal of Health Psychology*, 7, 253–265.
- Leone, M. (2009). Global standards alive and kicking, SEC accounting chief says. *CFO.com*. Retrieved from http://cfo.com/article.cfm/14440802/c_14440953?f=home_today_infinance. Accessed on September 17.
- Lowe, D. J., & Reckers, P. M. J. (1994). The effects of hindsight bias on jurors' evaluations of auditor decisions. *Decision Sciences*, 25(3), 401–426.

- Lowe, D. J., Reckers, P. M. J., & Whitecotton, S. M. (2002). The effects of decision-aid use and reliability on juror's evaluations of auditor liability. *The Accounting Review*, 77(1), 185–202.
- Nelson, M. W. (2003). Behavioral evidence on the effects of principles and rules-based standards. *Accounting Horizons*, 17(1), 91–104.
- Nelson, M. W., Elliott, J. A., & Tarpley, R. L. (2002). Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review*, 77(Suppl.), 175–202.
- Paetzold, R. L., & Huss, H. F. (1993). The standard of care for independent public accountants: Insights for self-regulation and standard setting. *Research in Accounting Regulation*, 7, 61–80.
- Pennington, N., & Hastie, R. (1981). Juror decision-making models: The generalization gap. *Psychological Bulletin*, 89, 246–287.
- Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision-making. *Journal of Personality & Social Psychology*, 51, 242–258.
- Pennington, N., & Hastie, R. (1988). Explanation-based decision-making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 14, 521–533.
- Pennington, N., & Hastie, R. (1993). The story model for juror decision-making. In: R. Hastie (Ed.), *Inside the juror: The psychology of juror decision-making* (pp. 192–221). New York, NY: Cambridge University Press.
- Prophix Software. (2010). Navigating the IFRS conversion process: A white paper prepared by PROPHIX Software. PROPHIX Software (March). Retrieved from <http://www.prophix.com>
- Rachlinski, J. J. (1998). A positive psychological theory of judging in hindsight. *The University of Chicago Law Review*, 65(2), 571–625.
- Reason, T. (2009). IFRS critic to leave accounting firm regulator. CFO.com. Retrieved from <http://www.cfo.com/article.cfm/14361880>. Accessed on September 2.
- Reilly, D. (2009). Grant Thornton CEO, Edward E. Nusbaum, International accounting standards: Going from the talk to doing the walk. *The CPA Journal*, 79(July), 34–39.
- Segovia, J., Arnold, V., & Sutton, S. G. (2009). Do principles vs. rules-based standards have a differential impact on U.S. auditors' decisions? *Advances in Accounting Behavioral Research*, 12, 61–84.
- Tan, H. T., & Lipe, M. G. (1997). Outcome effects: The impact of decision process and outcome controllability. *Journal of Behavioral Decision Making*, 10, 315–325.
- Walster, E. (1966). Assignment of responsibility for an accident. *Journal of Personality and Social Psychology*, 3(1), 73–79.
- WebCPA Staff. (2009). New SEC chair may delay IFRS roadmap. WebCPA.com. Retrieved from <http://www.accountingtoday.com/news/30446-1.html>. Accessed on January 16.
- Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, 71(1), 3–25.
- Whitehouse, T. (2011). Study IDs common themes in IFRS adoption errors. Compliance-week.com. Retrieved from <http://www.complianceweek.com/study-ids-common-themes-in-ifrs-adoption-errors/article/196446/>. Accessed on February 15.
- Winter, R. J., & Greene, E. (2007). Juror decision-making. In: F. Durso, R. Nickerson, S. Dumsid, S. Lewandowsky & T. Perfect (Eds.), *Handbook of applied cognition* (2nd ed., pp. 739–761). New York, NY: Wiley. Retrieved from <http://www.wiley.com/Wiley.CDA/WileyTitle/productCd-0470015349,descCd-tableOfContents.html>

APPENDIX: TYPE OF ACCOUNTING STANDARD AND OUTCOME MANIPULATIONS (ABBREVIATED FROM CASE SCENARIOS)

Part A: Manipulation of Type of Standard

Rules-Based Accounting Standard

Guidelines provided to accountants and auditors, for the proper accounting for leases, are as follows:

If the lease transfers substantially all of the benefits and risks of ownership of the property (equipment), then the company should recognize the obligation (liability) on its balance sheet.

To provide further guidance, the following questions should be answered:

1. Is the lease term equal to 75 percent or more of the economic life of the asset (equipment)?
2. Is the present value of the minimum lease payments equal to 90 percent or more of the fair market value of the leased asset (equipment)?

If the answer to *either* of the above questions is yes, PEP *must* show the \$10 million as a liability on their balance sheet. If the answer to both questions is no, then PEP *does not* show the \$10 million as a liability on their balance sheet.

Management and the leasing company structured the lease agreements such that the lease term was equal to 70 percent of the economic life of the equipment (see question #1 above). They further structured the agreements so that the present value of the minimum lease payments equaled 88 percent of the fair market value of the leased equipment (see question #2 above). Based upon their assessment, PEP classified their leases as operating leases.

Principles-Based Accounting Standard

Guidelines provided to accountants and auditors, for the proper accounting for leases, are as follows:

If the lease transfers substantially all of the benefits and risks of ownership of the property (equipment), then the company should recognize the obligation (liability) on its balance sheet.

To provide further guidance, the following questions should be answered:

1. Does the lease term include most of the economic life of the leased asset (equipment)?

2. Does the present value of the minimum lease payments include most of the fair market value of the leased asset (equipment)?

If the answer to *either* of the above questions is yes, PEP *must* show the \$10 million as a liability on their balance sheet. If the answer to both questions is no, then PEP *does not* show the \$10 million as a liability on their balance sheet.

Management and the leasing company structured the lease agreements such that the lease term was equal to 70 percent of the economic life of the equipment (see question #1 above). They further structured the agreements so that the present value of the minimum lease payments equaled 88 percent of the fair market value of the leased equipment (see question #2 above). Based upon their assessment, PEP classified their leases as operating leases.

Part B: Manipulation of Severity of Client-Loss Outcome

Moderately Negative Outcome Information

Because of the stockholders lawsuit, several creditors called their loans. PEP was unable to pay the creditors. As a result, the creditors sustained large losses. While the present situation looks bleak, PEP is expected to eventually recover.

Severe Negative Outcome Information

Because of the stockholders' lawsuit, several creditors called their loans. PEP could not pay. As a result, PEP declared bankruptcy in late 2003. Stockholders and creditors suffered large losses, and many employees lost their jobs.

Part C: Participant's Decisions

1. Do you feel that the auditors made the correct decision in concluding that the lease should be accounted for as an operating lease? (Circle a number)

Correct decision 0 1 2 3 4 5 6 7 8 9 Incorrect decision

2. How competent did you perceive the auditors to be in performing their duties in the audit of PEP? (Circle a number)

Completely competent 0 1 2 3 4 5 6 7 8 9 Completely incompetent

Participant's Perceptions

1. To what extent do you believe that stockholders must assume normal investment risks when purchasing stock, and therefore are largely responsible for their own losses? (Circle a number)

Responsible 0 1 2 3 4 5 6 7 8 9 Not responsible

WHAT FACTORS INFLUENCE THE NUMBER OF COAUTHORS IN THE PUBLISHED RESEARCH OF THE MOST PRODUCTIVE AUTHORS IN ACCOUNTING LITERATURE? A LONG-TERM STUDY

Robert W. Rutledge, Khondkar E. Karim
and Alan Reinstein

ABSTRACT

This study examines possible influences on the level of collaboration in published research by the most productive authors of accounting literature. Understanding the collaboration tendencies of these authors should benefit early-career-stage accounting faculty. Seven factors are examined for the publications of 93 of the most productive accounting authors. These productive authors are found to include fewer coauthors on their publications early in their careers. The number of coauthors increases through their first 16 to 17 years and then decreases through the remainder of their careers. The results also indicate that productive accounting researchers include a greater number of coauthors on more recently published articles and on longer articles. Fewer coauthors

Advances in Accounting Behavioral Research, Volume 14, 191–225

Copyright © 2011 by Emerald Group Publishing Limited

All rights of reproduction in any form reserved

ISSN: 1475-1488/doi:10.1108/S1475-1488(2011)0000014011

are included when a productive author is affiliated with a “top-10” university or on articles published in highly ranked accounting journals. Lastly, the results show that prolific authors seek out coauthorship throughout their careers and usually include one or more coauthors on their publications. Implications from these results and specific suggestions for accounting faculty are discussed.

INTRODUCTION

Most accounting faculty dedicate much of their time to research – seeking to publish their efforts to advance their careers, to attain promotion, tenure and merit-based salary increases, and to maintain their marketability. Less-experienced accounting faculty should understand the methods that “already successful” accounting researchers use, particularly as they attempt to establish their research techniques and reputations. Successful researchers often coauthor their works to use the synergy of working with other researchers.

Although most research in business is coauthored, a significant portion is not (e.g., Manton & English, 2007). An unanswered question is, “What factors determine the level of coauthoring by accounting faculty on their publications?” Of particular interest is understanding the factors that influence the level of coauthoring by the most productive accounting researchers, who can serve as models for accounting faculty. The current study helps answer this question.

