ADVANCES IN MANAGEMENT ACCOUNTING

VOLUME 17

MARC J. EPSTEIN
JOHN Y. LEE
Editors

ADVANCES IN MANAGEMENT ACCOUNTING

Series Editors: Marc J. Epstein and John Y. Lee

Volumes 1–16: Advances in Management Accounting

ADVANCES IN MANAGEMENT ACCOUNTING VOLUME 17

ADVANCES IN MANAGEMENT ACCOUNTING

EDITED BY

MARC J. EPSTEIN

Rice University, USA

JOHN Y. LEE

Pace University, USA

JAI

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

JAI Press is an imprint of Emerald Group Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2009

Copyright © 2009 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: 978-1-84855-266-1
ISSN: 1474-7871 (Series)

ISOQAR QUALITY ASSURED ISO 9001 UKAS QUALITY MANAGEMENT 026
Certificate number 1985

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

INVESTOR IN PEOPLE

LIST OF CONTRIBUTORS

Mohamed E. Bayou	School of Management, University of Michigan-Dearborn, Dearborn, MI, USA
Lisa Bryant-Kutcher	College of Business, Colorado State University, Fort Collins, CO, USA
Adriana Rejc Buhovac	Faculty of Economics, University of Ljubljana, Slovenia
Yee-Ching Lilian Chan	DeGroote School of Business, McMaster University, Ontario, Canada
Bea Chiang	School of Business, The College of New Jersey, Ewing, NJ, USA
Harry Zvi Davis	Zicklin School of Business, CUNY – Baruch College, New York, USA
Marc J. Epstein	Jones Graduate School of Management, Rice University, Houston, TX, USA
Arron Scott Fleming	College of Business and Economics, West Virginia University, Morgantown, WV, USA
Laura Francis-Gladney	State University of New York Institute of Technology, Utica, NY, USA
Clinton Free	Queen's Business School, Queen's University, Ontario, Canada
Jean-François Henri	Université Laval, Québec, Canada
Denise A. Jones	Mason School of Business, College of William and Mary, Williamsburg, VA, USA

John Y. Lee	Lubin School of Business, Pace University, Pleasantville, NY, USA
Norman Macintosh	Queen's Business School, Queen's University, Ontario, Canada
Nace Magner	Western Kentucky University, Bowling Green, KY, USA
Roger Mesznik	Columbia University, New York, NY, USA
Sebastiaan Morssinkhof	University of Twente, The Netherlands
Parunchana Pacharn	Faculty of Business, Brock University, St. Catharines, Ontario, Canada
Alfred Seaman	
Luk Warlop	Katholieke Universiteit Leuven, Belgium
Robert B. Welker	School of Accountancy, Southern Illinois University, Carbondale, IL, USA
Sally K. Widener	Jones Graduate School of Management, Rice University, Houston, TX, USA
Marc Wouters	University of Twente, The Netherlands

EDITORIAL BOARD

Jacob G. Birnberg
University of Pittsburgh

Germain B. Boer
Vanderbilt University

Adriana Rejc Buhovac
University of Ljubljana

Donald K. Clancy
Texas Tech University

Robin Cooper
Emory University

Srikant M. Datar
Harvard University

Antonio Davila
Stanford University

Alan S. Dunk
University of Canberra

Nabil S. Elias
University of North Carolina

Kenneth J. Euske
Naval Postgraduate School

Eric G. Flamholtz
University of California

George J. Foster
Stanford University

Eli M. Goldratt
Avraham Y. Goldratt Institute

John Innes
University of Dundee

Larry N. Killough
Virginia Polytechnic Institute

Carol J. McNair
U.S. Coast Guard Academy

Neale O'Connor
University of Hong Kong

James M. Reeve
University of Tennessee

Karen L. Sedatole
Michigan State University

George J. Staubus
University of California

Lourdes White
University of Baltimore

Sally K. Widener
Rice University

STATEMENT OF PURPOSE AND REVIEW PROCEDURES

Advances in Management Accounting (AIMA) is a professional journal whose purpose is to meet the information needs of both practitioners and academicians. We plan to publish thoughtful, well-developed articles on a variety of current topics in management accounting, broadly defined.

AIMA is to be an annual publication of quality-applied research in management accounting. The series will examine areas of management accounting, including performance evaluation systems, accounting for product costs, behavioral impacts on management accounting, and innovations in management accounting. Management accounting includes all systems designed to provide information for management decision-making. Research methods will include survey research, field tests, corporate case studies, and modeling. Some speculative articles and survey pieces will be included where appropriate.

AIMA welcomes all comments and encourages articles from both practitioners and academicians.

REVIEW PROCEDURES

AIMA intends to provide authors with timely reviews clearly indicating the acceptance status of their manuscripts. The results of initial reviews normally will be reported to authors within eight weeks from the date the manuscript is received. Once a manuscript is tentatively accepted, the prospects for publication are excellent. The author(s) will be accepted to work with the corresponding Editor, who will act as a liaison between the author(s) and the reviewers to resolve areas of concern. To ensure publication, it is the author's responsibility to make necessary revisions in a timely and satisfactory manner.

EDITORIAL POLICY AND MANUSCRIPT FORM GUIDELINES

1. Manuscripts should be type written and double-spaced on $8^1/_2''$ by $11''$ white paper. Only one side of the paper should be used. Margins should be set to facilitate editing and duplication except as noted:
   a. Tables, figures, and exhibits should appear on a separate page. Each should be numbered and have a title.
   b. Footnote should be presented by citing the author's name and the year of publication in the body of the text; for example, Ferreira (1998); Cooper and Kaplan (1998).
2. Manuscripts should include a cover page that indicates the author's name and affiliation.
3. Manuscripts should include on a separate lead page an abstract not exceeding 200 words. The author's name and affiliation should not appear on the abstract.
4. Topical headings and subheadings should be used. Main headings in the manuscript should be centered, secondary headings should be flush with the left hand margin. (As a guide to usage and style, refer to the William Strunk, Jr., and E.B. White, *The Elements of Style*.)
5. Manuscripts must include a list of references, which contain only those works actually cited. (As a helpful guide in preparing a list of references, refer to Kate L. Turabian*, A Manual for Writers of Term Papers, Theses, and Dissertations*.)
6. In order to be assured of anonymous review, authors should not identify themselves directly or indirectly. Reference to unpublished working papers and dissertations should be avoided. If necessary, authors may indicate that the reference is being withheld for the reason cited above.
7. Manuscripts currently under review by other publications should not be submitted. Complete reports of research presented at a national or regional conference of a professional association and "State of the Art" papers are acceptable.
8. Four copies of each manuscript should be submitted to John Y. Lee at the address below under Guideline 11.

9. A submission fee of $25.00, made payable to Advances in Management Accounting, should be included with all submissions.
10. For additional information regarding the type of manuscripts that are desired, see "*AIMA* Statement of Purpose."
11. Inquires concerning Advances in Management Accounting may be directed to either one of the two editors:

Marc J. Epstein
Jones Graduate School of Administration,
Rice University, Houston,
Texas 77251-1892

John Y. Lee
Lubin School of Business,
Pace University,
Pleasantville, NY 10570-2799

INTRODUCTION

This volume of *Advances in Management Accounting (AIMA)* begins with a article by Bryant-Kutcher, Jones, and Widener on the issues involving strategic human capital.

Economic theory posits that production factors that are both difficult to imitate and capable of creating organizational efficiencies can generate economic rents and sustain long-term competitive advantage. Based on a survey of 106 firms, they measure 4 dimensions of strategic human capital and find that the market values strategic human capital that has the capability to create efficiencies in the organization. They also discuss implications for the reporting of human capital in intellectual capital reports and offer suggestions for future research.

The next article by Epstein and Buhovac provides a model and a methodology for evaluating performance in information technology to help management justify and evaluate their information system initiatives and make better resource allocation decisions. The IT Contribution Model and the subsequent IT Payoff Methodology is illustrated and empirically tested in an international firm's operations context. The study shows that the methodology's requirement for active employee involvement in the identification of the critical drivers of success, the expected outputs of the IT initiative, in particular, substantially facilitates the IT initiative implementation by increasing the level of understanding and acceptance.

The next article by Morssinkhof, Wouters, and Warlop addresses purchasing decisions and the use of total cost of ownership information which is based on a monetary quantification of nonfinancial attributes and aggregation into a summary measure, such as cost per hour, per wafer, or per kilometer. From an accounting point-of-view, one intricate issue is the accuracy of the monetary quantification and how this affects decision-making. The authors distinguish three different kinds of inaccurate monetary quantification, and investigate the weights that decision makers attach to those attributes that are inaccurately quantified and included in the information. They investigate whether those weights depend on reflective thinking and experience. This question is relevant in all decision-making situations that involve monetary quantification of attributes and subsequent aggregation, such as in activity-based costing, net present value calculations for capital budgeting, or cost–benefit analyses in public administration. The authors found support for the hypothesis that reflective thinking increases the weight decision makers attach to the attributes, but not for the hypothesis that reflective thinking would reduce the weight.

In the next article, Fleming examines the concern over rising levels of executive compensation. Individuals on the compensation committee of the board of directors collectively determine executive compensation and are responsible for maintaining the pay-for-performance standard. The author examines the process of exaggeration of a group decision over individual beliefs and the impact of leadership upon a committee's outcome when making compensation awards. Using an experiment with 98 subjects role-playing as compensation committee members, the author shows that in a committee of individuals where a coterie and a majority belief is present, group polarization occurs and the compensation results are exaggerated as compared to individual beliefs. Fleming finds that the appointment of a leader as chair of the committee, either in the majority or minority view, has a moderating effect on the group outcome. These results highlight the potential for agency costs in the group decision process that may be found in the executive compensation-setting environment.

The article by Chan and Seaman investigates the alignment of performance management system with the strategy, structure, and organizational outcome in Canadian health care organizations using balanced scorecard as the framework for assessing the health care organization's performance management system and outcome. CEO and clinical unit managers were surveyed for their perceptions on their organization's strategy, autonomy structure, performance management system, and organizational performance. The results indicate that patient satisfaction is the primary and most significant perspective of the depicted balanced scorecard in organizational performance. Patient satisfaction and research criteria, however, are the significant perspectives of a balanced scorecard in an organization's performance management system, which are linked to strategy, autonomy structure, and organizational performance. The results further show that the strategy-structure links operated as suggested. Surprisingly, strategy on service innovation has a negative impact on the organizational outcome of patient satisfaction. Uncertainty from continuous development and organizational change in pursuing service innovation and cost-cutting measures in response to fiscal constraints are plausible explanations of the adverse impact reported.

In the next article, Chiang reports on a related study that investigates how system integration in different forms is related to the success of using the balanced scorecard for performance measurement. The use of a BSC in performance evaluation is considered in five contexts: determining cost, measuring efficiency, ensuring quality and customer satisfaction measure, promoting continuous innovation, and monitoring contract negotiation. The findings indicate that system integration defined in the study is positively related to the success of the BSC in all five decision perspectives. The author concludes that hospitals need a streamlined information integration across the continuum of care to better assess the operation results, in both organizational and technical perspectives.

The article by Davis, Mesznik, and Lee attempts to make a contribution to the fuzzy logic application literature in accounting by examining the key issue in the use of fuzzy logic: how to find an optimum where the costs of reducing fuzziness are justified by decision makers. To address the issue of finding the optimal number of classes, the authors define the objective function as being cost minimization, and seek to determine the costs and benefits of increasing the number of classifications and ask whether an internal optimum is identifiable and achievable. The article assumes, *ceteris paribus*, less fuzziness is preferable to more fuzziness, but fuzziness can only be reduced through the use of more categories whose creation is costly. More fuzziness is costly, but so is the creation of additional categories to alleviate the fuzziness. When the authors arrive at the optimal number of clusters that corresponds to a minimal total cost, that number may not be the same as the ‘‘natural’’ number of categories. It is, nonetheless, a very useful and certainly practical way of deciding on the number of classifications. The approach employed in this study is not confined to a management accounting information environment. It is versatile for application to any information environment where measurable classifications exist.

In the next article, Francis-Gladney, Welker, and Magner examines two situational factors that may affect perceptions of pseudo-participation in budgeting: budget favorability (receiving a much better or much worse budget than requested) and disclosure of budget intention (the decision maker discloses or does not disclose a preliminary budget before the budget decision, with the final budget exactly matching the preliminary budget). As hypothesized, budget participants had a self-serving tendency to discount pseudo-participation as the cause of low influence when they received a favorable budget. However, contrary to a hypothesized effect, budget participants did not have a self-serving tendency to inflate pseudo-participation as the cause of low influence when they received an unfavorable budget. Instead, they formed strong, unbiased pseudo-participation perceptions. Also contrary to a hypothesized effect, the budget decision maker's disclosure of an intended budget, which should have provided clear indications of an insincere request for budget input, did not increase perceptions of pseudo-participation. Budget outcomes that indicate low influence may evoke such strong perceptions of pseudo-participation as to override other information that suggests pseudo-participation.