Prior coauthorship research focuses mainly in the area of economics (cf., Barnett, Ault, & Kaserman, 1988; Durden & Perri, 1995; Hudson, 1996; Laband & Wells, 1998; Maske, Durden, & Gaynor, 2003; McDowell & Melvin, 1983; McDowell, Singell, & Stater, 2006; Taylor, Fender, & Burke, 2006). Scant accounting research has examined the function of coauthorship (cf., Beattie & Goodacre, 2004; Hasselback, Reinstein, & Schwan, 2000; Nathan, Hermanson, & Hermanson, 1998; Urbancic, 1992). Urbancic (1992) finds that coauthoring in accounting has increased over time and that most research in the top accounting journals is coauthored. Nathan et al. (1998) survey accounting faculty and administrators’ opinions regarding coauthorship. Hasselback et al. (2000) examine the quantity and the quality of accounting faculty research productivity to determine benchmarks for evaluating accounting faculty. They find the rate of coauthorship in accounting research has significantly increased over the period from 1971 to 1995. Furthermore, Beattie and Goodacre (2004) find that 24% of UK and

Irish accounting and finance researchers' academic articles have three or more authors, with an average of 1.93 authors per article. Lastly, [Rutledge and Karim \(R&K, 2009\)](#) find that productive accounting researchers produce longer articles when they coauthor, and coauthorship is found to be marginally related to the half-life of the journal containing the article and to the number of previous publications by the author.

The current study examines the literature to determine factors that are likely to influence authors' decisions to work with others on published research. These factors are then studied in the context of the most productive authors of accounting literature. Again, these authors are of particular interest because they are the faculty who should be emulated by other faculty who strive to be successful accounting researchers. The results indicate that productive authors with low academic rank (e.g., assistant professors) and those affiliated with highly ranked universities tend to have fewer coauthors. They also have fewer coauthors when publishing in higher-ranked accounting journals. Lastly, productive accounting authors have increasing numbers of coauthors for their more recent articles and for their longer articles.

An unexpected and interesting result is found for the relationship between coauthorship and productive authors' experience (as measured in years since doctoral graduation, academic rank, or total previous publications). The results suggest a life cycle model of coauthorship, whereby productive accounting authors have (1) few or no coauthors early in their academic career, (2) increasing numbers of coauthors through the middle of their careers (e.g., approximately 21 to 50 published articles, or 16 to 17 years of academic experience), and (3) subsequently, fewer coauthors in the late career stages (but not as low as in their early careers).

Lastly, the results suggest that prolific accounting authors appear to continually seek out coauthorship throughout most of their careers. The cultivation of coauthor relationships early in one's career is vitally important. The prolific authors studied are likely to have one or more coauthors on their publications under all of the conditions examined in this study with one exception, that is, prolific authors with more than 80 publications are likely to publish without any coauthors.

The remainder of this study is organized as follows. First, a review of the relevant literature is presented to determine factors that may influence the productive authors' likelihood to coauthor an article. Two testable hypotheses and five research questions (RQs) are developed from this review. Next is a discussion of the research design including a description of the variables, the data collection, and statistical methods used to analyze the data. This is followed by a discussion of the results. Lastly, concluding

remarks are provided, including implications for accounting faculty that want to be successful in their research efforts.

FACTORS THAT ARE LIKELY TO HAVE AN INFLUENCE ON THE LEVEL OF COAUTHORSHIP

Seven variables are suggested from the literature that should have a significant effect on the number of coauthors included in a journal article for productive authors of accounting. These include the (1) productive author's years of experience at the time of publication, (2) rank/quality of the productive author's university, (3) current age of the journal article, (4) rank/quality of the journal publishing the article, (5) academic rank of the productive author at the time of the publication, (6) total number of the productive author's publications at the time of the current publication, and (7) the length of the article.

Experience

Several prior studies examine the effects of experience on coauthorship. McDowell and Melvin (1983) analyze the determinants of coauthorship in the economics literature. They examine data gathered from the top-eight economics journals for the period 1946 through 1976. Their results indicate a significant positive relationship ($p < 0.01$) between author experience (as measured by journal year minus Ph.D. year) and the probability of coauthorship. They suggest that "less experienced" authors tend to work by themselves more than "more experienced" authors. Similarly, McDowell and Smith (1992) find that experience is the most significant variable for explaining the decision to coauthor, with more experience leading to a greater likelihood of coauthoring.

Oh, Choi, and Kim (2006) find that experience is the most significant variable in explaining the academic impact of researchers on other scholars in the area of information systems (IS) ($p < 0.01$). Therefore, experience is strongly tied to their success in assisting other researchers. McDowell et al. (2006) examine the relationship between experience and coauthorship and find that the probability of coauthoring published research increases for economists through the first 19 years of their careers, but no significant relationship exists beyond that time frame. No prior studies examined the relationship between experience and coauthorship in accounting. However, based on the prior research in economics, a greater number of coauthors are

expected to be found on publications of productive accounting authors as their years of experience increase. The following hypothesis is suggested:

- H₁.** The number of coauthors for a given article published by a productive author of accounting literature is positively related to the author's years of experience.

Author at a Top-10 University at the Time of Publication

McDowell and Melvin (1983) rank universities by the number of pages published in the top economics journals over a five-year period. They find university rankings were not significantly related to the probability of coauthorship. Sauer (1988) examines the economic returns associated with coauthorship in academia. He finds that the top-ranked universities will most heavily discount the value of coauthored articles. This suggests that faculty at higher ranked universities will exhibit less tendency to coauthor.

Piette and Ross (1992) propose that economics faculty at highly ranked universities will feel pressure to publish. Because of the pressure to increase publications, they expected to find a greater incidence of coauthorship at the top universities. However, their findings suggest that universities' ranks do not explain the number of coauthors on journal articles (i.e., there is no significant relationship). Subsequently, Sutter and Kocher (2004) find a significant relationship between the rank of authors' institutions and the amount of coauthoring, whereby higher-ranked institutions have less coauthoring. Their results may be different than those of Piette and Ross (1992) because they considered 15 different economics journals, whereas the Piette and Ross study examined only one journal (among other differences).

Lastly, McDowell et al. (2006) report a negative relationship between top-ranked departments and the probability of coauthoring. They suggest two possible explanations: (1) "more able" economists at top departments are less dependent on other authors, and (2) incentives provided at top departments reward sole-authored research publications. The preponderance of prior research suggests that productive accounting authors at highly ranked institutions will have fewer coauthors on their publications compared to authors at lower-ranked institutions. Thus, the following hypothesis is suggested:

- H₂.** The number of coauthors for a given article published by a productive author of accounting literature is negatively related to the rank/quality of the productive author's academic institution.

Age of the Journal Article

Many fields of research have seen an increase in the incidence of coauthorship over time. That is, there are more coauthors included on journal articles now (or more recently) as compared to journal articles published in the past. For example, Glanzel (2002) finds such an increase in the areas of math and science. The premier chemistry journal has had increased coauthorship for each decade in the 20th century, where multiple-authored articles have increased from 44% in the first decade of the 20th century to almost 99% in the 1990s (Cronin, Shaw, & LaBarre, 2004). Laband and Tollison (2000) report that biology coauthorship has increased both in percentage of papers coauthored and in number of coauthors during the period 1950–1994. A similar increase in the percentage of papers coauthored and number of coauthors is found for the top public choice journal during the period from 1973 through 2002 (Sobel & Taylor, 2004). Oh et al. (2006) note that the percentage of articles in the top-four IS journals has increased from 40% in 1978 to over 80% in 2002.

Several studies find increased coauthorship over time in the field of economics (Barnett et al., 1988; Coupe, 2004; Durden & Perri, 1995; Laband & Tollison, 2000; McDowell & Melvin, 1983). Barnett et al. (1988) examine the increasing incidence of coauthorship in the *American Economic Review* for the years 1960 through 1985. They suggest that opportunities for specialization have been a significant cause of the increase in coauthorship. Durden and Perri (1995) suggest that the increase in coauthorship is associated with increased production.

In the area of business research, Manton and English (2007) find a significant increase in the number of coauthors and in percent of articles coauthored for a top journal in each of six different areas, including *The Accounting Review* for the area of accounting ($p < 0.001$). Urbancic (1992) also finds an increase in coauthorship over the last several decades in the top-ranked accounting journals. Hasselback et al. (2000, p. 79) find the rate of coauthorship in accounting research has increased significantly from 1971 to 1995, and this rate is “growing at a pace of 0.017 authors per article per year.” Swanson (2004) adds that the average number of coauthors in the premier four accounting journals, three finance journals, four management journals, and three marketing journals has grown significantly from 1980 through 1999.

On the basis of prior research in the sciences, economics, and business, more recent articles published by productive accounting authors are expected to include more coauthors than older articles. However, because little

theoretical background exists to support framing this relationship as a hypothesis, the following RQ is proposed:

- RQ₁.** Is the number of coauthors for a given article published by a productive author of accounting literature negatively related to the age of the journal article (current year minus journal publication year)?

Journal Rank

Prior research has investigated the relationship between journal quality or rank and the incidence of coauthorship. For example, studies of economics journals (cf., Johnson, 1997; Piette & Ross, 1992) and science journals (cf., Glanzel, 2002) find higher quality articles have fewer coauthors. Assuming that higher rank/quality journals publish higher quality articles, this implies that journal rank is negatively associated with the number of coauthors.

C. L. Brown, Chan, and Lai (2006) report contrary results, finding that the top-four marketing journals have a higher rate of coauthorship than the other top-36 marketing journals for the period 1991–2000. They suggest that “publishing an article in a top-four marketing journal requires more work, making coauthorship more likely” (p. 19). Minimal research is available to predict the direction of the relationship between journal quality and coauthorship. Therefore, because the prior research is inconclusive, the following nondirectional RQ is suggested:

- RQ₂.** Is the number of coauthors for a given article published by a productive author of accounting literature related to the rank/quality of the journal that publishes the article?

Academic Rank of Productive Author

McDowell and Melvin (1983) find that assistant professors of economics are significantly less likely to coauthor compared to their senior colleagues ($p < 0.01$). They also show that “rank” and “experience” are different variables, each with a unique contribution to the likelihood of coauthorship. This may be because many tenured faculty members never go beyond the rank of associate professor, regardless of their years of experience. Next, Piette and Ross (1992) analyze several factors that are likely to influence coauthorship.