The article by Henri provides an integrated view of performance measurement systems by developing a taxonomy reflecting the interdependencies among system components. The study investigates the extent to which similar patterns across various system dimensions occur with regularity. Using a survey of manufacturing firms, this taxonomy develops three aspects of the system process: the design involving the mix of financial, customer, internal processes, innovation and learning measures; the use of monitoring, strategic decision-making, attention focusing, and legitimization; and the revision in performance indicators. Three patterns of relationships reflecting the role and importance of performance measurement systems within the organization emerge: the system as an outcomes surveillance mechanism, the system as a management support tool, and the system as an institutionalized organizational process. The study contributes to the management accounting literature by providing a different understanding of the various levels of integration of performance measurement systems within organizational routines.

In the next article, Pacharn examines three structural properties of accounting commonly embedded in generally accepted accounting principles in a two-period principal-agent model. These structural properties are conservation of income, consistency, and selective recognition. The article illustrates that these properties are essential for the use of accounting information in management performance evaluation: they are necessary conditions for an accounting mechanism to be more efficient than a direct revelation mechanism. The tradeoff between the gain from the information revelation and the incentive cost of discretion determines whether contracting is more efficient under the accounting mechanism or under the direct revelation mechanism.

The research note by Bayou addresses the issue of cost allocation in operating a governmental project either as an independent, self-supporting municipal enterprise insulated from political influence or as a special revenue fund financed by tax levies. The cost allocation issue plays an important role in this decision since the development of an acceptable user charge requires calculations of the full-cost per unit of service. The article selects one of the largest municipal enterprises in the United States whose pricing practices are typical of those followed by many cities in the United States. The second research note by Free and Macintosh identifies the radical change in Enron's corporate culture that took place from the Lay–Kinder era (1986–1996) to the Lay–Skilling era (1997–2001). It argues that this was a major cause of neutralizing the Enron controls, which in turn proved to be a major factor in Enron's fall into bankruptcy. The article contributes to the literature by drawing attention to the rich but untold story of Enron's governance and control and also extends the research linking corporate culture and control systems.

We believe the 12 articles in Volume 17 represent relevant, theoretically sound, and practical studies the discipline can greatly benefit from. These manifest our commitment to providing a high level of contributions to management accounting research and practice.

Marc J. Epstein
John Y. Lee
Editors

MARKET VALUATION OF INTANGIBLE RESOURCES: THE USE OF STRATEGIC HUMAN CAPITAL☆

Lisa Bryant-Kutcher, Denise A. Jones and Sally K. Widener

ABSTRACT

Economic theory posits that production factors that are both difficult to imitate and capable of creating organizational efficiencies can generate economic rents and sustain long-term competitive advantage. Using survey data for 106 firms, we measure four dimensions of strategic human capital and find that the market values strategic human capital that has the capability to create efficiencies in the organization and is also difficult for competitors to imitate. We discuss implications for the reporting of human capital in intellectual capital reports and offer suggestions for future research.

☆Part of the data in this study was obtained through a survey constructed for Sally K. Widener's dissertation.

Advances in Management Accounting, Volume 17, 1–42
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017003

INTRODUCTION

Corporate executives often state, ''people are our most important asset.'' An August 28, 2000 Business Week editorial proclaimed that ''human capital is the *only* asset.'' The International Federation of Accountants (IFAC, 1998) positions human capital as the foundation for the building of intellectual capital, which they conclude is becoming increasingly important to companies. Similarly, the Steering Committee of the Financial Accounting Standards Board (FASB, 2001, p. 22) states ''the important assets of enterprises are increasingly intangible. There is general agreement among business observers and analysts that the big contributors to business success are a company's people ...'' These statements highlight the increasing reliance and value managers place on human capital in order to compete in today's global and rapidly changing economy (Lev, 2001).

An early stream of accounting research on human capital attempted to develop a measure of human capital to record as an asset in the financial statements (e.g., Flamholtz, 1971, 1985; Lev & Schwartz, 1971). More recently, interest has shifted to investigating whether the market values human capital even though organizations do not formally measure and record human capital as an asset in their financial statements (e.g., Abdel-khalik, 2003; Ballester, Livnat, & Sinha, 2002). Additionally, recent studies have found that human resource practices, such as workplace attitudes and fees resulting from football players' contractual obligations are positively correlated with market values (Amir & Livne, 2005; Ballou, Godwin, & Shortridge, 2003). However, Lev (2001, p. 76) states, ''Of the various intangible assets ..., we have the least systematic information on human resources.''[1]

One difference between human capital and other firm resources is that the firm does not own its employees (Coff, 1997). Since the employees can leave at any time, employee-related expenditures will not necessarily translate into firm value. As Coff (1997, p. 374) states, ''Merely having talented employees does not mean that a sustainable advantage exists.'' Much of the past literature attempts to measure an overall value for the workforce. In this article, we discuss why the workforce cannot be treated homogeneously and develop a hypothesis about which portion of the workforce is likely to add value to the firm. More specifically, we extend the previous literature by focusing on the value of the *strategic* human capital of the firm. Having a workforce (i.e., human capital) alone is not sufficient for a firm to earn a competitive advantage. Rather a firm must utilize the workforce as a strategic resource to sustain a competitive advantage.

By using a theoretically driven measure of human capital, we are able to separate human capital into various components that should be valued differently by the market. This makes for a more powerful and insightful test.

Before continuing, we must first precisely define a strategic resource. Resources, or production factors,[2] that are both difficult to imitate and capable of creating organizational efficiencies are labeled *strategic* resources (Barney, 1991). Firms possess a stock of *capable* resources that create efficiencies and enhance effectiveness, with which they can deploy their strategy. Capable resources are necessary to establish a competitive advantage; however, they are not sufficient to sustain a long-term competitive advantage since competitors may be able to replicate the source of the competitive advantage. Capable resources generate *sustainable* organizational rents and help maintain long-term competitive advantage only when they are also difficult for competitors to imitate (Barney, 1991; Dierickx & Cool, 1989; Lippman & Rumelt, 1982; Lev & Schwartz, 1971; Becker, 1962). Three characteristics of human capital that make it difficult for competitors to imitate are: (1) firm-specific or idiosyncratic knowledge and skills, (2) causal ambiguity (tasks or processes that are not clearly defined or linked to firm performance), and (3) high levels (stock) of the resource (Amit & Schoemaker, 1993; Barney, 1991; Dierickx & Cool, 1989; Lippman & Rumelt, 1982; Williamson, 1979). Both attributes – difficult to imitate and capable of creating organizational efficiencies – are necessary for a firm to sustain its competitive advantage over the longer term, thus defining a strategic resource (Lev, 2001; Barney, 1991; Lippman & Rumelt, 1982). Assuming that the market can obtain adequate information regarding a firm's strategic resources, they should be valued in the marketplace since they have the ability to generate sustainable economic rents. One potential type of strategic resource is human capital.[3]

Using survey and archival data for a sample of 106 firms, we find, after controlling for the book value of equity, earnings, and salary levels (as proxied by pension and retirement plan costs), a positive relation between market value and strategic human capital that is *both* capable of creating efficiencies and difficult to imitate. These findings are consistent with underlying economic theory. We also show that this finding is robust across alternative model specifications.

In addition, we find a positive relation between market value and the spread of human capital throughout the organization, suggesting that the market values human capital that is difficult to imitate. Finally, we find that the market negatively values human capital that has the capability to create organizational efficiencies. At first glance, a negative coefficient on this type of human capital seems somewhat surprising. The intuition behind this finding is that the market is treating investments in a mobile, *capable* resource that can be imitated by other firms similar to other expenses. Although the firm may gain a short-lived advantage from these investments, the market recognizes that this is not a sustainable condition since the human capital can be imitated by a competitor.

To explain further, we provide the following example. Consider the service counter help at McDonalds. Assume that these workers are capable of creating efficiencies for the organization; however, they are free to leave at any time. A rival firm, such as Burger King, may be able to either hire away McDonalds' employees or hire similar workers and train them in providing effective counter service. The market should treat the investment McDonalds makes in this type of worker as an expense since McDonalds is unable to generate sustainable competitive advantage using this strategy. On the other hand, assume that McDonalds' employees are not valued in the labor market because their training is wholly firm-specific, or the operational process of the employees' efforts is not transparent to Burger King, or the employees' capabilities are encompassed in a large stock of employees all working together. Since the employees are capable of creating organizational efficiencies *and* this value is difficult to imitate by competitors, McDonalds is able to sustain its competitive advantage and generate future rents. Now the market should positively value the investment McDonalds makes in its employees.

As unrecorded intangibles become more critical to firms' ability to succeed, it becomes even more important that we increase our understanding of the definition, measurement, and valuation of strategic human capital. This study increases our knowledge of human capital and makes both theoretical and practical contributions to the literatures on intangibles and intellectual capital as follows. Theoretically, we demonstrate that strategic human capital possesses properties that we can characterize as being similar to an unrecognized asset. However, a specific type of human capital – which is capable of creating efficiencies but able to be imitated by competitors – can be characterized as being similar to an expense. We are not proposing that human capital necessarily be recorded as an asset. Rather, we are saying that a firm's workforce is not homogeneous and should not be treated as such. Roslender and Fincham (2001) discuss how "what gets measured gets managed." One important contribution of this article is to shift the focus of managers away from thinking about employees in terms of the amounts spent on salaries, training, and development, and to move toward thinking about the situations where employees add value to the firm.

From a practical perspective, this study contributes to the literature by informing managers about when investments in human capital are positively valued. This is particularly important in determining disclosures, such as those found in the intellectual capital statements that are prevalent in Europe (Lev & Zambon, 2003). As human capital becomes increasingly important to firms, it will be imperative for managers to demonstrate to stakeholders that their human capital is an investment in the firm and, accordingly, should be valued as such. Lev and Zambon (2003, pp. 597–598) state, "... intangibles will continue to be vital to companies, and the challenge of how to manage, measure and visualize them has to be addressed in theoretical and practical terms." IFAC (1998) echoes this message, acknowledges that this area is wide-open, and asserts that managers will want to experiment with various measurements and reporting practices for intellectual capital. Moreover, there is a growing movement for accounting harmonization across countries. Yet, the definition and accounting for intangibles, of which human capital is an increasingly important component, varies widely (Stolowy & Jeny-Cazavan, 2001). In summary, a better understanding of when human capital creates value helps in developing performance measures, providing better information for decision-making, valuing the firm, and being able to agree on a common definition and understanding of various intangible assets across firms and countries. This study suggests that information regarding the strategic human capital that a firm employs is useful information.

This study is organized as follows. The second section provides an overview of intellectual capital, the strategic human capital literature, and develops the hypothesis. The third section presents the research design, measurement of the variables, and the sample. In fourth section, we present the analyses and discuss the results. Finally, in fifth section, we provide concluding comments and limitations of this research study.

HUMAN CAPITAL AND MARKET VALUATION

Importance of Human Capital Across Firms

Albeit simplistic, the market value of the firm can be thought of as the sum of the tangible or physical capital and the intangible or intellectual capital controlled by the firm less liabilities. While the specifics of intellectual capital vary, there are various components that comprise it including human capital, structural capital, and customer capital (Mouritsen, 1998). Edwards (1997, p. 21) states that "whatever its composition, intellectual capital is essentially an intelligence-derived production input that companies combine with other production inputs – raw materials, producer goods and physical labor – to create goods and services for sale."[4] Human capital, which is the knowledge contained in the minds of the employees or "employee know-how" is an important component of intellectual capital (Edwards, 1997, p. 23).

While physical capital is usually recorded as an asset by the firm, intellectual capital is often not recorded due to the difficulty in reliably measuring the value of intangibles. Thus, the book value of common equity in the financial statements is often lower than the value that the stock market places on common equity, resulting in a market-to-book ratio greater than one. Over the past 20 years, there has been a steady increase in the aggregate market-to-book ratio of U.S. firms (see Fig. 1, panel A). By the late 1990s, the aggregate market-to-book ratio was over 4. After the U.S. stock market correction in 2000 and 2001, the average market-to-book ratio was just under 3. It is important to note that this gap exists across industries. Fig. 1, panel B shows that while research and development (R&D) intensive industries have the highest market-to-book ratios, the average market-to-book ratio of firms in service, retail, and manufacturing industries has also been over 2 in recent years. It is apparent that factors other than stock market exuberance, such as unrecorded intangible assets or future growth opportunities, are causing the market value of equity to be greater than the book value of equity for all firms, not just high-technology firms.

The market-to-book ratios provide empirical evidence that there is unrecorded value across industries, which is likely due, at least in part, to intellectual capital. Existing literature and anecdotal evidence support the claim that human capital comprises part of the unrecorded value. While some types of intellectual capital (e.g., patents) may be more prevalent in specific industries, human capital is pervasive throughout most industries. Human capital is often the primary value-creating resource for service firms since it is usually the primary strategic interface between the firm and the customer. One of the fundamental tenets of the service-value-profit chain is that the employee is a key component of the production function necessary for firm success (Heskett, Sasser, & Schlesinger, 1997). For example, in the airline industry there is evidence that the interaction between the consumer and the provider's employees significantly influences the consumer's satisfaction (Anderson, Davis, & Widener, 2005). In R&D firms, employees are a critical component

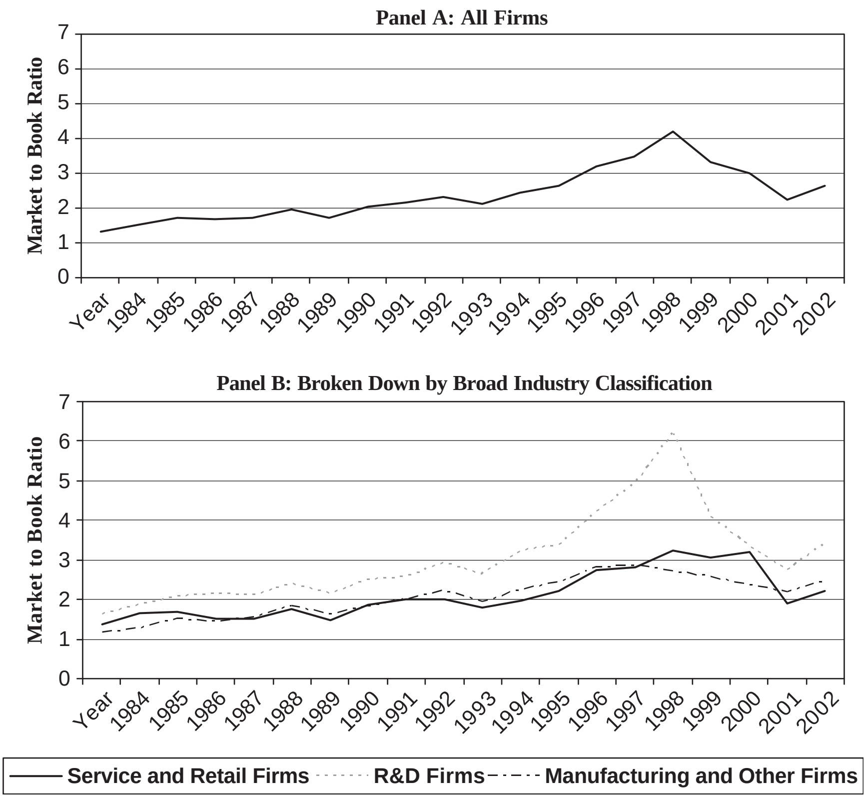


Fig. 1. Market-to-Book Ratio over the Past 20 Years. Figure Shows the Aggregate Market-to-Book Ratio for All Firms in the Compustat Database Reporting Book Value of Common Equity, Common Stock Price, and Common Shares Outstanding. The Aggregate Market-to-Book Ratio is the Aggregate Market Value of All Firms Divided by Their Aggregate Book Value at Year-End. Service and Retail Firms are All Firms with a Two Digit SIC Code of 50 or Higher. R&D Firms are All Firms in R&D-Intensive Industries (SIC Codes 28, 35, 36, 38). Manufacturing and Other Firms are All Firms in SIC Codes 01–49 Except R&D-Intensive Industries.

in creating more effective and efficient processes for the firm by designing, developing, and producing new products (Edwards, 1997). The importance of human capital to R&D-intensive firms is illustrated in the following passage from the 10-K of Donnelly Corporation, a firm in this study (1998, p. 13):

> Continued emphasis on effective research and product development is a key part of the Company's strategy for future growth ... The Company has a corporate applied research

> group, including several PhD's located at research facilities in ... The Company believes its human resources are one of its fundamental strengths ... The Company believes that this approach has increased productivity by emphasizing employee opportunity and participation aimed at continuous improvement. The Company believes this emphasis has resulted in enhanced long-term productivity, cost control and product quality and has helped the company attract and retain capable employees.

Human capital is also important to manufacturing firms, especially in today's age of advanced manufacturing technologies. Firms that compete on the basis of flexible manufacturing systems, just-in-time, total quality management, and lean production invest heavily in training costs (Snell & Dean, 1992). Public documents from Boeing, one of the firms in this study, illustrate the use of advanced manufacturing technologies and emphasis on training. They state in their 10-K (1998, p. 50)

> The 777, the Next-Generation 737, the Joint Strike Fighter, and other recent commercial and government developmental programs included early commitment of resources for integrated product teams ... and increased use of automated manufacturing processes. Although these measures have required significant current investments, substantial long-term benefits are anticipated ... Major long-term productivity gains are being aggressively pursued, with substantial resources invested in education and training.

On the basis of the importance of human capital across a variety of industries, our study investigates whether the market values strategic human capital for a cross-section of firms.

Relevant Background Literature

Beginning in the 1960s a body of research focused on human resource accounting (HRA).[5] These studies were primarily interested in how to measure and value human assets in order to record employees as an asset on the financial statements (see, e.g., the Stochastic Rewards Valuation Model developed by Flamholtz, 1971). Later research used the Stochastic Rewards Valuation Model to value members of a certified public accountant (CPA) firm (Carper & Posey, 1976) and to demonstrate how firms could value the human capital component of an acquisition (Flamholtz, 1987). One of the primary criticisms of HRA was that the focus was on "coming up with numbers rather than coming up with ways to make better management decisions" (Edwards, 1997, p. 22). Moreover, the interest in HRA in the United States seemed to wane due to unreliable measures. Now that companies compete in a global competitive market characterized as the information-age, interest in human capital is increasing.

Countering the criticism of HRA, a recent stream of literature investigates the effective management of human capital within the firm. Research has shown that management control systems emphasize ex ante behavioral controls in an environment characterized by a heavy reliance on human capital (Widener, 2004). Moreover, firms must be careful to not stifle creativity when attempting to increase productivity in knowledge-intensive organizations (Chang & Birkett, 2004). Managers also are experimenting with processes to make their human capital more transparent to both internal and external stakeholders. Skandia (1998) issues a substantial report annually to their stakeholders describing their intellectual capital, of which human capital is a component.

Academic research concludes that firms use narratives, visualization, and numbers in intellectual capital (IC) reports in order to make their intellectual capital more transparent (Mouritsen, Larsen, & Bukh, 2001). Moreover, the act of constructing and classifying intellectual capital helps facilitate organizational understanding and learning of their knowledge processes (Leitner & Warden, 2004; Grojer, 2001; Mouritsen, 1998). In addition to constructing and classifying intellectual capital, Johanson, Martensson, and Skoog (2001) identify a total of seven sub-routines in management accounting[6] that enable the firm to increase the value of its stock of knowledge and enhance organizational learning. One of the relevant takeaways, albeit implicit, from this literature, is that human capital is an important area to study. Both managers and researchers are concerned with appropriately managing human capital, engaging in organizational learning, and conveying the state of the resource to external stakeholders. In the conclusion, we discuss the implications of our study for this line of research.

If firms are making the state of their human capital more transparent to external stakeholders, then the market should value it. A large body of literature has documented that structural capital, such as R&D expenditures and patent development costs, which are immediately expensed, are value relevant (e.g., Ballester, Garcia-Ayuso, & Livnat, 2003; Green, Stark, & Thomas, 1996; Lev & Sougiannis, 1996). In addition, studies have found a positive association between market value and customer satisfaction, a type of customer capital (e.g., Ittner & Larcker, 1998). However, there is less known about human capital, most likely due to problems associated with obtaining data and constructing an appropriate measure of human capital (Abdel-khalik, 2003).

Studies have investigated the relation between value and human resource practices. For example, a study by Watson Wyatt, a consulting firm, shows that market values are positively correlated with their measure of an index of human capital that captures how organizations carry out their human resource practices (Watson Wyatt, 1999).[7] Skandia (1998) asserts that their intellectual capital encompasses the entire workforce. One relevant question is whether the market perceives this as well? Amir and Livne (2005) find a positive association between market value and transfer fees paid to buy players of a small sample of UK football clubs. However, it is difficult to generalize those results to the typical firm in which the workforce does not have a contractual obligation to remain with the firm. Using archival data on labor expense disclosed by a subset of US firms, Ballester et al. (2002) find that the market values a portion of a firm's labor costs. Ballester et al. (2002) do not identify what types of labor costs are valued. Finally, Abdel-khalik (2003) investigates the valuation of human capital using archival data on executive compensation and firm-specific data to proxy for managerial skills and finds that the market values managerial skills. One limitation of this study is that it only investigates CEOs and other executive members of the board. Contrary to these results, Bontis (1998) does not find evidence of a direct relation between human capital and firm performance; rather, he finds that human capital is associated with both structural and customer capital, which are associated with performance.

Our study differs from prior literature in three primary ways. First, we do not use self-reported measures of performance that are inherently noisy; rather, we are interested in understanding how the market views the various elements of strategic human capital. Second, we are not limited to archival executive compensation data that is only reported for the top executives in the organization nor are we limited to the study of only contractually obligated employees. Thus, our results are more generalizable across workforces and firms. Third, prior studies treat the workforce homogeneously without regard to what types of human capital add value to the firm. For example, an imitable workforce should not allow the firm to earn profits in excess of its cost of capital since there is no foundation for the "establishment of competitive advantage over rivals" (Grant, 1991, p. 117). Although this is also the case for all strategic resources, it is particularly true for human capital as the firm does not own its employees (Coff, 1997). We draw on the strategic-based resource literature that theorizes that only human capital capable of providing efficiencies *and* that is difficult to imitate will sustain competitive advantage. Thus, we take the view that the market will only value a subset of a firm's workforce (see discussion in the next sub-section). This is consistent with Barney (1991, p. 102) who states, "Of course, not all aspects of a firm's physical capital, human capital, and organizational capital are strategically relevant resources." Therefore, we contribute to the above literature stream by ascribing market value to several components of human capital, which should aid managers in determining which disclosures are relevant for investors and assist in understanding which types of human capital are investments that are valued by the market.

Valuation of Human Capital

Although human capital is not reported on the balance sheet, economists consider human capital to be valuable to the firm. Some economists (e.g., Alfred Marshall) even state that human capital is the most valuable type of capital (Lev & Schwartz, 1971). While human capital is the knowledge and/or skills possessed by the firm's workforce (Lev, 2001; Becker, 1962), *strategic* human capital is the part of the workforce that helps the firm to sustain its competitive advantage (Lev, 2001; Barney & Wright, 1998; Amit & Schoemaker, 1993).

For a firm to sustain its competitive advantage, it is necessary that strategic resources (1) have the capability to create organizational efficiencies and (2) are difficult for competitors to imitate. Resources that have the capability to create organizational efficiencies are those that "enable a firm to conceive of or implement strategies that improve its efficiency and effectiveness" (Barney, 1991, p. 106). Firms use human capital to create efficiencies, which decrease costs or enable the sale of goods for premium prices. In a service firm, emphasizing and exploiting relationships that exist between employees and their customers often allows a firm to be more effective in capitalizing on opportunities, whereas in manufacturing firms, employees often drive the effectiveness and efficiencies of advanced manufacturing techniques (Snell & Dean, 1992).

Even in the face of free entry and fully competitive behavior firms can realize persistent economic rents if there is uncertainty in the underlying resources, such as human capital (Lippman & Rumelt, 1982). This uncertainty, or lack of imitability, thus sustains the competitive advantage.[8] Lippman and Rumelt (1982, p. 419) state

> We find that uncertain imitability can lead to supernormal industry profits together with a lack of entry. Additionally, uncertain imitability provides a theoretical connection between the height of this apparent 'entry barrier' and the stable dispersion of interfirm profit rates. Finally while the standard view is that excess industry profits induce entry, this theory suggests that high profits, ceteris paribus, may well signal the presence of very successful and difficult to imitate competitors and thereby impede rational entry attempts.

In order for competitors to imitate a firm's competitive advantage, the competitor must be able to acquire the underlying resources that the firm uses to implement its strategy. If, under similar conditions, circumstances, and costs, competitors can acquire resources that rival firms use to drive their competitive advantage, then the rival's competitive advantage will be short lived (Grant, 1991). We measure three characteristics specified in the resource-based strategy literature that can cause the strategic use of human resources to be difficult to imitate: firm-specificity, causal ambiguity, and the mass or spread of the resource.[9]

Firm-specific knowledge and skills are difficult to transfer among firms (Amit & Schoemaker, 1993; Williamson, 1979), resulting in low imitability. Firm-specificity is created through investments in training designed to develop employees in operational methods, systems, and processes specific to the firm (Williamson, 1979; Becker, 1962). Because the skills and knowledge are firm-specific, human capital cannot earn returns in the labor market. Thus, investing in firm-specific human capital reduces some of the threat of voluntary turnover and facilitates the firm's long-term sustained competitive advantage (Coff, 1997; Becker, 1962).