They hypothesize that “assistant professors nearing the tenure decision may have an added incentive to co-author papers” (p. 279). However, their data supports the contention that assistant professors are less likely to coauthor compared to higher professor ranks. Similarly, for scientists and engineers, [Bozeman and Corley \(2004\)](#) find greater numbers of collaborators on research as academic rank increases.

Limited research has examined the relationship between academic rank and coauthorship. The research that does exist (as referenced earlier) indicates that economics, science, and engineering faculty are more likely to coauthor as they move up in academic rank. Intuitively, lower-ranked researchers (particularly assistant professors) should try to “make a name for themselves” by producing sole-authored articles. Evaluators of their research will not question their abilities and contributions to research as they might for a multiple-authored article. The value of producing sole-authored articles will likely diminish after an author reaches higher academic ranks and their reputation has been established. Additionally, lower-ranked professors may not have established significant networks of potential coauthors. This would suggest that they would have fewer coauthors as assistant professors compared to when they become associate or full professors. Therefore, higher-ranked productive authors of accounting may have more coauthors included on their publications than lower-ranked authors. The following RQ is suggested based on the discussion earlier:

- RQ₃.** Is the number of coauthors for a given article published by a productive author of accounting literature positively related to the academic rank of the productive author at the time the article is published?

Total Publications at the Time of Current Publication

[Rutledge and Karim \(2009\)](#) find a significant relationship between the number of coauthors and the total articles published by the accounting author at the time of the current publication. That is, as the number of past publications increases for an accounting researcher, the number of coauthors on their articles also increases. Therefore, this is a factor that needs to be considered in the current study.

However, [Rutledge and Karim \(2009\)](#) has a significant set of limitations. They use [Cranfield's \(2003\)](#) half-life impact factors and “type of review” to proxy for journal quality. Most of the productive authors' publications

(63.4%) were not considered for their study because of unavailability of impact factors or type of review. A major improvement over the [Rutledge and Karim \(2009\)](#) study is the current study's use of a journal rank variable to proxy for journal quality. Furthermore, they only examine a short time frame (1993–1997), rather than the current study's 1970–1997 accounting doctoral graduates. Another limitation of the R&K study is that it includes only 187 article observations compared to nearly 3,000 observations in the current study. Lastly, the [Rutledge and Karim \(2009\)](#) study does not control for the age of the journal article, rank of the journal, rank of the author, or rank of the author's university.

The R&K study finds that a researcher's production of articles is positively associated with the level of coauthorship. This suggests that faculty with no (or few) publications may not have developed a network of potential coauthors, and therefore, they work alone, which results in the establishment of their reputation (i.e., a brand name). Their reputation improves as they increase their total prior publications. They are likely to find it easier to attract potential coauthors as their reputation for having the ability to publish becomes established. Thus, it is expected that more coauthors will be included on the current publications of productive authors if they have published more in the past. The following RQ is suggested:

- RQ₄.** Is the number of coauthors for a given article published by a productive author of accounting literature positively related to their total number of publications at the time of the current publication?

Length of Journal Article

[McDowell and Melvin \(1983\)](#) propose that greater effort and value are associated with producing longer articles. Therefore, shorter articles are more likely to be sole-authored and longer articles are more likely to include coauthors. [Petry \(1988\)](#) measures publication success in business journals using "number of authored pages." Of several variables examined, he finds that the number of pages published has the most significant association with percentage of coauthored pages.

[Laband and Tollison \(2000\)](#) study this collaboration relationship in economics. They find a greater occurrence of coauthorship as the number of pages in an article increases. [Dombrow and Turnbull \(2002\)](#) find that page counts of articles by individual real estate researchers are associated with the

extent of coauthorship. Rutledge and Karim (2009) find a similar result for accounting researchers.

The number of pages for an article is significantly related to the quality of the paper as measured by future cites and perceived quality of the paper (Laband & Tollison, 2006; Mason, Steagall, & Fabritius, 1997; Piette & Ross, 1992). Similar to McDowell and Melvin's argument earlier, the higher actual and perceived quality of longer papers involves increased time to produce, which implies a higher likelihood of employing additional coauthors.

In summary, the preceding research implies that the number of coauthors included on an article is likely to be positively related to article length as measured by the number of pages included in the article. However, the cause/effect of this relationship is uncertain. That is, longer articles (e.g., more complicated studies) could cause the number of authors to increase, or the number of authors (i.e., more writers) could cause the article to increase in length. The following RQ is suggested:

RQ₅. Is the number of coauthors for a given article published by a productive author of accounting literature positively related to the total number of pages included in the publication?

RESEARCH DESIGN

Method

To test the hypotheses proposed earlier, the following equation is estimated using least squares regression:

$$\begin{aligned} \text{TNOA}_i = & \alpha + \beta_1 \text{EXP} + \beta_2 \text{TOPUNIV} + \beta_3 \text{AGEJ} + \beta_4 \text{JRANK} \\ & + \beta_5 \text{AURANK} + \beta_6 \text{TPUB} + \beta_7 \text{NOP} + \varepsilon_i \end{aligned} \quad (1)$$

where, TNOA_i is number of authors on article i (productive author plus coauthors); α is intercept; β_x is regression coefficient estimate for variable x (where $x = 1-7$); EXP is productive author experience (publication year of article i minus graduation year of author); TOPUNIV is author at a "top-10 university" at the time of publication of article i (0 = no; 1 = yes); AGEJ is age of journal article i (current year minus publication year of article i); JRANK is rank/quality of journal containing article i (1 = top-4 journals; 2 = journals ranked 5 through 12; 3 = journals ranked 13 through 22; 4 = journals ranked 23 through 40; 5 = non-top-40 journals); AURANK is

author rank at the time of publication of article i (1 = assistant professor; 2 = associate professor; 3 = professor); TPUB is total number of publications by the “productive” author on article i at the time article i was published; NOP is number of pages in article i ; and ε_i is error term.

Data and Variable Descriptions

Identification of the Productive Authors

The list of the most productive authors of accounting literature is derived from [Hasselback, Reinstein, and Schwan \(2003\)](#). The current study employs their listing of the top-10 prolific authors (not including ties) for every third doctoral graduation year starting from 1970 to 1997 (the most recent year listed).¹ The final sample includes 93 subjects. The names of the productive authors included in this study and their doctoral-granting institutions are given in [Exhibit 1](#).

Productive Author Publications

This study is unique in the breadth of journal articles included for the productive authors who are being studied and examined for a long time frame. All publications available for each of the productive authors are identified through an author search of the 2008 ABI/Inform Global database and the 2008 J. L. Heck database. All published journal articles for each productive author from their graduation year through the end of 2006 are included.

Together, the two databases include 2,940 different academic and practitioner journal articles for the prolific authors considered in this study. These articles include 788 sole-authored (26.8%); 1,194 dual-authored (40.6%); 761 tri-authored (25.9%); 155 four-authored (5.3%); and 42 with more than 4 authors (1.4%). In addition to identifying each of the productive authors' publication titles, these two databases are used to identify the (1) name of the journal containing the article; (2) journal year when the article was published; (3) coauthors included on the article; and (4) total number of pages for the article.

Journal Quality

Journal quality is ranked consistent with [Englebrecht, Bisping, Anderson, and Hasselback \(2008\)](#) and [Hasselback et al. \(2003, 2000\)](#).² That is, the top-40 journals are selected based on a composite ranking from five prior journal ranking studies including [Jolly, Schroeder, and Spear \(1995\)](#), [Smith \(1994\)](#),

Exhibit 1. Productive Authors Included in the Study (including Doctoral Graduation Year, Name, and Doctoral Program).

1997

Ashbaugh, Hollis	Iowa
Bushee, Brian	Michigan
Johnstone, Karla	Connecticut
Mahoney, Lois S.	Central Florida
Mauldin, D. Shawn	Mississippi
Nichols, Nancy B.	North Texas
Pacini, Carl J.	Florida State
Seida, Jim A.	Texas A&M
Tinkelman, Daniel	New York

1994

Beasley, Mark S.	Michigan State
Behn, Bruce K.	Arizona State
Hunton, James E.	Texas-Arling.
Hwang, Lee-Seok	New York
Iyer, Govind S.	Georgia State
Vafeas, Nikos	Kansas
Walker, Paul L.	Colorado
Wilkins, Michael	Arizona

1991

Adhikari, Ajay	Virg. Comm.
Amir, Eli	Berkeley
Balsam, Steven	Baruch
Cullinan, Charles	Kentucky
Ghosh, Dipankar	Penn. State
Ramsay, Robert J.	Indiana
Ruhl, Jack M.	Case Western
Young, Joni J.	Illinois
Zimmermann, R.	Texas Tech

1988

Bonner, Sarah E.	Michigan
Geiger, Marshall	Penn. State
Kachelmeier, S.	Florida
Roberts, Michael	Georgia State
Ryan, Stephan G.	Stanford
Sivaramakrishnan, K.	Northwestern
Schadewald, M.	Minnesota
Wheeler, Steven	Arizona State

1985

Anderson, Urton L.	Minnesota
Bedard, Jean C.	Wisconsin
Bline, Dennis M.	Arkansas

Exhibit 1. (Continued)