Causal ambiguity exists when employees perform ill-defined tasks or when the link between factor inputs or effort and firm performance is not clear (Barney, 1991).[10] The accumulation and use of these resources is not deterministic or continuous, instead it is stochastic and discontinuous (Dierickx & Cool, 1989). In other words, there may be uncertainty regarding how to control the human capital, how to accumulate the important variables in the process, and/or how the human capital process produces benefit. Causal ambiguity prohibits transparency, prevents competitors from imitating the source of competitive advantage, and limits factor mobility (Lippman & Rumelt, 1982).

Lastly, when talented and valuable human capital is spread throughout the organization, it is more difficult to imitate. Barney and Wright (1998) state "… the synergistic value from a large number of individuals who work together is quite costly if not impossible for competitors to imitate." Williamson (1979) argues that frequent transactions between people lead to idiosyncratic, transaction-specific skills that cannot be specified. Therefore, the human capital is immobile because replication of its value is uncertain (Lippman & Rumelt, 1982). Dierickx and Cool (1989, p. 1507) describe it as "asset mass efficiencies." In other words, building a sufficient mass of resources enhances sustained competitive advantage since it is more difficult for competitors to replicate synergies within a large mass of resources than it is to replicate within a low level of a resource.[11]

In summary, the economics literature and the resource-based view of strategy literature assume that firms hold and control unique strategic resources or production factors that help establish a competitive advantage, earn economic rents, and provide the firm with the basis for sustaining that competitive advantage. A firm establishes competitive advantage by competing with resources that increase organizational efficiency and effectiveness and by acquiring resources that are difficult for competitors to imitate. By themselves, neither characteristic is sufficient to sustain long-term competitive advantage and to generate economic rents. A firm sustains its competitive advantage when the same resource both creates efficiencies and cannot be easily imitated (Barney, 1991). Thus, both characteristics are necessary to generate economic rents (Lippman & Rumelt, 1982). In summary, *strategic* human capital is the part of the workforce that helps the firm sustain its competitive advantage and must possess the following attributes:

1. Have the capability to create strategic organizational efficiencies, *and*
2. Be difficult to imitate as achieved by one of the following:
   a. Have firm-specific attributes, *or*
   b. Possess causal ambiguity, *or*
   c. Be part of a large mass of stock (i.e., frequency or spread).

In this study, we investigate whether the market positively values the component characteristics of strategic human capital and their joint effect. On the basis of the theory discussed above, we expect that only the joint effect (i.e., capable of achieving organizational efficiencies and difficult to imitate), which generates sustainable economic rents, will be positively associated with market value. Thus, our formal hypothesis is

H1. Human capital that is both capable of creating efficiencies and difficult to imitate is positively associated with a firm's market value.

RESEARCH METHODS

Research Design and Measurement of Financial Variables

We are interested in whether the market values the use of strategic human capital. The use of strategic human capital is measured using a survey, which is discussed in the next section. Because the survey responses capture the use of strategic human capital as of a specific date, we use a valuation model to examine whether the use of this human capital is reflected in stock prices at the time of the survey. As discussed in the sensitivity analysis section, a firm employs human capital as part of an overall strategy encompassing other firm resources. Therefore, it is important to control for other resources owned by the firm, many of which are recognized as part of book value. As such, we use a market valuation model which models price as a function of book value, net income, and other information. Consistent with prior research, we begin with the following valuation model (e.g., Ballester et al., 2003; Barth, Beaver, & Landsman, 1998):

$$P_i = \beta_0 + \beta_1 BV_i + \beta_2 NI_i + \varepsilon_i \tag{1}$$

where P_i is the fiscal year end share price for firm i, BV_i is the fiscal year end book value of equity adjusted for net pension liabilities (discussed below) per share, and NI_i is the net income per share before extraordinary items.[12]

Under current accounting rules, strategic human capital is not capitalized on the balance sheet and is therefore not included in book value. If the market considers the use of strategic human capital relevant to firm valuation and sufficiently reliable to be reflected in share prices, then, after controlling for recorded book value of equity and net income, proxies for the use of strategic human capital should be positively valued by the market. Accordingly, we expand Eq. (1) to include the underlying components of strategic human capital, as well as a control variable for salary costs

$$\begin{aligned} P_i = {} & \beta_0 + \beta_1 BV_i + \beta_2 NI_i + \beta_3 PENS_i + \beta_4 CAP_i \\ & + \beta_5 FS_i + \beta_6 CA_i + \beta_7 SPR_i + \varepsilon_i \end{aligned} \tag{2}$$

where P_i, BV_i, and NI_i are as defined earlier. The human capital variables are: (1) whether human capital has the capability to create efficiencies (CAP), (2) the firm-specificity of human capital resources (FS), (3) the ambiguity of the work performed by the human resources (CA), and (4) how extensively human capital is spread throughout the firm (SPR). The measurement of these variables is discussed in the following section.

It is possible that firms with either more employees or higher salary and employee benefit expenditures have a greater use of strategic human capital. Therefore, we include pension and retirement costs after tax, deflated by shares outstanding ($PENS$), to control for the level of expenditures on salary and employee benefits. If a firm has a defined benefit pension plan, we include only the portion of pension costs related to employee service. We use pension and retirement costs for two reasons. First, it is highly correlated with salary expenditures, which are not disclosed by many firms. Second, companies typically offer benefits packages designed to foster long-term relationships with skilled employees. Prior studies document a positive relation between market values and pension contributions indicating that pension contributions may represent an unrecorded human capital asset (e.g., Barth, 1991). We also adjust the book value of equity for the net pension asset or liability (fair value of pension assets less the accumulated benefit obligation). Under U.S. accounting rules, this is often not recognized in the financial statements and may be correlated with the presence of human capital.[13]

Hypothesis 1 posits that to be associated with firm value, human capital should *both* be capable of creating efficiencies and difficult to imitate. We test this hypothesis in two ways. First, as discussed previously, for a resource to create a sustained competitive advantage, it must be able to create efficiencies (captured by *CAP*) and be difficult for competitors to imitate (captured by *FS*, *CA*, or *SPR*). We create a composite variable (*SUSTADV*) by first ranking all of the firms by industry on each of the four human capital variables. We then create an indicator variable equal to 1 if a firm scores above the industry median on *CAP* and above the industry median on *either FS*, *CA*, or *SPR*.[14] Because of the number of firms available, we define industries broadly. We identify three broad industry groups based on the importance of human capital versus physical capital to the firm, and the ease with which the knowledge and skills of the employees can be converted into a tangible asset owned by the firm: service firms, R&D-intensive firms, and manufacturing firms. We identify service firms as those firms with an standard industrial classification (SIC) code greater than 50 (i.e., retail, wholesale, financial, insurance, and services), R&D-intensive firms as those firms with SIC codes 28, 35, 36, and 38, and manufacturing firms as those firms with SIC codes 10–49, except for 28, 35, 36, and 38. To test Hypothesis 1, we add *SUSTADV* to Eq. (2)

$$P_i = \beta_0 + \beta_1 BV_i + \beta_2 NI_i + \beta_3 PENS_i + \beta_4 CAP_i + \beta_5 FS_i + \beta_6 CA_i + \beta_7 SPR_i + \beta_8 SUSTADV_i + \varepsilon_i \tag{3}$$

Theoretically, *FS*, *CA*, and *SPR* are three distinct variables that measure a broader construct, lack of imitability. Since our data set is small, we have limited observations that are above the industry median for both *CAP* and either *FS*, *CA*, or *SPR*. Therefore, examining the imitability variables individually may result in very low power tests. The benefit of *SUSTADV* is that it is a powerful variable that captures in one variable all of the information contained in the four human capital variables. However, the problem with this variable is that it does not provide information on which attribute is valued by the market. Therefore, the second way that we test Hypothesis 1 is by creating three sustained advantage variables; one for each different type of lack of imitability. This has the benefit of distinguishing between *FS*, *CA*, and *SPR*, but may be a less powerful test. To create these variables, we interact *CAP* with each of *FS*, *CA*, and *SPR*. We then add these interaction terms to Eq. (2) both separately and at the same time

$$\begin{aligned} P_i = \beta_0 + \beta_1 BV_i + \beta_2 NI_i + \beta_3 PENS_i + \beta_4 CAP_i + \beta_5 FS_i \\ + \beta_6 CA_i + \beta_7 SPR_i + \beta_8 CAP * FS_i + \varepsilon_i \end{aligned} \tag{4}$$

$$\begin{aligned} P_i = \beta_0 + \beta_1 BV_i + \beta_2 NI_i + \beta_3 PENS_i + \beta_4 CAP_i + \beta_5 FS_i \\ + \beta_6 CA_i + \beta_7 SPR_i + \beta_8 CAP * CA_i + \varepsilon_i \end{aligned} \tag{5}$$

$$\begin{aligned} P_i = \beta_0 + \beta_1 BV_i + \beta_2 NI_i + \beta_3 PENS_i + \beta_4 CAP_i + \beta_5 FS_i \\ + \beta_6 CA_i + \beta_7 SPR_i + \beta_8 CAP * SPR_i + \varepsilon_i \end{aligned} \tag{6}$$

$$\begin{aligned} P_i = \beta_0 + \beta_1 BV_i + \beta_2 NI_i + \beta_3 PENS_i + \beta_4 CAP_i + \beta_5 FS_i + \beta_6 CA_i \\ + \beta_7 SPR_i + \beta_8 CAP * FS_i + \beta_9 CAP * CA_i + \beta_{10} CAP * SPR_i + \varepsilon_i \end{aligned} \tag{7}$$

Measurement of Human Capital Variables

We gather the data for the human capital variables (refer to the previous section for a thorough discussion and definition of these variables) through a survey that is described in the next section. See appendix for an abbreviated version of the survey. As appropriate, the responses to individual survey questions are combined to form an average summary measure. Since this study uses survey measures, we are concerned with both content and construct validity. Content validity can be assessed by (1) the "plan and procedures of construction," (2) appearance, or face validity, and (3) a measure of internal consistency through an empirical measure of reliability (Nunnally, 1978, p. 92). Construct validity can be assessed by (1) specifying an appropriate domain of observables underlying the construct, (2) using factor analysis to find relationships among the observables, and (3) using correlation analysis to find relations among the constructs (Nunnally, 1978). To establish high degrees of both content and construct validity, we took the following steps: (1) reviewed existing literature to establish appropriate domains, (2) used previously validated measures whenever possible,[15] (3) performed four in-depth field visits to learn more about the domain being measured, (4) used guidelines set forth in Dillman (1978) for the construction of survey questions, (5) pre-tested the survey on several academicians and a pilot sample of 30 respondents, and (6) performed various empirical tests. Factor analysis reveals that all measures are uni-dimensional; the Cronbach's α, which range from 0.77 to 0.85, demonstrate acceptable internal reliability; and the α coefficients exceed the inter-item correlation coefficient in all cases, which also helps demonstrate discriminant validity (Nunnally, 1978). In addition, plausible behavior of the constructs was demonstrated through a review of correlation analysis (Widener, 2004). Descriptive statistics for the multi-item variables are reported in Table 1.

We use the label "capabilities" (*CAP*) to capture what Barney (1991, p. 106) refers to as "valuable resources." We measure it using two questions taken from the underlying literature that ask whether the firm's strategic human capital enables the firm to be more efficient and effective in exploiting opportunities. We draw these questions from Barney (1991, p. 106) who specifically states that resources have the capability to be valuable "when they enable a firm to conceive of or implement strategies that improve its efficiency and effectiveness."

Since we are interested in the firm-specificity of human capital, we are interested in knowing the degree to which employees have skills and knowledge that are difficult to transfer among firms (Amit & Schoemaker, 1993; Williamson, 1979); therefore, we measure firm-specificity (*FS*) using four questions regarding the extent to which the knowledge base is specific to the firm, the ease with which experience workers could enter the firm and contribute without undergoing extensive firm-specific training, the time it would take a newly hired employee to become familiar with firm-specific customers and products and, finally, the time it would take a replacement employee to be equally effective as a current employee. These questions are drawn from underlying literature (see, e.g., Lohtia, Brooks, & Krapfel, 1994; Williamson, 1979).

Causal ambiguity is "when the link between a firm's resources and its sustained competitive advantage are poorly understood" (Barney, 1991, p. 109). We measure causal ambiguity (*CA*) using seven questions related to whether the employees' duties are repetitious, if tasks are standardized, the difficulty of monitoring and evaluating employees' effort, and if there is an understandable sequence of steps the employees follow. If activities are repetitious and lend themselves to the use of standard operating procedures,

Table 1. Reliability Measures and Descriptive Statistics for Human Capital Variables.

	Min.	Max.	Mean	Std. Dev.	Cronbach's α	Explained Variance (%)
Capabilities (*CAP*)	1.50	7.00	5.23	1.15	0.85	87
HC enables firm to be more efficient	1.00	7.00	5.14	1.22		
HC enables firm to be more effective	1.00	7.00	5.33	1.22		
Firm-specificity (*FS*)	1.75	6.50	3.75	1.05	0.77	60
Knowledge base specific	1.00	7.00	3.68	1.58		
Additional firm-specific training	1.00	7.00	3.87	1.44		
Time to learn firm-specific products/customers	1.00	6.00	3.50	1.21		
Time needed for firm-specific training	2.00	6.00	3.90	1.16		
Causal ambiguity (*CA*)	1.29	5.86	3.81	0.90	0.84	54
Repetitive activities	1.00	7.00	3.87	1.39		
Same tasks daily	1.00	7.00	3.94	1.26		
Nature of job	1.00	7.00	3.95	1.35		
Follow sequence of steps	1.00	7.00	3.51	1.30		
Routineness of work	1.00	6.00	3.92	1.18		
Established procedures/policies	1.00	7.00	3.63	1.19		
Repetitious duties	1.00	6.00	3.88	1.16		
Spread (*SPR*)	2.33	7.00	4.24	1.07	0.79	71
% of workforce strategic human capital	2.00	7.00	3.74	1.49		
Skills found throughout organization	2.00	7.00	4.53	1.17		
Knowledge found throughout organization	2.00	7.00	4.51	1.21		

Note: Shown are the descriptive statistics for the variables that proxy for four characteristics of human capital. Each variable is constructed from the survey questions described in detail in the appendix and is the simple mean of the survey responses to the questions related to each variable. All survey questions are on a scale from 1 to 7.

Variable descriptions: *CAP*, the capability of the human capital; *FS*, the firm-specificity of the human capital; *CA*, the ambiguity of the work performed by human resources; *SPR*, the extent to which human capital is spread throughout the firm.

are easily monitored and evaluated by superiors, and there is an understandable sequence of steps employees follow, then the resources may be easier to imitate. This construct was originally used in Abernethy and Brownell (1997).

Activities that are low in causal ambiguity because they are characterized as repetitious; easily monitored and understood; and subject to standard operating procedures, may still be inimitable if they are embedded in social and/or complementary relations. Thus, we also measure the mass or spread of resources throughout the firm (*SPR*). This is consistent with Barney and Wright (1998) and Dierickx and Cool (1989) who argue that it is difficult for competitors to replicate a competitive advantage when it is embedded in a mass of resources. There is no previously validated measure of *SPR*; therefore, this is a novel measure that uses three questions based on the concepts put forth by Barney and Wright (1998) and Dierickx and Cool (1989) to capture the spread of knowledge and skills throughout the firm. This construct was originally used in Widener (2004).

Survey Procedures and Sample Population

The survey was undertaken during February–April, 1999. Highly diversified firms may pursue multiple strategies and rely on various strategic resources across business units or segments. Since firms reported in Compustat range in size and complexity, we exclude firms that are highly diversified from the sample population. More specifically, we included in the sample population only Compustat firms reporting sales for a single four-digit SIC code or reporting sales for between one and five four-digit SIC codes within the same overall SIC division.[16] To analyze non-response bias and to validate the variable measures, firms are required to report sales from 1993 to 1996 and depreciation expense for 1996, and have at least 250 employees. After deleting firms that were either used in the pretest or that are foreign-owned, the population is 1,662 firms, of which we randomly surveyed 800 firms. The largest concentration of firms (43%) is classified as manufacturing (SIC codes 2000–3999). Other concentrated segments include financial services firms (17%), other service firms (13%), and transportation, communication, and utilities (11%).

To enhance the validity of the survey instrument, we followed the Dillman's (1978) "total design approach." We visited four firms and interviewed various members of management in order to better inform the survey. After designing the survey, we pretested it on 30 firms. Upon final revisions, we sent the survey along with a personalized cover letter and a stamped return envelope to the chief financial officer of 800 firms. We promised to provide the respondents with a summary of results as an incentive to respond. In addition, we performed three follow-up mailings, along with a postcard reminder.

The mailing process resulted in 118 responses (15% response rate), comparable to other survey results for top executives in U.S. firms. We tested for non-response bias to determine potential effects on our findings. We found that the pattern of SIC classifications for respondents mirrors both the sample and the Compustat population. We compared respondents to non-respondents for the Compustat variables. Using return date, we also divided the respondents into early, middle, and late respondents. We then compared early and late respondents for the four human capital variables that are derived from the survey. There are no statistically significant differences among groups of respondents, which provides some comfort regarding the lack of response bias. Non-response tests are reported in Table 2.

Table 2. Investigation of Non-Response Bias[a].

Human Capital Variable	*N*	Mean	Std. Dev.
Panel A: Early respondents[b]			
CAP	34	5.19	0.93
FS	34	3.79	1.03
CA	34	3.63	0.93
SPR	34	4.04	0.95
Panel B: Late respondents[b]			
CAP	28	5.50	1.24
FS	28	3.67	1.25
CA	28	3.81	0.95
SPR	28	4.49	1.15

Note: See Table 1 for a description of the variables.

[a]Response bias was investigated based on comparisons of 1997 sales, depreciation, total assets, and number of employees. Although not reported, differences in means between (a) respondents and non-respondents, (b) early and middle respondents, (c) early and late respondents, and (d) middle and late respondents are not statistically significant on these archival variables. Additionally, there are no significant differences between respondents and the population from which the sample was drawn.

[b]Early respondents are those surveys returned prior to the second mailing. Middle respondents are those surveys returned prior to the third mailing. Late respondents are those surveys returned subsequent to the last mailing.

Final Sample and Descriptive Statistics

We matched the survey data of the 118 respondent firms to Compustat data as of the fiscal year end closest to February–April, 1999. For most firms, this was December 31, 1998. We require that the following items be available in Compustat: price per share (data item #199), net income before extraordinary items (data item #18), and number of common shares outstanding (data item #25). These restrictions result in a final sample of 106 firms.[17]

Table 3, panel A, presents descriptive statistics of the sample firms. The firms have an average market value of equity of \$1.9 billion. The average market price per share is \$19.03. The average number of employees is 10,000. Our sample firms contribute on average \$0.12 in pension expense per share, whereas the 23 firms that disclose salary information spend on average \$8.30 on salary expenditures per share. Although there are several loss firms in the sample, firms on average record \$0.72 per share in profits before extraordinary items. The average market-to-book ratio of 2.33 is consistent with the notion that our sample firms have a number of valued, unrecorded resources. That is, the market values more than just the recorded assets of the firm. This further reinforces the need to examine off-balance sheet items, such as human capital.

As shown in panel B, manufacturing firms dominate our sample (34%); however, we also have a number of financial service firms (20%), wholesale and retail trade (12%), and transportation and utility firms (11%). As discussed earlier, the sample composition broadly mirrors the Compustat population.

ANALYSIS AND RESULTS

Correlation Analysis

Table 4 shows the correlation between the variables. Previously, we argued that three characteristics of strategic human capital can make it difficult for competitors to imitate the firm's use of its strategic resources. It is important to note that, with the exception of the correlation between *CA* and *SPR* ($r = -0.226$, $p < 0.05$), the three dimensions of imitability, *FS*, *CA*, and *SPR*, are not significantly correlated. Both *CAP* and *SPR* are significantly positively correlated with *SUSTADV*. Also of significance is the fact that none of the human capital variables are correlated with the number of

Table 3. Descriptive Statistics.

Panel A: Descriptive statistics

Variable	Mean	Median	Std. Dev.	Min.	Max.
Regression variables					
PRICE (*P*)	19.03	15.88	15.61	0.15	87.44
BV	11.99	8.69	11.32	−3.71	65.33
NI	0.72	0.88	1.60	−6.56	4.31
PENS	0.12	0.04	0.36	−0.03	3.38
SUSTADV	0.36	0	0.48	0	1
Other variables					
Market value of equity (in millions)	1,899	253	5,074	2.81	30,598
Market-to-book	2.33	1.75	2.39	−1.01	15.69
Number of employees (in millions)	0.01	0.001	0.03	0.0001	0.23
Salary expense per share	8.30	2.89	14.92	1.16	71.99
R&D per share	0.29	0	0.65	0	3.62
Net property, plant, and equipment per share	7.66	3.02	12.66	0	108.87

Panel B: Industry Breakdown

1-Digit SIC Code	Industry Description	Number	Percent
1	Metal and construction	6	5.7
2	Food, textile, and chemicals	8	7.5
3	Rubber, metal, and machine products	36	34.0
4	Transportation and utilities	12	11.3
5	Wholesale and retail trade	13	12.3
6	Financial services	21	19.8
7	Hotel and other services	6	5.7
8	Health and other services	4	3.7
	Total	106	100

Note: Shown are descriptive statistics for the entire sample of 106 firms, except for salary expense per share, which is only available for 23 firms.

Variable descriptions: PRICE (*P*), price per share; *BV*, book value of equity adjusted for pension contributions, deflated by shares outstanding; *NI*, net income before extraordinary items, deflated by shares outstanding; *PENS*, pension and retirement cost, after tax and adjusted for non-service cost items, deflated by shares outstanding; *SUSTADV*, composite variable used to proxy for sustained competitive advantage. It is equal to 1 if a firm is above the industry median on *CAP* and above the industry median on either *FS*, *CA*, or *SPR* (see Table 1 for definitions), otherwise it is equal to zero.

Table 4. Pearson Correlation Coefficients (*p*-values in parentheses).

	BV	*NI*	*PENS*	*CAP*	*FS*	*CA*	*SPR*	*SUSTADV*	Number of Employees	Salary Expense	R&D Expense
NI	0.444										
	(0.0001)										
PENS	0.377	0.171									
	(0.0001)	(0.08)									
CAP	0.010	−0.096	0.036								
	(0.92)	(0.33)	(0.71)								
FS	0.223	−0.050	−0.082	0.090							
	(0.02)	(0.61)	(0.40)	(0.36)							
CA	0.046	−0.066	0.204	−0.026	0.032						
	(0.64)	(0.50)	(0.04)	(0.79)	(0.74)						
SPR	−0.036	0.057	−0.126	0.142	−0.099	−0.226					
	(0.71)	(0.56)	(0.20)	(0.15)	(0.31)	(0.02)					
SUSTADV	−0.021	−0.100	0.036	0.676	0.001	0.103	0.218				
	(0.83)	(0.31)	(0.71)	(0.0001)	(0.99)	(0.29)	(0.02)				
Number of employees	0.312	−0.003	0.237	−0.027	−0.031	0.014	0.043	−0.061			
	(0.001)	(0.98)	(0.01)	(0.78)	(0.76)	(0.89)	(0.66)	(0.53)			
Salary expense per share	0.288	−0.101	0.960	0.204	−0.265	0.457	−0.483	0.206	0.199		
	(0.18)	(0.65)	(0.0001)	(0.35)	(0.22)	(0.03)	(0.02)	(0.35)	(0.36)		
R&D expense per share	0.060	−0.163	0.085	−0.001	0.122	0.260	0.009	0.043	0.170	0.060	
	(0.54)	(0.09)	(0.39)	(0.98)	(0.21)	(0.01)	(0.93)	(0.66)	(0.08)	(0.78)	
Net property, plant, and equipment per share	0.371	−0.072	0.222	0.015	0.191	0.001	−0.097	0.148	0.238	0.306	−0.036
	(0.0001)	(0.47)	(0.02)	(0.88)	(0.05)	(0.99)	(0.33)	(0.13)	(0.01)	(0.17)	(0.71)

Note: The correlation coefficients are estimated for the entire sample of 106 firms, except for salary expense per share, which is only available for 23 firms. See Tables 1 and 3 for a description of the variables.

employees. This gives assurance that the survey variables are not simply capturing the size of the workforce. Finally, as expected, pension expense per share is highly correlated with both the number of employees and salary expense per share ($p<0.01$), indicating that this variable is a good proxy for the overall level of spending on employee salary and benefits. We test for and find no evidence of multicollinearity in the final models presented in Tables 5 and 6.[18]

Looking at the individual strategic human capital variables, you can see that *CA* is increasing with respect to two measures of labor intensity (Graham, 2000): pension expense per share ($r = 0.204$, $p<0.05$) and salary per share ($r = 0.457$, $p<0.05$). Thus, firms that are more labor intensive have employees with less transparent jobs and employees who perform ambiguous tasks receive higher employee benefits. *CA* is also positively correlated with the amount spent on R&D per share ($r = 0.260$, $p<0.01$), suggesting that firms in more R&D-intensive industries have more ambiguous tasks. Finally, *FS* is positively correlated with net property, plant, and equipment per share ($r = 0.191$, $p<0.05$), suggesting that firms with a higher proportion of physical assets have more employees with firm-specific skills and knowledge. These correlations are discussed again in the sensitivity analysis section.