Datar, Srikant M.	Stanford
Reiter, Sara A.	Missouri
Rezaee, Zabihollah	Mississippi
Shaw, Wayne H.	Texas-Austin
Siegel, Philip	Memphis
Strawser, Jerry R.	Texas A&M
Zarowin, Paul A.	Chicago
<i>1982</i>	
Abdolmohammadi, M.	Indiana
Bernard, Victor L.	Illinois
Borthick, A. Faye	Tennessee
Kaplan, Steven E.	Illinois
Lambert, Richard	Stanford
Limberg, Stephen	Arizona State
Lys, Thomas Z.	Rochester
Schneider, Arnold	Ohio State
Stout, David E.	Pittsburgh
Wilson Earl R.	Missouri
<i>1979</i>	
Baldwin, Bruce	Arizona State
Brownell, Peter	Berkeley
Covaleski, Mark	Penn. State
Giroux, Gary A.	Texas Tech
Mensah, Yaw M.	Illinois
Messier, William	Indiana
Raman, Kris K.	Indiana
Smith, David B.	Illinois
Solomon, Ira	Texas-Austin
Wright, Arnold M.	Southern Cal.
<i>1976</i>	
Bloom, Robert	New York
Dillard, Jessie F.	South Carolina
Englebrecht, Ted	South Carolina
Gibbins, Michael	Cornell
Graham, Lynford	Pennsylvania
Maples, Lawrence	Mississippi St.
Pastena, Victor S.	New York
Patton, James M.	Wash. Univ.
Porcano, Thomas	Indiana
Ro, Byung T.	Michigan State
<i>1973</i>	
Ashton, Robert H.	Minnesota
Boatsman, James	Texas-Austin

Exhibit 1. (Continued)

Coffman, Edward	George Wash.
Collins, Daniel W.	Iowa
Epstein, Marc J.	Oregon
Gordon, Lawrence	Rensselaer
Imhoff, Eugene A.	Michigan State
Schnee, Edward J.	Michigan State
Sunder, Shyam	Carn. Mellon
<i>1970</i>	
Felix, William L.	Ohio State
Gibson, Charles	Kent State
Loeb, Stephen E.	Wisconsin
Most, Kenneth S.	Florida
Nichols, Donald R.	Oklahoma
Robertson, Jack C.	North Carolina
Ronen, Joshua	Stanford
Stickney, Clyde P.	Florida State
Seago, W. Eugene	Georgia
Williams, Jan R.	Arkansas

Source: Derived from Hasselback et al. (2003)

Hall and Ross (1991), Hull and Wright (1990), and Schroeder, Payne, and Harris (1988). These composites are sensitive to major, but not minor, differences in perceptions of quality between individual journals (Hasselback et al., 2000). Therefore, the Morris, Cudd, and Crain (1990) methodology is used to reduce the 40 journal rankings down to fewer journal quality categories.

Thus, the top-40 journals are assigned to 4 journal quality categories: category 1 – top-4 journals; category 2 – journals ranked 5 through 12; category 3 – journals ranked 13 through 22; and category 4 – journals ranked 23 through 40. All productive author publications that are not published in the top-40 journals are assigned to category 5. Exhibit 2 provides the five category descriptions and lists the journal titles included in the top four categories.

Top Accounting Programs

The literature includes many studies that seek to rank top accounting programs, using different criteria as bases for program rankings. Some early studies use subjective opinion surveys to rank the quality of accounting programs. Estes (1970) surveys faculty, administrators, and nonacademics to rank accounting programs. Similarly, Carpenter, Crumbley, and Strawser

(1974) survey faculty members to rank such programs. Both the Estes (1970) and the Carpenter et al. (1974) studies allow respondents to determine their own subjective criteria to assess accounting department quality.

Citation analysis is also used to rank the top accounting programs. For example, L. D. Brown and Gardner (1985) look at the references in the top-four accounting journals from 1976 to 1982. They determine the number of times articles written by an accounting program's faculty are cited (i.e., referenced in the top-four journals' articles). The total cites of each program's researchers determine the rank of the program. L. D. Brown and Laksmana (2004) choose a different measure to rank the top accounting programs. They count Social Science Research Network (SSRN) downloads to rank accounting programs.

Several studies use research production to rank accounting programs. Bazley and Nikolai (1975) rank accounting departments based on the number of published journal articles by the departments in four accounting journals. They rank universities by the authors' location at the time articles are published, the authors' location during the 1974–1975 academic year, and the university where the authors earned their doctorates. Jacobs, Hartgraves, and Beard (1986) rank the top-25 universities based on the publication productivity of their Ph.D. graduates in 8 journals over a 13-year period and develop time- and size-adjusted productivity measures. They find rankings change over time. Trieschmann, Dennis, Northcraft, and Niemi (2000) use page counts of articles in the top-three accounting research journals to determine research production and to rank the top accounting departments. Chan, Chen, and Cheng (2007) provide a global ranking of accounting programs based on “number of articles published” by their faculty in the top-5 and top-24 accounting journals during the period 1991–2005.

Stammerjohan and Hall (2002) provide four rankings of accounting departments based on the institution in which their graduates are initially placed. The first ranking of accounting departments is based on initial placement of graduates at *US News and World Report's* 1997 best colleges (e.g., the top-ranked accounting department is the one that placed its graduates at the highest ranked *US News and World Report* colleges). The other three Stammerjohan and Hall (2002) rankings were based on accounting research productivity of initial-placement institutions including (1) the overall productivity of initial-placement institution, (2) the productivity of “AACSB accredited” initial-placement institutions, and (3) the productivity of “doctoral granting” initial-placement institutions.

The current study defines top accounting programs as those ranked in the top-10 of any of Stammerjohan and Hall's (2002) 4 rankings. This results in

Exhibit 2. Quality Ranking of Journals Included in the Study.

Category 1 – Top-4 Journals

Journal of Accounting Research
The Accounting Review
Journal of Accounting and Economics
Journal of Finance

Category 2 – Ranks 5 through 12

Accounting, Organizations and Society
Contemporary Accounting Research
Journal of Accounting, Auditing and Finance
Journal of the American Taxation Association
Journal of Business
Journal of Finance and Quantitative Analysis
Journal of Financial Economics
Management Science

Category 3 – Ranks 13 through 22

Auditing: A Journal of Practice & Theory
Journal of Accounting and Public Policy
Journal of Business, Finance and Accounting
Journal of Management Accounting Research
Journal of Taxation
National Tax Journal
Abacus
Accounting and Business Research
Behavioral Research in Accounting
Journal of Accounting Literature

Category 4 – Ranks 23 through 40

Accounting, Auditing & Accountability
Accounting Horizons
Financial Analysts Journal
Issues in Accounting Education
Journal of Accountancy
Advances in Accounting
International Journal of Accounting Education and Research
Journal of Accounting Education
Advances in International Accounting
Advances in Taxation
Critical Perspectives on Accounting
The Journal of Information Systems
Research in Accounting Regulation
Research in Governmental and Nonprofit Accounting
Accounting Educators' Journal
Accounting and Finance
The CPA Journal
Management Accounting

Category 5 – All other journals listed in the ABI/Inform Database or Heck Data Base

a group of accounting institutions that are both highly productive in research and capture general academic environmental issues.³ The 13 accounting programs included in the study are summarized in [Exhibit 3](#).

There are significant similarities between the ranking used in the current study and those from the review of the literature. For example, 8 of this study's top accounting institutions are included in [Chan et al.'s \(2007\)](#) and [L. D. Brown and Gardner's \(1985\)](#) top-11 institutions (based on article counts); 10 of this study's top institutions are included in the [Trieschmann et al.'s \(2000\)](#) top-12 (based on page counts); and 9 institutions from this study's list are included in [L. D. Brown and Laksmans's \(2004\)](#) top-10 ranked accounting programs (based on SSRN downloads). Therefore, much agreement exists among the rankings of accounting programs using differing criteria (i.e., "Top" universities this study considers are essentially the same as other "top university" rankings, regardless of criteria used to determine such rankings).⁴

Other Variables This Study Includes

The year of graduation for the productive authors is determined by reference to [Hasselback et al. \(2003\)](#) and confirmed by reference to [Hasselback's directory \(Hasselback, 1974–1996, 1997–2003, 2004–2007\)](#). Author rank and location at the time of publication are determined by reference to the [Hasselback directory](#) for the year of the publication. Lastly,

Exhibit 3. Top-10 Accounting Programs^a (in Alphabetical Order).

Univ. of California-Berkeley
 Carnegie Mellon University
 University of Chicago
 Columbia University
 Cornell University
 Harvard University
 Massachusetts Institute of Technology
 University of Michigan
 Northwestern University
 University of Pennsylvania
 University of Rochester
 Stanford University
 University of Washington

^aTop-10 in one or more of four ranking measures ([Stammerjohan & Hall, 2002](#)).

the quantity of total publications at the time of the current publication is determined by totaling the number of the given author's prior listed journal articles.

DATA ANALYSIS AND RESULTS

Descriptive Statistics

Descriptive statistics for the data in this study are summarized in [Table 1](#), which notes the “median” productive author in the study is an associate professor with 11 years of experience. This mean (median) number of publications by a productive accounting author is 28 (20). The mean (median) number of authors per publication (including the productive

Table 1. Descriptive Statistics.

Variable	Median	Mean	Standard Deviation	Minimum	Maximum
Number of authors on article (TNOA)	2	2.16	1.01	1	9
Productive author experience (publication year of article minus graduation year of author (EXP))	11	12.84	8.78	0	36
Author at a “top-10 university” at the time of publication (TOPUNIV) [0 = no; 1 = yes]	0	0.08	0.28	0	1
Age of journal article (AGEJ) [current year minus publication year of article]	11	12.52	8.52	0	37
Rank/quality of journal (JRANK) [1 = top-4 journals; 2 = journals ranked 5 through 12; 3 = journals ranked 13 through 22; 4 = journals ranked 23 through 40; 5 = non-top- 40 journals]	4	3.64	1.40	1	5
Productive author rank at the time of publication (AURANK) [1 = assistant professor; 2 = associate professor; 3 = professor]	2	2.26	0.80	1	3
Total number of publications by the “productive” author on article at time article was published (TPUB)	20	28.89	29.00	0	175
Number of pages in article (NOP)	12	14.01	10.21	1	100

author) is 2.16 (2), and the average (mean) number of pages per publication is 14.01 (12). The mean as well as the median publication quality is a top “23 through 40” journal.