Market Valuation of Human Capital

The results for our primary analysis are presented in Table 5. Panel A reports the results for the entire sample. Because the valuation model may not be well specified for certain regulated industries, in panel B we present the results for a sub-sample of 81 non-regulated firms, which excludes insurance, banking, and utility firms. Column 1 of panel A reports the regression results from estimating Eq. (2). The model significantly explains stock price and has an adjusted R^2 of 49% ($F = 15.39$, $p<0.01$). Consistent with prior research, book value and net income are significantly and positively associated with stock price ($p<0.01$).

Examining the individual coefficients in column 1, we find that the associations between price and both the capabilities of human capital and firm-specificity are not significant. This is not unexpected since economic theory predicts that resources will only be valued when they are *both* capable of creating efficiencies and difficult to imitate. However, somewhat unexpectedly we find a significant and positive association between firm value and both causal ambiguity ($p<0.10$) and the spread of strategic

Table 5. Valuation Model Regression Results.

Variable	Predictions	Basic Model	Sustained Advantage Composite Variable	Source of Sustained Advantage			
		(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All firms[a]							
Intercept	Np	−2.74	8.35	−5.89	−3.37	−4.72	−7.79
		(−0.29)	(0.80)	(−0.63)	(−0.35)	(−0.50)	(−0.81)
BV	+	0.52***	0.53***	0.52***	0.53***	0.51***	0.52***
		(4.35)	(4.44)	(4.40)	(4.36)	(4.26)	(4.32)
NI	+	4.34***	4.41***	4.41***	4.35***	4.40***	4.47***
		(5.55)	(5.75)	(5.71)	(5.55)	(5.65)	(5.76)
PENS	+	−2.75	−2.79	−1.96	−2.14	−3.82	−2.38
		(−0.80)	(−0.83)	(−0.57)	(−0.59)	(−1.09)	(−0.66)
CAP	Np	−1.55	−3.47^{+++}	−0.95	−1.50	−1.59^{+}	−1.06
		(−1.59)	(−2.68)	(−0.94)	(−1.53)	(−1.64)	(−1.03)
FS	Np	0.80	0.96	0.50	0.70	0.99	0.61
		(0.71)	(0.86)	(0.44)	(0.61)	(0.87)	(0.53)
CA	Np	2.08^{+}	1.51	2.42^{++}	2.21^{+}	2.19^{+}	2.61^{++}
		(1.64)	(1.19)	(1.92)	(1.71)	(1.74)	(2.03)
SPR	Np	2.32^{++}	1.84^{+}	2.20^{++}	2.35^{++}	2.68^{+++}	2.58^{++}
		(2.19)	(1.73)	(2.10)	(2.20)	(2.46)	(2.35)
SUSTADV	+		6.86**				
			(2.21)				
CAP∗FS	+			1.84**			1.47*
				(1.93)			(1.46)
CAP∗CA	+				−0.68		−0.78
					(0.56)		(−0.61)
CAP∗SPR	+					−1.60	−1.45
						(−1.40)	(−1.17)
Adjusted R^2		0.49	0.51	0.50	0.49	0.49	0.50
Panel B: Firms in non-regulated industries[b]							
Intercept	Np	−6.69	7.87	−11.27	−8.82	−7.37	−12.23
		(−0.68)	(0.77)	(−1.16)	(−0.90)	(−0.74)	(−1.23)
BV	+	0.43***	0.46***	0.43***	0.42***	0.43***	0.43***
		(3.10)	(3.57)	(3.22)	(3.05)	(3.09)	(3.14)
NI	+	2.77***	2.82***	2.83***	2.71***	2.81***	2.79***
		(3.27)	(3.55)	(3.44)	(3.23)	(3.28)	(3.35)
PENS	+	0.44	−0.43	1.57	2.83	−0.12	3.01
		(0.12)	(−0.13)	(0.44)	(0.73)	(−0.03)	(0.75)
CAP	Np	−1.49	−3.98^{+++}	−0.66	−1.42	−1.46	−0.71
		(−1.41)	(−3.22)	(−0.61)	(−1.36)	(−1.37)	(−0.65)
FS	Np	1.68	1.69	1.20	1.58	1.66	1.19
		(1.41)	(1.52)	(1.02)	(1.33)	(1.39)	(1.01)
CA	Np	1.61	0.99	1.99	1.89	1.69	2.15^{+}
		(1.21)	(0.79)	(1.54)	(1.43)	(1.26)	(1.62)

Table 5. *(Continued)*

Variable	Predictions	Basic Model	Sustained Advantage Composite Variable	Source of Sustained Advantage			
		(1)	(2)	(3)	(4)	(5)	(6)
SPR	Np	2.73^{+++} (2.53)	2.02^{++} (1.95)	2.75^{+++} (2.62)	2.88^{+++} (2.68)	2.82^{+++} (2.55)	2.85^{+++} (2.63)
SUSTADV	+		10.24^{***} (3.36)				
*CAP*FS*	+			2.35^{***} (2.36)			2.06^{**} (1.87)
*CAP*CA*	+				−1.98 (−1.59)		−1.36 (−1.04)
*CAP*SPR*	+					−0.57 (−0.43)	−0.06 (−0.05)
Adjusted R^2		0.36	0.44	0.40	0.37	0.35	0.39

Note: Column 1 shows the results from a regression of price per share on book value of equity adjusted for pension contributions (*BV*), net income before extraordinary items (*NI*), pension and retirement cost after tax and adjusted for non-service cost items (*PENS*), and four characteristics of strategic human capital: the capability of the strategic human capital (*CAP*), the firm-specificity of the strategic human capital (*FS*), the ambiguity of the work performed by strategic human resources (*CA*), and the extent to which strategic human capital is spread throughout the firm (*SPR*). Column 2 adds a composite variable to proxy for sustained competitive advantage (*SUSTADV*). Columns 3–6 add interaction terms between *CAP* and the other strategic human capital variables. All financial variables are deflated by shares outstanding and are as of the fiscal year end closest to the survey period of February–April, 1999. For most firms this was December 31, 1998. The characteristics of human capital are constructed from survey data and are described in Table 1.

Np: no prediction.

***, **, *Significant at $p<0.01$, 0.05, 0.10, respectively, using a one-tailed test.

$^{+++}$, $^{++}$, $^{+}$Significant at $p<0.01$, 0.05, 0.10, respectively, using a two-tailed test.

[a]The sample is 106 firms responding to the survey.

[b]The sample is 81 firms in non-regulated industries.

human capital within the firm ($p<0.05$). Firms have a higher market value when the employees perform ambiguous tasks, although the significance level is only marginal. Uncertainty underlying firm resources can lead to excess profits because it is difficult for competitors to identify the firm's strategy and replicate it. This also has the benefit of potentially impeding entry of new firms into the market (Lippman & Rumelt, 1982). This result should be interpreted cautiously as it does not hold for the non-regulated industry sub-sample or in either of the sensitivity tests. In addition, firms

Table 6. Sensitivity Analysis Using Market-Adjusted Stock Returns.

Variable	Predictions	All Firms	Profit Firms	Profit Firms Excluding Regulated Industries
		(1)	(2)	(3)
Intercept	Np	−0.30	0.15	0.24
		(−1.12)	(0.49)	(0.64)
NI	+	0.53**	0.04	−0.48
		(2.31)	(0.06)	(−0.56)
PENS	+	3.44*	4.10**	4.58**
		(1.57)	(1.82)	(1.80)
ΔNI	+	0.03	1.45**	1.65**
		(0.25)	(2.32)	(2.33)
$\Delta PENS$	+	17.75***	17.28**	18.19**
		(2.34)	(1.90)	(1.70)
CAP	Np	−0.01	−0.04	−0.04
		(−0.41)	(−1.20)	(−1.10)
FS	NP	0.01	−0.01	−0.03
		(0.34)	(−0.38)	(−0.69)
CA	Np	−0.01	−0.06	−0.07
		(−0.32)	(−1.50)	(−1.62)
SPR	NP	0.05+	0.03	0.04
		(1.81)	(1.09)	(1.07)
SUSTADV	+	0.11*	0.19**	0.24**
		(1.39)	(2.04)	(2.22)
Adjusted R^2		0.12	0.13	0.17
Number of observations		103	82	58

Note: Shown are the results from a regression of market-adjusted stock returns (*BHAR*) on the level of net income per share before extraordinary items (*NI*), the level of pension and retirement cost per share, after tax and adjusted for non-service cost items (*PENS*), the change in net income per share before extraordinary items (ΔNI), the change in pension and retirement cost per share ($\Delta PENS$), and five characteristics of strategic human capital: the capability of the strategic human capital (*CAP*), the firm-specificity of the strategic human capital (*FS*), the ambiguity of the work performed by strategic human resources (*CA*), the extent to which strategic human capital is spread throughout the firm (*SPR*), and a composite variable to proxy for sustained competitive advantage (*SUSTADV*). All financial variables are deflated by the beginning of the year stock price. The characteristics of human capital are constructed from survey data and are described in Table 1.

Np: no prediction.

***, **, *Significant at $p<0.01$, 0.05, 0.10, respectively, using a one-tailed test.

+Significant at $p<0.10$ using a two-tailed test.

have a higher market value when employees that possess valuable skills and knowledge are spread throughout the organization. When talents and skills are spread over a larger group of employees it is more difficult for competitors to imitate the process and receive the benefits. In addition, if the strategy involves employees throughout the organization, firms are able to spread costly start up expenditures (e.g., selection and staffing costs) necessary to establish a strategic workforce over a greater number of employees (Williamson, 1979). Finally, implementing a strategy that involves many employees may be less risky since the loss of one talented employee may not be as detrimental to the firm's competitive advantage when the pool of talented employees is large. Thus, although the employees do not necessarily rank high in their capability of producing organizational efficiencies, the market still values a large pool of human capital.

Column 2 of Table 5, panel A, reports the regression results from estimating Eq. (3). This is identical to Eq. (2), except that the composite variable capturing sustained competitive advantage is added to the model. The model significantly explains stock price and has an adjusted R^2 of 51% ($F = 14.61$, $p < 0.01$). There is also a positive association between price and the group of human capital variables ($F = 2.75$, $p < 0.05$). When *SUSTADV* is added to the model, *CA* is no longer significant. In addition, we find a negative and significant association between the capabilities of human capital (*CAP*) and firm price ($p < 0.01$); however, the relation between the composite variable (*SUSTADV*) and price is positive and significant ($p < 0.05$). The latter result provides support for Hypothesis 1. Since the composite variable is defined as 1 if a firm is above the median on *CAP* and above the median on *FS*, *CA*, or *SPR*, it is essentially an interaction variable capturing the interaction between capability and imitability.[19] The intuition behind these two results is that the market treats investments in human capital, which is capable of creating organizational efficiencies, similar to other expenses unless the firm takes steps to protect their source of competitive advantage by making the strategic resource difficult to imitate. Relying on human capital to create efficiencies while that resource can be imitated by rival firms increases the risk that the firm will lose its competitive advantage; however, when the human capital has the capability to create efficiencies and is difficult for competitors to imitate, the firm is able to realize a long-term benefit and the market positively values the investment in human capital.

Columns 3–6 of Table 5, panel A, report the regression results from estimating Eqs. (4)–(7). This is identical to Eq. (2), except that interaction terms of *CAP* with *FS*, *CA*, and *SPR* individually are added to the model both separately and at the same time. In all cases, the model significantly explains stock price and has an adjusted R^2 of 49–50% ($p<0.01$). When the interaction terms are added to the model, both *CA* and *SPR* remain positive and significant. In addition, the interaction between *CAP* and *FS* is positive and significant ($p<0.05$), although the interactions of *CAP* with both *CA* and *SPR* are insignificantly different from zero. The market positively values a sustained competitive advantage achieved through firm-specific human capital. This is in line with arguments made by Coff (1997), who argues that firm-specific investments are the best way to reduce worker mobility and reduce the threat of voluntary turnover. We should also note that the significance level decreases to 10% when all of the interaction terms are included in the model at the same time, likely due to power issues with a small sample.

Table 5, panel B, reports the results for the sub-sample of non-regulated firms. The format is the same as panel A. The results are similar to those discussed above for the full sample, except that *CA* is not significant and *SPR* is significant at the 5% or above level in all cases. In addition, *SUSTADV* is significant at the 1% level, and the interaction between *CAP* and *FS* is significant at the 1% level when it is included in the model by itself and 5% when all interaction terms are included.