Tests for Multicollinearity

Variance inflation factors (VIFs) are calculated for each independent variable in the regression model. The VIFs are included in the ordinary least squares (OLS) results provided in Table 2. A VIF measures the degree of multicollinearity of an independent variable with the other independent

Table 2. OLS Regression Results.

$$TNOA_i = \alpha + \beta_1 EXP + \beta_2 TOPUNIV + \beta_3 AGEJ + \beta_4 JRANK + \beta_5 AURANK + \beta_6 TPUB + \beta_7 NOP + \varepsilon_i$$

Independent Variable	Coefficient	t-Value	p-Value	VIF
EXP	-0.212***	-5.01	0.0001	4.49
TOPUNIV	-0.115*	-1.75	0.0803	1.08
AGEJ	-0.035***	-12.24	0.0001	1.90
JRANK	0.037***	2.65	0.0082	1.22
AURANK	0.260***	7.18	0.0001	2.73
TPUB	-0.007***	-8.09	0.0001	2.09
NOP	0.015***	7.77	0.0001	1.28

Notes: $R^2=0.113$; α , intercept; AGEJ, age of journal article i (current year minus publication year of article i); AURANK, productive author rank at the time of publication of article i (1 = assistant professor; 2 = associate professor; 3 = professor); β_x , regression coefficient estimate for variable x (where $x=1-7$); EXP, productive author experience (publication year of article i minus graduation year of author); NOP, number of pages in article i ; ε_i , error term; JRANK, rank/quality of journal containing article i (1 = top-4 journals; 2 = journals ranked 5 through 12; 3 = journals ranked 13 through 22; 4 = journals ranked 23 through 40; 5 = non-top-40 journals); TNOA _{i} , number of authors on article i (productive author plus coauthors); TOPUNIV, author at a “top-10 university” at the time of publication of article i (0 = no; 1 = yes); TPUB, total number of publications by the “productive” author on article i at the time article i was published; VIF, variance inflation factor (less than 10 indicates inconsequential collinearity).

*Significant at the .10 level.

***Significant at the .01 level.

variables in the model (O'Brien, 2007). VIFs less than 10 are generally considered indicative of inconsequential collinearity (Hair, Anderson, Tatham, & Black, 1995; StataCorp, 1997; Stine, 1995). None of the VIFs for the independent variables included in this study exceed 4.5; therefore, the variables considered in the current study do not have a significant multicollinearity problem.

Regression Analysis

The results of the regression analysis are reported in Table 2. The coefficients for the variables considered in the study are all significant at $p = 0.0803$ or less. The results related to each hypothesis or RQ are presented later.

Experience (H_1)

H_1 predicts that productive authors include more coauthors on their published research as they gain more experience. Table 2 presents the results of the least squares regression that was performed. These results show that a significant, negative relationship ($p < 0.0001$) exists between the total number of authors on a given article and the productive author's experience at the time of publication. This result is contrary to what was expected.

Additional analysis is performed to gain further insight into the relationship between the productive authors' experience and the number of coauthors included on their published research. Interestingly, an increase in the number of coauthors included on published research is found only up to a certain level of experience. That is, the relationship is positive (more authors are included as experience increases), but only through the first 16 or 17 years. Subsequently, productive authors' levels of coauthorship remain relatively consistent through about 24 years and then declines significantly through the remainder of their careers (Fig. 1). The descriptive figure provides evidence of our expectation of a positive relationship between number of coauthors and prolific author experience, but only for about the first 16 or 17 years. An "inverted-U" describes the relationship over their entire career. Our results may explain why some prior studies that only consider shorter durations of experience find this relationship to be positive. In particular, the McDowell et al. (2006) study only examines 19 years of experience. We do not find a significant decline in coauthorship until about 25 years experience.

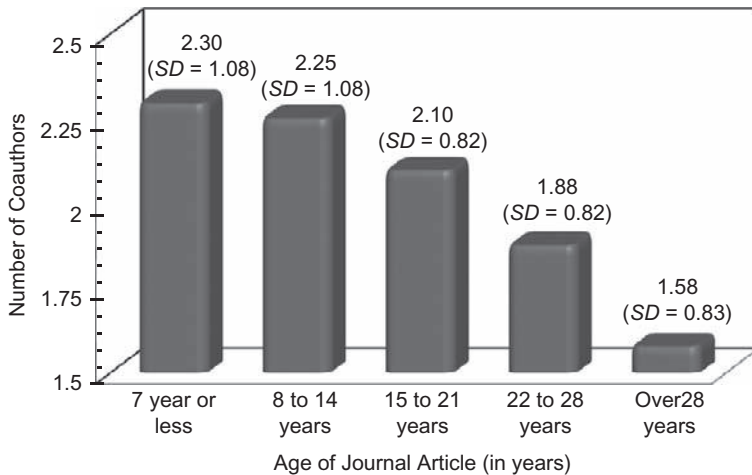


Fig. 1. Distribution of the Number of Authors by Productive Author Experience in Years (i.e., Publication Year of Article i Minus Graduation Year of Author). *Notes:* The one-sample t -test shows that the means are significant at 0.01 ($p = 0.000$, two-tailed). The one-way ANOVA results to test hypotheses regarding equality of means show a statistically significant difference in the average number of authors across authors' experience in years. The Scheffe test was used to see if pairs of means are different and finds that all pairs of means are statistically different ($p < 0.05$) except between 2.31 and 2.30 or 2.04 and 1.92.

Productive Author at a Top-10 University (H_2)

The number of coauthors for a given article published by a productive author of accounting literature is predicted to be negatively related to the rank/quality of the productive author's academic institution (H_2). That is, productive authors affiliated with highly ranked institutions should have fewer coauthors on their publications. Table 2 indicates H_2 is partially supported. A closer look at the means shows that productive authors at top-10 institutions averaged 2.1 authors per article, whereas authors at non-top-10 institutions averaged 2.2 authors per article, which is a marginally significant difference ($p < 0.0803$). The data also indicates that prolific authors at non-top-10 universities published 2,691 articles, whereas those at top-10 programs published 249 articles (however, there were considerably more prolific authors at non-top-10 universities).

Age of Journal Article (RQ₁)

RQ₁ asks whether the number of coauthors for a given article published by a productive author is negatively related to the age of the journal article (current year minus journal publication year)? That is, do the most productive accounting researchers include more coauthors on their journal articles now (i.e., more recently) than they did in the past? Table 2 indicates that a significant, negative relationship ($p < 0.0001$) exists between the total number of authors on a given article and the age of the journal article. The relationship between “age of journal article” and “number of coauthors” is the strongest of all independent variables included in this study. This relationship is visually illustrated in Fig. 2.

Journal Rank/Quality (RQ₂)

RQ₂ examines whether the number of coauthors for a given article published by a productive author of accounting literature is related to the rank/quality of the journal that publishes the article? The RQ is not directional because the prior research is inconclusive. The results from

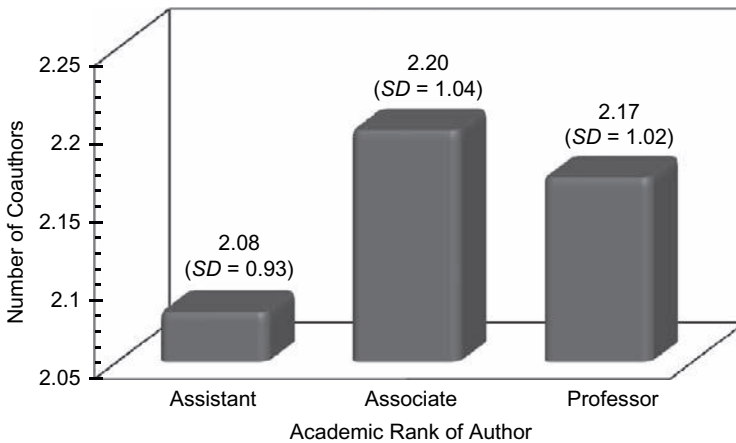


Fig. 2. Distribution of the Number of Coauthors by Age of Journal Article in Years (i.e., Current Year Minus Journal Year). Notes: The one-sample *t*-test shows that the means are significant at 0.01 ($p = 0.000$, two-tailed). The one-way ANOVA results to test hypotheses regarding equality means show a statistically significant difference in the average number of authors across age of journal articles in years ($p = 0.000$). The Scheffe test was used to see if pairs of means are different and finds that all pairs of means are statistically different ($p < 0.05$) except between 2.30 and 2.25.

Table 2 indicate a significant, positive relationship ($p < 0.0082$). Thus, for this study, productive authors' publications in higher ranked journals contain fewer coauthors.

Productive Author Rank (RQ₃)

RQ₃ asks whether the number of coauthors for a given article published by a productive author of accounting literature is positively related to the academic rank of the productive author at the time the article is published? Table 2 notes that the data supports this relationship (i.e., higher academic rank is associated with a greater number of coauthors on publications). The coefficient for the productive author's rank (AURANK) is significant and positive ($p < 0.0001$). Closer examination of the data reveals that the number of coauthors on productive author publications significantly increases ($p < 0.05$) as they move from assistant to associate professor (Fig. 3).

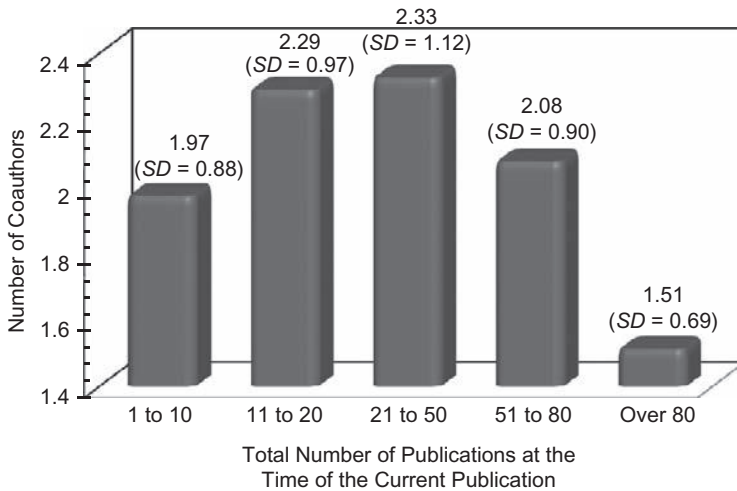


Fig. 3. Distribution of the Number of Coauthors by Academic Rank of Prolific Author. Notes: The one-sample *t*-test shows that the means are significant at 0.01 ($p = 0.000$, two-tailed). The one-way ANOVA results to test hypotheses regarding equality means show that there is a statistically significant difference in the average number of authors across academic ranks of author at 0.01 ($p = 0.000$). The Scheffe test was used to see if pairs of means are different and finds that the difference between assistant professor (2.08) and associate professor (2.20) is statistically different ($p < 0.05$).

However, there is no significant change in the number of coauthors on productive author publications when they move from associate professor to full professor.

Total Number of Publications (RQ₄)

RQ₄ asks whether the number of coauthors for a given article published by a productive author of accounting literature is positively related to their total publications at the time of their current publication? In other words, do productive accounting authors include more coauthors as they publish more articles? Table 2 indicates that this relationship is significant ($p < 0.0001$), but the coefficient is negative – suggesting a decrease in number of coauthors as the total number of previous publications increases.

This unexpected result is investigated further (Fig. 4). The number of coauthors on published research increases, but only through the first “21 to 50” articles. Then, the number of coauthors decreases substantially beyond the “21 to 50” level of total publications. Publications by productive authors with over 80 prior publications average only 1.51 total authors. The graph in

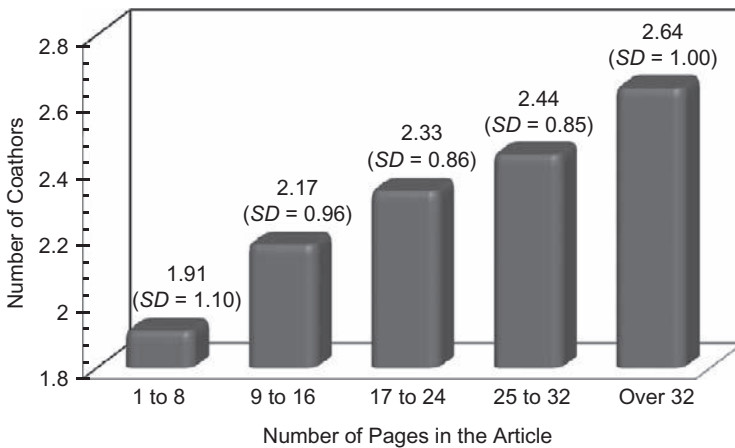


Fig. 4. Distribution of the Number of Coauthors by Total Number of Publications at the Time of the Current Publication. Notes: The one-sample t -test shows that the means are significant at 0.01 ($p = 0.000$, two-tailed). The one-way ANOVA results to test hypotheses regarding equality means show that there is a statistically significant difference in the average number of authors across age of journal articles in years ($p = 0.000$). The Scheffe test was used to see if pairs of means are different and finds that all pairs of means are statistically different ($p < 0.05$) except between 1.97 and 2.08 or 2.29 and 2.33.

Fig. 4 shows that some successful researchers reduce their coauthoring when they achieve a large number of publications. Only prolific authors with over 80 publications are more likely than not to publish without coauthors. In summary, Fig. 4 suggests that the number of coauthors on published research tends to increase for the first 21 to 50 articles and then decreases for the rest of prolific authors' careers.

Number of Pages (RQ₅)

Whether the number of coauthors for a given article published by a productive author of accounting literature is positively related to the total number of pages included in the publication is addressed in RQ₅? The estimated coefficient for NOP (number of pages) is significantly positive ($p < 0.0001$) – see Table 2. This provides strong evidence that longer papers by productive authors have more coauthors. Thus, the additional work required to produce longer articles is likely to be shared among more authors. The relationship between number of authors and article length is illustrated in Fig. 5. In summary, Table 2 and Fig. 5 confirm a positive

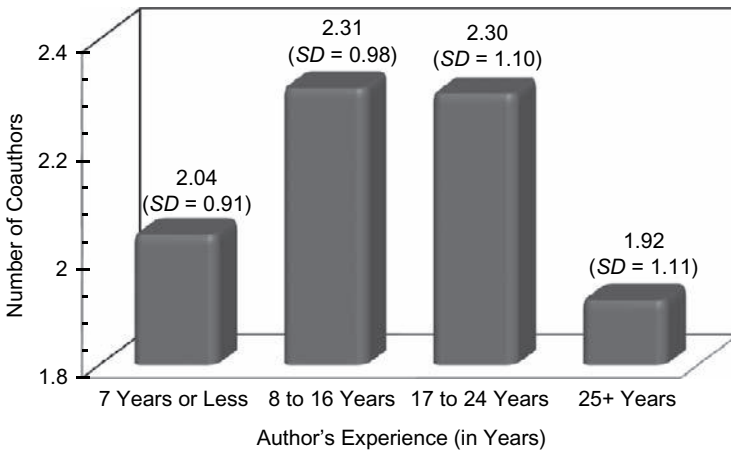


Fig. 5. Distribution of the Number of Coauthors on a Published Article by Number of Pages in the Article. Notes: The one-sample *t*-test shows that the means are significant at 0.01 ($p = 0.000$, two-tailed). The one-way ANOVA results to test hypotheses regarding equality means show that there is a statistically significant difference in the average number of authors across age of journal articles in years ($p = 0.000$). The Scheffe test was used to see if pairs of means are different and finds that all pairs of means are statistically different ($p < 0.05$) except between 2.33 and 2.44 or 2.44 and 2.64.

relationship between the number of coauthors and the number of pages included in prolific authors' publications.

Multinomial LOGIT Analysis

In this section, we consider how the OLS results (above) are affected if a nonparametric logistic regression (LOGIT) approach is used. The LOGIT model is used to predict the probability of dichotomous outcome in terms of the dependent variable, total number of authors (TNOA), as compared to the OLS regression, which measures association of the independent variables with the dependent variable. The LOGIT results are provided in Table 3.

Table 3 notes the model chi-square (χ^2) value is 641.06, which is significant at $p < 0.0001$. The null hypothesis that there was no difference between the model without independent variables and the model with independent variables is rejected. In other words, the existence of a relationship between the independent variables and the dependent variable is supported. Furthermore, the LOGIT results provided in Table 3 indicate that there is a statistically significant relationship between all the independent variables and the dependent variables at $p < 0.05$, with all but one independent variable (i.e., RANK) being significant at $p < 0.0001$. Therefore, each independent variable considered in this study has a significant contribution to the model.

The LOGIT results are more robust than those of OLS because many of the OLS assumptions are relaxed. For example, the independent variables are not required to be normally distributed or to have equal variance in each group, and the model does not assume a linear relationship between the independent and the dependent variables. Furthermore, there is no homogeneity of variance assumption. A comparison of the LOGIT results with the OLS in our primary analysis suggests our results and conclusions would be essentially the same using either method of analysis.

Additional Analysis – Tests of Robustness

To test the robustness of our empirical model, the generalized additive models (GAM) procedure is used to nonparametrically explore the data and to visualize the relationship between the dependent variable and the independent variables. No notable differences exist between the regression results and those of the GAM procedure. Next, the generalized linear

Table 3. LOGIT Regression Results.

$$\text{Ln}\left(\frac{pn}{p1}\right)_i = \gamma_0 + \gamma_1 \text{EXP} + \gamma_2 \text{TOPUNIV} + \gamma_3 \text{AGEJ} + \gamma_4 \text{JRANK} + \gamma_5 \text{AURANK} + \gamma_6 \text{TPUB} + \gamma_7 \text{NOP} + v_i$$

Independent Variable	Coefficient	χ^2 Value	p-Value
EXP	-0.0069***	-80.20	0.0001
TOPUNIV	-0.0070***	-30.27	0.0001
AGEJ	-0.0071***	-182.98	0.0001
JRANK	0.0069**	14.97	0.0364
AURANK	0.0070***	96.84	0.0001
TPUB	-0.0070***	-65.80	0.0001
NOP	0.0071***	168.62	0.0001

Notes: Model $\chi^2 = 641.06$; pseudo $R^2 = 0.212$; AGEJ, age of journal article i (current year minus publication year of article i); AURANK, productive author rank at the time of publication of article i (1 = assistant professor; 2 = associate professor; 3 = professor); EXP, productive author experience (publication year of article i minus graduation year of author); JRANK, rank/quality of journal containing article i (1 = top-4 journals; 2 = journals ranked 5 through 12; 3 = journals ranked 13 through 22; 4 = journals ranked 23 through 40; 5 = non-top-40 journals); $\text{Ln}(pn/p1)_i$, multinomial LOGIT function for TNOA_i (number of authors) on article i (productive author plus coauthors) of group n keeping TNOA_i of group 1 as a base; NOP, number of pages in article i ; TPUB, total number of publications by the “productive” author on article i at the time article i was published; γ_0 , Intercept; γ_x , regression coefficient estimate for variable x (where $x=1-7$); TOPUNIV, author at a “top-10 university” at the time of publication of article i (0 = no; 1 = yes); v_i , error term.