Sensitivity Analysis – Returns Model

There are two potential limitations associated with the price model used in the prior analysis. First, the choice of variable used to control for heteroscedasticity could lead to spurious scale effects (Brown, Lo, & Lys, 1999). Second, as can be seen in Fig. 1, during this time period stock prices were generally inflated. Therefore, we also examine a returns model. Following Easton and Harris (1991), we relate returns to both the level and change in net income and pension expense per share. Ideally, we would also include both the level and change in the human capital variables, but we do not have survey data available for the prior year. However, we do not expect that the use of strategic human capital would change dramatically from one year to the next and, therefore, including only the level of the human capital variables is a reasonable approach (i.e., the change would be zero and drop out of the model). We use the following model:

$$\begin{aligned} BHAR_i = {} & \beta_0 + \beta_1 NI_i + \beta_2 PENS_i + \beta_3 \Delta NI_i + \beta_4 \Delta PENS_i + \beta_5 CAP_i \\ & + \beta_6 FS_i + \beta_7 CA_i + \beta_8 SPR_i + \beta_9 SUSTADV_i + \varepsilon_i \end{aligned} \tag{8}$$

where NI_i and ΔNI_i are the level and change in net income per share before extraordinary items, deflated by beginning of the year stock price, $PENS_i$ and $\Delta PENS_i$ the level and change in pension and retirement costs per share, after tax, deflated by beginning of the year stock price, and all other independent variables are as described previously. *BHAR* is the 12-month buy-and-hold, market-adjusted stock return.[20] The advantage of this model is that there are no spurious scale effect issues and the impact of inflated stock prices is mitigated by adjusting for the market return. The disadvantage is determining what period to accumulate stock returns over, since the survey was conducted over a several month period after the fiscal year end of most firms. We selected an accumulation period to cover both the dates during which the annual report was released and the dates during which the survey was administered. Specifically, buy-and-hold returns are cumulated over the 12-month period ending with the fifth month after the fiscal year end. For most firms this was June 1998 to May 1999.

The results for the sensitivity analysis are presented in Table 6. Stock return data was obtained from the Center for Research in Security Prices (CRSP) database. The information was not available for three of the sample firms, resulting in a sample size of 103 for this analysis. In addition to running the regression on the entire sample (column 1), we also look at a subset of the sample where loss firms are removed (column 2). Losses are less informative than profits about a firm's future prospects and the stock market reacts differently to firms in a loss position (Hayn, 1995). To ensure that the human capital variables are not proxying for future earnings potential when a firm has a loss, we exclude loss firms from the sample, resulting in a subset of 82 profit firms. Finally, consistent with the previous analysis, we also look at a subset of the sample excluding regulated firms (column 3), resulting in a subset of 58 profit firms in non-regulated industries.

The level of net income is positive and significant ($p<0.05$) for the entire sample (column 1), while the change in net income is positive and significant ($p<0.05$) when the sample is limited to profit firms (columns 2 and 3). This is consistent with Ali and Zarowin (1992), who show that the change in earnings is more appropriate when earnings are permanent, and the level of earnings acts as a proxy for unexpected earnings when earnings are not purely permanent, which would be the case when there are loss firms in the sample. The level and change in pension costs are both positively associated with stock returns for the three sub-samples. While we do not find results for *CAP* that are consistent with those in Table 5, we do find some support that the spread of strategic human capital within the firm ($p < 0.10$ in column 1) is positively associated with market-adjusted stock returns. More importantly, we find an association between the composite variable (*SUSTADV*) ($p < 0.05$ in columns 2 and 3; $p < 0.10$ in column 1) and market-adjusted stock returns, which is robust across all three sub-samples and consistent with the findings presented in Table 5. We conclude that the support for Hypothesis 1 is robust to alternative model specifications (price model versus returns model), as well as different sub-samples (all firms, profit firms and non-regulated firms).

Sensitivity Analysis – Other Firm Resources

As discussed previously, the resource-based view of the firm posits that resources are important to achieve a particular competitive position. Barney (1991) classifies resources into three categories: physical capital, human capital, and organizational capital. A firm can implement a strategy through a single resource or through the use of bundles of resources – such as a particular mix of physical, human, and organizational capital (Barney, 1991). Different resources can have no relation to each other, they may be complements, and they may also be substitutes (Amit & Schoemaker, 1993). For example, a firm may provide firm-specific training on a manufacturing process that is used only by the firm using specialized equipment. In this case, there is a complementary relation between the human capital resource and the physical resource. This will differ for each firm and it is difficult to develop a systematic variable that captures the relationship between similar resources across firms.

Similarly, intellectual capital is often divided into three components: human capital, organizational capital, and customer capital (see, e.g., Mouritsen et al., 2001; Bontis, 1998). Bontis (1998) finds that human capital is associated with firm performance through customer and structural capital, but not directly. Similarly, Bontis and Fitz-enz (2002) find that human capital has an effect on human capital effectiveness (a measure capturing the return on employee expenditures) through its effect on relational capital (similar to customer capital). Bontis posits that this is because employee knowledge must be codified into organizational knowledge to impact performance (Bontis, 1998, p. 71). This idea is very similar to the resource-based view of the firm, which states that human capital must not be imitable to create value and one way to do this is to provide firm-specific skills and knowledge.

The focus of this article is on strategic human capital. Although we have controlled for the book value of assets, future earnings potential as captured by net income, and employee cash expenditures as captured by pension expense, we may have missing variables in the form of physical, structural, and customer capital. Recalling from Table 4 that causal ambiguity is correlated with R&D expenditures and that firm-specificity is correlated with property, plant, and equipment. To ensure that our results are not driven by variables that are missing from the model we perform two sensitivity analyses. First, we include two additional archival variables in the model. The first variable is property, plant, and equipment, which captures physical capital. Since property, plant, and equipment is a part of book value, it was already in the model, although not separately. The second variable is R&D expenditures, which captures a portion of structural capital. In untabulated results, we find that these variables are significant; however, they do not change the statistical inferences drawn above. Second, we include two additional survey variables in the model. The first variable captures the importance of structural capital to the firm and is measured through questions on technology and innovation capital such as patents. The second variable captures the importance of physical capital to the firm and is measured through questions on the firm's fixed assets. In untabulated results, we find that neither of these variables is significant and including them in the model does not change the statistical inferences drawn above. In addition, to capture complementarities between strategic resources, we included interaction terms of *SUSTADV* with both the structural capital and the physical capital survey variables. In untabulated results, we find that the interaction terms are not significant.

CONCLUSION

On the basis of the strategy-based resource framework, we measure four dimensions of human capital – firm-specificity, causal ambiguity, spread, and capability for producing efficiencies – and demonstrate that human capital that is both difficult to imitate and capable of producing efficiencies is positively valued by the market. On the other hand, human capital that is *not* difficult to imitate, but still capable of producing efficiencies is valued as an expense. We also find some evidence that large numbers of human capital are valued positively by the market.

Consistency across countries in accounting standards is becoming more important, yet there is little agreement even on the underlying definition of intangibles, intellectual capital, and more specifically, human capital (Stolowy & Jeny-Cazavan, 2001). Since our research demonstrates that the market values *strategic human capital*, perhaps our findings can inform the debate on the conceptualization and definition of the type of human capital that is important to firms and their stakeholders.

One area where the parameters of human capital visibly matter is in the reporting of intellectual capital. Firms continue to experiment with producing and distributing intellectual capital reports. Johanson et al. (1998, p. 7) discuss the notion of linking measurement and reporting of human capital to the balanced scorecard (BSC) framework, but they state, ''However, before the application of HRCA [human resource cost accounting], BSC or any other model (i.e., before measuring) you have to know what to measure. For instance, which are the important value drivers in the firm?'' Additional research has attempted to develop a structured approach to the reporting of intellectual capital in terms of the content of the report (van der Meer-Kooistra & Zijlstra, 2001). Our study has theoretical implications for this line of research and sheds insights on the type of human capital that could be measured. Our findings provide a starting point for developing metrics that firms can measure and use both internally for decision-making and externally in the reporting of intellectual capital. A valuable extension of this study would be to develop archival measures of these survey constructs for use as a future performance measurement tool and for possible inclusion in an IC report.

In contrast to a structured, quantifiable, IC report, Roslender and Fincham (2001) suggest that the reporting of intellectual capital might provide more useful information if it was fluid, flexible, and perhaps more qualitative in nature. This is consistent with Mouritsen's discourse on the importance of the narrative in IC reports. Our study also informs this approach to the reporting of human capital by providing some structure to guide the narrative. Certainly, our results indicate that the firm would want to make it known the extent to which their human capital is both capable of providing strategic efficiencies and difficult to imitate.

Contrary to Johanson (2002), who concludes that ''capital market actors'' do not value *human capital* indicators, this study documents that the U.S. stock market values *strategic* human capital. This is an important distinction. Johanson (2002, p. 34) points out that, ''financial analysts are not stupid persons." Our results provide support for Johanson's statement and demonstrate that capital market participants are able to parse through the available information and attach a value to strategic human capital that is *both* capable of providing efficiencies and difficult for competitors to imitate. However, this study does not shed insight on how market participants form their valuation. Mouritsen (2003) presents an interesting high-level discussion of the mental processes that market participants must undertake in order to value intangibles and intellectual capital. Mouritsen (2003, p. 27) states, "... a valuable avenue for research is exactly to develop more of an appreciation of how it is that capital market participants become knowledgeable about the projects they investigate. Such a process view may add significantly to our understanding of the role of intangibles and intellectual capital in valuing firms and projects." Since we document that strategic human capital is valued by market participants, future research could investigate the process by which it is valued. Case studies in this area could be particularly illuminating.

This study offers insights about a critical intangible asset that is increasingly important in today's organizations. However, consistent with most empirical studies, this one does have its limitations. Although we ran numerous sensitivity and robustness tests, and incorporate several different model specifications, we have a sample of 106 firms, which results in low power. Future studies could expand on the sample size to provide more definitive support of the findings and an extension of this study that incorporates firm data, such as compensation expense. We also only have data at a point in time. Tracking the use of strategic resources over time would provide better information with which to test associations with future earnings, and allow more informative inferences regarding the timing of the future benefits and the strength of the associations. We also use survey measures to construct our independent measures, which may contain noise. This study examines the value relevance of quantitative non-financial data. While this is an important and innovative extension of the current valuation literature, to the extent future studies could capture the four characteristics of strategic human capital in financial terms, corroborating evidence would be obtained. Finally, we limit our investigation to one type of strategic resource because we are interested in understanding how the market values various components of strategic human capital. Recalling Barney's (1991) statement that not all resources are *strategic* resources, future research could incorporate our findings with that of Bontis (1998) and investigate how a broader set of *strategic* resources (e.g., physical, organizational, and human capital) are bundled to create strategic capabilities, and how these strategic capabilities are then valued by the market.

NOTES

1. Studies have also provided systematic evidence that certain intangible resources, such as R&D expenditures and patents, are valued by the market (e.g., Lev, 2001; Lev & Sougiannis, 1996).

2. We use the terms "resources" from the resource-based strategy literature and "production factors" from the economics literature interchangeably. Resources include organizational routines, physical assets, and human capital. Production factors include land, labor, and capital goods. One element common to both is labor, or human capital, which is the focus of this article.

3. Barney (1991) classifies strategic resources into three categories: physical capital, human capital, and organizational capital. We discuss this further in the sensitivity analysis section.

4. For a thorough discussion of intellectual capital, see Mouritsen et al. (2001).

5. For a thorough discussion of HRA, see Johanson et al. (1998).

6. They are recognition and measurement, reporting, evaluation, attention, motivation, commitment, and follow up (Johanson et al., 2001).

7. Also, see Becker and Huselid (1998) for a review of literature investigating the association between human resource practices and firm performance.

8. The strategy-based resource literature has an interesting connection to the early measurement models proposed in HRA. For example, Carper and Posey (1976) model the likelihood of employee mobility when assessing the value of the *CPA* employees. Strategy-based resource is more encompassing since it considers imitability and not only mobility.

9. Another variable that may cause human capital to be difficult to imitate is the presence of stock options. Some argue that stock options are used to retain key employees. In fact, Ittner, Lambert, and Larcker (2003) find that firms in their sample rank retention as the most important objective of option grants.But, the empirical evidence in Ittner et al. (2003) suggests that the importance of the retention objective is only for new employee option grants and not for ongoing option grants. This suggests that options do not necessarily bind an employee to a particular firm; rather options may initially attract a new employee. Many firms use options as a compensation tool, so any firm could use options to attract a new employee. Therefore, we do not include stock options as an imitability variable. Instead we rely on the strategy and economics literature, which provides us with three characteristics of imitability that we measure through a survey instrument.