**Significant at the 0.05 level.

***Significant at the 0.01 level.

models (GENMOD) is used, which emphasizes estimation and inference for the parameters of the model. Therefore, the GENMOD procedure is used to fit and assess the corresponding parametric model after having used the GAM procedure to discover an appropriate form of dependence of TNOA (total number of authors) on each of the independent variables.

An iterative procedure is followed whereby each of the independent variables is raised to a higher order to see if it is possible to “better fit”

the data. After a large number of iterations, the best model fit from the GENMOD procedures has only small changes (some higher, some lower) in the level of significance for the independent variables with one exception: TPUB (total number of publications) fit better when raised to the second order. This procedure is merely a test of robustness and addresses the empirical issue of a “better fit.” It does not invalidate the relationship between the independent variables and the dependent variables. Thus, the validity of using OLS in this study is not likely to be unduly affected by departures from model assumptions. The original evidence still holds that there is a significant relationship between the independent variables examined in this study and the total number of authors on an article.

SUMMARY AND CONCLUDING COMMENTS

Discussion of Research Findings

This study develops a model to determine coauthorship behavior. Such behavior is examined for 93 of the most productive authors of accounting literature. The study considers an extensive collection of 2,940 publications over a 37-year period. Also considered is a broad set of seven factors that could affect productive accounting researchers’ publishing patterns. These seven factors were considered in the development and testing of two hypotheses and five RQs.

The results of the study indicate that the most prolific researchers coauthor throughout their entire careers. Specifically, over 73% of the publications by prolific authors are coauthored. However, different patterns of coauthorship are found during their careers based on specific factors. Productive authors of accounting literature appear to have a life cycle element to coauthoring that may be characterized as an “inverted U” or concave function. The stages in productive authors’ careers can be measured in terms of years of experience (H_1), academic rank (RQ_3), or number of prior publications (RQ_4). Our tests for multicollinearity indicate that each of these variables makes its own contribution to this effect. Early in their careers, prolific accounting faculty members have significantly fewer coauthors on their publications. This may be to establish their reputation and to aid in attaining tenure and promotion. Alternatively, this may be a result of not having developed a network of potential coauthors.

As they become more experienced, prolific authors have an opportunity to meet other accounting faculty and develop collaboration networks,

which leads to more potential research coauthors. Also, with an established reputation, they have more associates who want to work with them. This appears to lead to an increase in the number of coauthors during mid-career.

Late in productive authors' careers (i.e., full professors with many publications and years of experience), there is a tendency to stabilize or reduce the number of coauthors on their publications. This may be because knowledge gained through experience reduces their need to coauthor. As elite researchers can more selectively choose their research projects and coauthors, they may behave differently than other researchers and reduce the number of coauthors as their careers progress. Productive authors most always include one or more coauthors on their published research. However, they appear to increase and then decrease the number of coauthors on published research as they progress through their careers.

In addition to the life cycle model discussed earlier, the results suggest that productive accounting researchers increase the number of coauthors on published research over time (RQ₁). Articles published over 28 years ago include fewer than 1.6 authors each, whereas articles published since 2000 include approximately 2.3 authors each. The current study considers a broader set of journals and a longer time frame but is consistent with the findings of prior business research on this result (e.g., [Hasselback et al., 2000](#); [Manton & English, 2007](#)).

A slight tendency exists for the number of coauthors to decrease when productive accounting authors are affiliated with top-10 universities (H₂) and when they publish in higher-ranked accounting journals (RQ₂). This suggests that a different incentive structure may be present in "heavy research" universities as compared to other universities. These incentives influence productive accounting faculty at top universities to minimize their number of coauthors and to concentrate their efforts on producing articles in the top journals. Researchers should recognize that some of these high-level programs have unwritten requirements that at least one sole-authored article is necessary to achieve tenure, an effect of which future studies may examine. Furthermore, accounting faculty should recognize that [Hasselback et al.'s \(2003\)](#) and other rankings show accounting faculty give higher rankings to "primarily archival" journals as compared to behavioral journals.

Productive authors of accounting literature also include a greater number of coauthors when they work on longer articles (RQ₅). This result is interesting in that it points out potential advantages to productive authors working with others when writing longer articles. Working with others is

likely to provide the productive author a greater variety of expertise and more available “man-hours” to produce such longer articles.

Suggestions for Future Research

The authors considered including a dummy variable for practitioner versus basic academic research, suspecting that it could capture part of the effect of the number of pages (NOP) variable. However, the type of research is not a dichotomous variable, as some of the journals that are included have both practitioner- and academic-focused articles, and many research papers include both academic and practitioner components. Therefore, this variable is not included in this study. Future research may attempt to look closer at this potential variable.

Additional variables that may aid in explaining the number of coauthors on prolific authors' published research include the following: (a) whether the author teaches at a doctoral-granting accounting program; (b) whether the researcher coauthors works with his or her doctoral students; (c) whether the researcher specializes in archival or behavioral research (although many papers focus on both or neither of these areas); (d) whether the faculty member specializes in financial, managerial, auditing, tax, systems or other specialty areas; and (e) whether the employers provide spring/summer research relief, grants, or other support. Future researchers may include some or all of the above variables in explaining the number of coauthors or probability of coauthoring.

Another topic for researchers may be to consider the coauthoring activity of accounting researchers that are not considered prolific and attempt to determine if there is a difference in the behavior between prolific and non-prolific authors. Future researchers may also want to compare the coauthoring activities of accounting researchers with non-accountants.

Implications

Important implications can be derived from the current study. As we only examine the research records of elite accounting researchers, we cannot extend the implications to less-productive colleagues. The results of this study are particularly important to “early stage” accounting faculty who should understand the most successful accounting researchers' publication habits if they want to emulate their behavior. For most faculty members,

successful publishing is a key to achieving favorable tenure and promotion decisions and for obtaining salary increases, recognition, and other rewards.

To follow the patterns of successful accounting researchers, faculty should seek out coauthors throughout their careers. A critically important factor for new faculty is to develop coauthor relationships early in their careers, as prolific authors generally have one or more coauthors working with them on their published research. All faculty should minimize, but not necessarily eliminate, coauthoring if they are affiliated with top universities or publish in top-ranked journals. Coauthorship will generally increase in the middle stages of prolific authors' careers and decrease in the late career stages. Lastly, faculty working on longer (more complex) articles may want to take advantage of shared work and the varied expertise that comes from working with others.

The results should encourage accounting administrators to reward their productive senior faculty for mentoring their talented newer colleagues, thereby broadening the departments' research productivity, and producing an increased quantity of high-level articles. This process may also reduce the costs of replacing high-priced junior faculty with even higher-priced newer ones. Thus, accounting faculty members should not be "punished" for coauthoring works. Rather, accounting department administrators and deans may be well advised to encourage research collaboration.

NOTES

1. Examining every n th year is not uncommon to allow a broader time frame to be studied without overburdening data collection (cf., Glanzel, 2002; McDowell & Melvin, 1983).

2. We use the journal quality classifications originally developed in the Hasselback et al. studies. Many publications reference the Hasselback et al. studies (e.g., Almer & Single, 2007; Bernardi, 2004; Gallivan & Benbunan-Finch, 2008; Hermanson, 2008; Hutchison & White, 2003; McMillan & Guffey, 2006) and their ranking of the top-40 journals in the accounting discipline (e.g., Bernardi, Melton, Roberts, & Bean, 2008). For example, Bernardi et al. (2008) use Hasselback et al.'s (2003) top-40 journals to determine the level of ethics research in the accounting discipline's top journals.

3. Fogarty and Ruhl (1997) and Maranto and Streuly (1994) find the rankings of institutions, and the rankings of the initial-hiring institutions of their doctoral graduates are highly correlated.

4. Similarities in accounting program rankings may be caused by authors who graduate from the top programs having "a disproportionate share of publications in top-notch journals" (Chan et al., 2007, p. 187).

ACKNOWLEDGMENTS

We thank the reviewers and participants at the 2009 Annual Congress of the European Accounting Association and the 2009 Annual Meeting of the American Accounting Association. Also, we thank the two anonymous reviewers for their comments, particularly editor Vicky Arnold for her valuable input.

REFERENCES

- Almer, E. D., & Single, L. E. (2007). Shedding light on the AICPA work/life and women's initiatives research: What does it mean to educators and students? *Issues in Accounting Education*, 22(1), 67–77.
- Barnett, A. H., Ault, R. W., & Kaserman, D. L. (1988). The rising incidence of co-authorship in economics: Further evidence. *The Review of Economics and Statistics*, 70(3), 539–543.
- Bazley, J. D., & Nikolai, L. A. (1975). A comparison of published accounting research and qualities of accounting faculty and doctoral programs. *The Accounting Review*, 50, 605–610.
- Beattie, V., & Goodacre, A. (2004). Publishing patterns within the UK accounting and finance academic community. *The British Accounting Review*, 36(1), 7–44.
- Bernardi, R. A. (2004). Suggestions for providing legitimacy to ethics research. *Issues in Accounting Education*, 19(1), 145–146.
- Bernardi, R. A., Melton, M. R., Roberts, S. D., & Bean, D. F. (2008). Fostering ethics research: An analysis of the accounting, finance and marketing disciplines. *Journal of Business Ethics*, 82(1), 157–170.
- Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: Implications for scientific and technical human capital. *Research Policy*, 33(4), 599–657.
- Brown, C. L., Chan, K. C., & Lai, P. (2006). Marketing journal coauthorships: An empirical analysis of coauthor behavior. *Journal of Marketing Education*, 28(1), 17–25.
- Brown, L. D., & Gardner, J. C. (1985). Applying citation analysis to evaluate the research contributions of accounting faculty and doctoral programs. *The Accounting Review*, 60(2), 262–277.
- Brown, L. D., & Laksmna, I. (2004). Ranking accounting Ph.D. programs and faculties using Social Science Research Network downloads. *Review of Quantitative Finance and Accounting*, 22(3), 249–266.
- Carpenter, C. G., Crumbley, D. L., & Strawser, R. H. (1974). A new ranking of accounting faculties and doctoral programs. *Journal of Accountancy*, 137(6), 90–94.
- Chan, K. C., Chen, C. R., & Cheng, L. T. W. (2007). Global ranking of accounting programmes and the elite effect in accounting research. *Accounting and Finance*, 47, 187–220.
- Coupe, T. (2004). What do we know about ourselves? On the economics of economics. *Kyklos*, 57, 197–216.
- Cranfield University, School of Management. (2003). *Journal recommendations for academic publication*. Unpublished working paper. Cranfield University School of Management, Bedfordshire, UK.

- Cronin, B., Shaw, D., & LaBarre, K. (2004). Visible, less visible, and invisible work: Patterns of collaboration in 20th century chemistry. *Journal of the American Society for Information Science and Technology*, 55(2), 160–168.
- Dombrow, J., & Turnbull, G. K. (2002). Individuals and institutions publishing research in real estate: 1989–1998. *Journal of Real Estate Literature*, 10(1), 45–92.
- Durden, G. C., & Perri, T. J. (1995). Coauthorship and publication efficiency. *Atlantic Economic Journal*, 23(1), 69–76.
- Englebrecht, T. E., Bisping, T., Anderson, M. M., & Hasselback, J. R. (2008). A further inquiry into the scholarly productivity of academic accountants: Twenty years of evidence from classes of 1980–82. *Advances in Accounting, Incorporating Advances in International Accounting*, 24, 24–31.
- Estes, R. W. (1970). A ranking of accounting programs. *Journal of Accountancy*, 130(1), 86–90.
- Fogarty, T. J., & Ruhl, J. M. (1997). Institutional antecedents of accounting faculty research productivity: A LISREL study of the “best and brightest”. *Issues in Accounting Education*, 12, 27–48.
- Gallivan, M., & Benbunan-Finch, R. (2008). Exploring the relationship between gender and career outcomes for social scientists: Implications for research on IS scholarship. *Information Technology & People*, 21(2), 178–204.
- Glanzel, W. (2002). Coauthorship patterns and trends in the sciences (1980–1998): A bibliometric study with implications for database indexing and search strategies. *Library Trends*, 50(3), 461–473.
- Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis* (3rd ed.). New York: Macmillan.
- Hall, T. W., & Ross, W. R. (1991). Contextual effect in measuring accounting faculty perceptions of accounting journals: An empirical test and updated journal rankings. *Advances in Accounting*, 8, 161–182.
- Hasselback, J. R. (1974–1996). *Accounting faculty directory 1974–1996*. Englewood Cliffs, NJ: Prentice-Hall.
- Hasselback, J. R. (1997–2003). *Accounting faculty directory 1997–2003*. Upper Saddle River, NJ: Prentice-Hall.
- Hasselback, J. R. (2004–2006). *Accounting faculty directory 2004–2007*. Upper Saddle River, NJ: Pearson Education, Inc.
- Hasselback, J. R., Reinstein, A., & Schwan, E. S. (2000). Benchmarks for evaluating the research productivity of accounting faculty. *Journal of Accounting Education*, 18(2), 79–97.
- Hasselback, J. R., Reinstein, A., & Schwan, E. S. (2003). Prolific authors of accounting literature. *Advances in Accounting*, 20, 95–125.
- Hermanson, D. R. (2008). What I have learned so far: Observations on managing an academic accounting career. *Issues in Accounting Education*, 23(1), 53–66.
- Hudson, J. (1996). Trends in multi-authored papers in economics. *Journal of Economic Perspectives*, 10(3), 153–158.
- Hull, R. P., & Wright, G. B. (1990). Faculty perceptions of journal quality: An update. *Accounting Horizons*, 4(1), 77–98.
- Hutchison, P. D., & White, C. G. (2003). The *Journal of the American Taxation Association* 1979–2000: Content, participation, and citation analyses. *The Journal of the American Taxation Association*, 25(1), 100–121.
- Jacobs, F. A., Hartgraves, A. L., & Beard, L. H. (1986). Publication productivity of doctoral alumni: A time-adjusted model. *The Accounting Review*, 61(1), 179–187.

- Johnson, D. (1997). Getting noticed in economics: The determinants of academic citations. *American Economist*, 41(1), 42–53.
- Jolly, S. A., Schroeder, R. G., & Spear, R. K. (1995). An empirical investigation of the relationship between journal quality ratings and promotion and tenure decisions. *Accounting Educators' Journal*, 7, 47–68.
- Laband, D. N., & Tollison, R. D. (2000). Intellectual collaboration. *The Journal of Political Economy*, 108(3), 632–662.
- Laband, D. N., & Tollison, R. D. (2006). Alphabetized coauthorship. *Applied Economics*, 38(14), 1649–1653.
- Laband, D. N., & Wells, J. M. (1998). The scholarly journal literature of economics: A historical profile of the AER, JPE, and QJE. *American Economist*, 42(2), 47–58.
- Manton, E. J., & English, D. E. (2007). The trend toward multiple authorship in business journals. *Journal of Education for Business*, 82(3), 164–168.
- Maranto, C. L., & Streuly, C. A. (1994). The determinants of accounting professors' publishing productivity: The early career. *Contemporary Accounting Research*, 10(2), 387–407.
- Maske, K. L., Durden, G. C., & Gaynor, P. E. (2003). Determinants of scholarly productivity among male and female economists. *Economic Inquiry*, 41(4), 555–564.
- Mason, P. M., Stegall, J. W., & Fabritius, M. M. (1997). Economics journal ranking by type of school: Perceptions versus citations. *Quarterly Journal of Business and Economics*, 36(1), 69–79.
- McMillan, J. J., & Guffey, D. M. (2006). Twenty years of *Advances in Accounting* – 1984–2003. *Journal of American Academy of Business*, 8(1), 32–39.
- McDowell, J. M., & Melvin, M. (1983). The determinants of co-authorship: An analysis of the economics literature. *The Review of Economics and Statistics*, 65(1), 155–160.
- McDowell, J. M., Singell, L. D., & Stater, M. (2006). Two to Tango? Gender differences in the decisions to publish and coauthor. *Economic Inquiry*, 44(1), 153–168.
- McDowell, J. M., & Smith, J. K. (1992). The effect of gender-sorting on propensity to coauthor: Implications for academic promotion. *Economic Inquiry*, 30(1), 68–82.
- Morris, J. L., Cudd, R. M., & Crain, J. L. (1990). A study of the potential bias in accounting journal ratings: Implications for promotion and tenure decisions. *Accounting Educators' Journal*, 2, 46–55.
- Nathan, S., Hermanson, D. R., & Hermanson, R. H. (1998). Co-authoring in refereed journals: Views of accounting faculty and department chairs. *Issues in Accounting Education*, 13(1), 79–92.
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673–690.
- Oh, W., Choi, J. N., & Kim, K. (2006). Coauthorship dynamics and knowledge capital: The patterns of cross-disciplinary collaboration in information systems research. *Journal of Management Information Systems*, 22(3), 265–292.
- Petry, G. H. (1988). A statistical analysis of worldwide coauthorship relationships in scholarly journals of business. *Journal of Economics and Business*, 40(2), 169–176.
- Piette, M. J., & Ross, K. L. (1992). An analysis of the determinants of co-authorship in economics. *Journal of Economic Education*, 23(2), 277–283.
- Rutledge, R. W., & Karim, K. E. (2009). Determinants of coauthorship for the most productive authors in accounting literature. *Journal of Education for Business*, 84(3), 130–134.
- Sauer, R. D. (1988). Estimates of the returns to quality and coauthorship in economic academia. *The Journal of Political Economy*, 96(4), 855–866.

- Schroeder, R. G., Payne, D. D., & Harris, D. G. (1988). Perceptions of accounting publication outlets. *Accounting Educators' Journal*, 1, 1–17.
- Smith, L. M. (1994). Relative contributions of professional journals to the field of accounting. *Accounting Educators' Journal*, 6(Spring), 1–31.
- Sobel, R. S., & Taylor, J. J. (2004). The last 30 years of public choice: An analysis of author and institution rankings. *Public Choice*, 120(3–4), 331–352.
- Stammerjohan, W. W., & Hall, S. C. (2002). Evaluation of doctoral programs in accounting: An examination of placement. *Journal of Accounting Education*, 20(1), 1–27.
- StataCorp. (1997). *Reference manual A-F (release 5)*. College Station, TX: Stata Press.
- Stine, R. A. (1995). Graphical interpretation of variance inflation factors. *The American Statistician*, 49(1), 53–56.
- Sutter, M., & Kocher, M. (2004). Patterns of co-authorship among economics departments in the USA. *Applied Economics*, 36(4), 327–333.
- Swanson, E. P. (2004). Publishing in the majors: A comparison of accounting, finance, management, and marketing. *Contemporary Accounting Research*, 21(1), 223–255.
- Taylor, S. W., Fender, B. F., & Burke, K. G. (2006). Unraveling the academic productivity of economists: The opportunity costs of teaching and service. *Southern Economic Journal*, 72(4), 846–859.
- Trieschmann, J. S., Dennis, A. R., Northcraft, G. B., & Niemi, A. W. (2000). Serving multiple constituencies in business schools: M.B.A. program versus research performance. *Academy of Management Journal*, 43(6), 1130–1141.
- Urbancic, F. (1992). The extent of collaboration in the production of accounting research. *Accounting Educators' Journal*, 4, 47–61.