10. It is important to note that this does not imply that the employees are not performing appropriately or performing services that are not required nor expected. It simply means that the value-added component of their work is not very transparent. For example, consider a manager who makes good, value-added decisions. The thought process is invisible and hard to detect by a superior or by an outsider. Documenting that decision-making process in order to make it transparent could be quite difficult.

11. Ford Motor Company is an example where talented human capital is spread throughout the firm, and competitive advantage is sustained through mass employee involvement and participation. There are circumstances when a valuable human talent may be centered on one key executive, such as Bill Gates. However, the more the use of strategic human capital is spread throughout the ranks of the firms' employees, the harder it will be for a competitor to determine how the competitive advantage is being achieved. Thus, the spread of the resource can foster the lack of imitability.

12. All variables are deflated by shares outstanding to control for heteroscedasticity. The choice of variable used to control for heteroscedasticity could lead to spurious scale effects (Brown et al., 1999). As a sensitivity test, we also deflate all of the financial variables by book value of common equity and the results are qualitatively the same.

13. We also run the analyses without adjusting for the pension contributions. The statistical inferences using unadjusted book value are unchanged.

14. For example, if the industry median (and firm score) for firm A was 5.5 (5.7) on *CAP*, 5.8 (5.6) on *FS*, 4.5 (4.4) on *CA*, and 5.0 (5.6) on *SPR*, the firm would be coded a "1" on *SUSTADV* since it scored over the industry median on *CAP* and at least one of the lack of imitability variables. If its score was 5.5 (5.7) on *CAP*, 5.8 (5.6) on *FS*, 4.5 (4.4) on *CA*, and 5.0 (4.9) on *SPR*, the firm would be coded as a "0" on *SUSTADV* since it did not score above the industry median on at least one of the lack of imitability variables, even though it scored above the industry median on *CAP*. Finally, if its score was 5.5 (5.2) on *CAP*, 5.8 (5.9) on *FS*, 4.5 (4.4) on *CA*, and 5.0 (5.2) on *SPR*, the firm would be coded as a "0" on *SUSTADV* since it did not score over the industry median on *CAP*, even though it scored above the industry median on two of the lack of imitability variables.

15. These constructs were validated in a prior study on strategic human capital (see Widener, 2004). For this study, we remove one question from *CAP* in order to better align the construct measurement with the theoretical definition of "effectiveness and efficiency" that we use in this study. Since there are different data requirements in this study as compared to Widener (2004), the final set of firms differs between the two papers. However, construct validity is reliable across samples. Moreover, the purpose of this study, which has an external focus and relies heavily on Compustat and CRSP data, differs from Widener (2004), which investigated the management control system within the firm.

16. For example, Ampco-Pittsburgh reports sales in three SIC codes: 3452, 3462, and 3561. All three of these codes are classified as Division D (manufacturing) and either, Major Group 34 (fabricated metal products), or Major Group 35 (industrial and commercial machinery and computer equipment). Thus, Ampco-Pittsburgh is included as part of the study's population. Analysis of the Compustat segment data reveals that our sample firms reported an average (median) of 1.76 (1) segments. We also note that the use of R&D expenditures (i.e., a type of structural capital) and the use of seasonal/part time employees (i.e., a type of human capital) is fairly consistent across segments within a firm, indicating that the use of strategic resources within the firm is likely consistent across segments. To the extent that our selection technique does not remove highly diversified firms, our measures will be subject to additional noise, which will bias against finding results.

17. One firm has negative book value of equity. In an unreported sensitivity test, we delete this observation and our results are qualitatively similar.

18. We calculate both the variance inflation factor (VIF) and the tolerance (TOL) for each coefficient. All coefficients are within the recommended guidelines of a maximum of 10 for the VIF and a minimum of 0.10 for the TOL (see Hair, Anderson, Tatham, & Black, 1995, p. 127).

19. This variable captures the high-high cell versus all other cells for the interaction of capability and imitability.

20. We also calculated *BHAR* in two other ways. First, using raw stock returns, we find qualitatively similar results. Second, to control for differences in size and growth opportunities in the sample firms, we also estimated *BHAR* as the 12-month buy-and-hold, book-to-market, and size adjusted stock return (Barber & Lyon, 1997; Fama & French, 1992). Following Chan, Lakonishok, and Sougiannis (2001), each sample firm is assigned a control portfolio at the beginning of the return accumulation period. There are 30 control portfolios corresponding to five possible ranks of book-to-market and six possible ranks of size. Following Chan et al. (2001), the breakpoints for the size portfolio are based on NYSE issues only. The first four groups correspond to the largest four quintiles, groups five and six are formed by splitting the lowest quintile in half, reflecting the fact that many NYSE firms are larger than the entire CRSP population. The book-to-market breakpoints are based on the book-to-market ratio at the end of the most recent fiscal year for all stocks included in the Compustat database. Once the control portfolio breakpoints were established, the portfolios were formed using all stocks in the CRSP database. The results using this measure are qualitatively similar.

ACKNOWLEDGMENTS

Sally K. Widener would like to thank her dissertation Chairman, Frank H. Selto and members of her committee. The support of the University of Colorado at Boulder is gratefully acknowledged. We would like to thank Dick Dietrich, Elizabeth Demers, Mark Kohlbeck, Robert Lipe, Glenn Pfeiffer, Kirk Philipich, Steve Rock, Phil Shane, Tom Turk, Wanda Wallace, Terry Warfield, and workshop participants at Chapman University, the University of Colorado at Boulder, the College of William and Mary, and the University of Wisconsin.

REFERENCES

Abdel-khalik, A. R. (2003). Self-sorting, incentive compensation and human-capital assets. *European Accounting Review*, *12*, 661–697.

Abernethy, M. A., & Brownell, P. (1997). Management control systems in research and development organizations: The role of accounting, behavior and personnel controls. *Accounting, Organizations and Society*, *22*, 233–248.

Ali, A., & Zarowin, P. (1992). The role of earnings levels in annual earnings-returns studies. *Journal of Accounting Research*, *30*(Autumn), 286–296.

Amir, E., & Livne, G. (2005). Accounting, valuation and duration of football player contracts. *Journal of Business, Finance and Accounting*, *32*, 549–586.

Amit, R., & Schoemaker, P. J. H. (1993). Strategic assets and organizational rent. *Strategic Management Journal*, *14*, 33–46.

Anderson, S. W., Davis, G., & Widener, S. K. (2005). *Operational performance as a driver of customer satisfaction and dissatisfaction in the U.S. airline industry*. Working Paper. Rice University.

Ballester, M., Garcia-Ayuso, M., & Livnat, J. (2003). The economic value of the R&D intangible asset. *European Accounting Review*, *12*, 605–633.

Ballester, M., Livnat, J., & Sinha, N. (2002). Labor costs and investments in human capital. *Journal of Accounting, Auditing, and Finance*, *17*, 351–373.

Ballou, B., Godwin, N. H., & Shortridge, R. T. (2003). Firm value and employee attitudes on workplace quality. *Accounting Horizons*, *17*, 329–341.

Barber, B. M., & Lyon, J. D. (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics*, *43*, 341–372.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*, 99–120.

Barney, J. B., & Wright, P. M. (1998). On becoming a strategic partner: The role of human resources in gaining competitive advantage. *Human Resource Management*, *37*, 31–46.

Barth, M. (1991). Relative measurement errors among alternative pension asset and liability measures. *The Accounting Review*, *66*, 433–463.

Barth, M., Beaver, W. H., & Landsman, W. R. (1998). Relative valuation roles of equity book value and net income as a function of financial health. *Journal of Accounting and Economics*, *25*, 1–34.

Becker, B. E., & Huselid, M. A. (1998). High performance work systems and firm performance: A synthesis of research and managerial implications. In: G. Ferris (Ed.), *Research in personnel and human resources*. Stamford, CT: JAI Press Inc.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *The Journal of Political Economy*, *70*, 9–49.

Bontis, N. (1998). Intellectual capital: An exploratory study that develops measures and models. *Management Decision*, *36*, 63–76.

Bontis, N., & Fitz-enz, J. (2002). Intellectual capital ROI: A causal map of human capital antecedents and consequents. *Journal of Intellectual Capital*, *3*, 223–247.

Brown, S., Lo, K., & Lys, T. (1999). Use of R^2 in accounting research: Measuring changes in value relevance over the last four decades. *Journal of Accounting and Economics*, *28*, 83–115.

Carper, W. B., & Posey, J. M. (1976). The valuation of selected surrogate measures of human resource value: A field study. *Accounting, Organizations and Society*, *1*, 143–151.

Chan, L., Lakonishok, J., & Sougiannis, T. (2001). The stock market valuation of research and development expenditures. *Journal of Finance*, *56*, 2431–2456.

Chang, L., & Birkett, B. (2004). Managing intellectual capital in a professional service firm: Exploring the creativity-productivity paradox. *Management Accounting Research*, *15*, 7–31.

Coff, R. W. (1997). Human assets and management dilemmas: Coping with hazards on the road to resource-based theory. *The Academy of Management Review*, *22*, 374–402.

Dierickx, I., & Cool, K. (1989). Asset stock accumulation and sustainability of competitive advantage. *Management Science*, *35*, 1504–1511.

Dillman, D. A. (1978). *Mail and telephone surveys: The total design method*. Wiley: New York.

Easton, P. D., & Harris, T. (1991). Earnings as an explanatory variable for returns. *Journal of Accounting Research*, *29*, 19–36.

Edwards, S. (1997). The brain gain. *CA Magazine*, April, pp. 21–25.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, *47*, 427–465.

Financial Accounting Standards Board (FASB). (2001). *Improving business reporting: Insights into enhancing voluntary disclosures*. Stamford, CT: Financial Accounting Standards Board.

Flamholtz, E. G. (1971). A model for human resource valuation: A stochastic process with service rewards. *The Accounting Review*, *46*, 253–267.

Flamholtz, E. G. (1985). *Human resource accounting*. San Francisco, CA: Josey Bass Inc.

Flamholtz, E. G. (1987). Valuation of human assets in a securities brokerage firm: An empirical study. *Accounting, Organizations and Society*, *12*, 309–318.

Graham, D. J. (2000). Spatial variation in labour productivity in British manufacturing. *International Review of Applied Economics*, *14*, 323–341.

Grant, R. M. (1991). The resource-based theory of competitive advantage: Implications for strategy formulation. *California Management Review*, *33*, 114–135.

Green, J. P., Stark, A. W., & Thomas, H. M. (1996). UK evidence on the market valuation of research and development expenditures. *Journal of Business Finance and Accounting*, *23*, 191–216.

Grojer, J. (2001). Intangibles and accounting classifications: In search of a classification strategy. *Accounting, Organizations and Society*, *26*, 695–713.

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis*. Englewood Cliffs, NJ: Prentice Hall.

Hayn, C. (1995). The information content of losses. *Journal of Accounting and Economics*, *20*, 125–153.

Heskett, J. L., Sasser, E., & Schlesinger, L. A. (1997). *The service profit chain*. New York, NY: The Free Press.

IFAC Report. (1998). *The measurement and management of intellectual capital: An introduction*. New York: International Federation of Accountants.

Ittner, C. D., Lambert, R. A., & Larcker, D. F. (2003). The structure and performance consequences of equity grants to employees of new economy firms. *Journal of Accounting and Economics*, *34*, 89–127.

Ittner, C. D., & Larcker, D. F. (1998). Are non-financial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research*, *36*(Suppl.), 1–35.

Johanson, U. (2002). Why are capital market actors ambivalent to information about certain indicators on intellectual capital? *Accounting, Auditing and Accountability Journal*, *16*, 31–38.

Johanson, U., Eklov, G., Holmgren, M., & Martensson, M. (1998). *Human resource costing and accounting versus the balanced scorecard: A literature survey of experience with the concepts*. MERITUM project: Stockholm University.

Johanson, U., Martensson, M., & Skoog, M. (2001). Mobilizing change through the management control of intangibles. *Accounting, Organizations and Society*, *26*, 715–733.

Leitner, K.-H., & Warden, C. (2004). Managing and reporting knowledge-based resources and processes in research organizations: Specifics, lessons learned and perspectives. *Management Accounting Research*, *15*, 35–51.

Lev, B. (2001). *Intangibles management, measurement, and reporting*. Washington, DC: Brookings Institution Press.

Lev, B., & Schwartz, A. (1971). On the use of the economic concept of human capital in financial statements. *The Accounting Review*, *46*, 103–112.

Lev, B., & Sougiannis, T. (1996). The capitalization, amortization and value-relevance of R&D. *Journal of Accounting and Economics*, *21*, 107–138.

Lev, B., & Zambon, S. (2003). Intangibles and intellectual capital: An introduction to a special issue. *European Accounting Review*, *12*, 597–603.

Lippman, S. A., & Rumelt, R. P. (1982). Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *The Bell Journal of Economics*, 418–438.

Lohtia, R., Brooks, C. M., & Krapfel, R. E. (1994). What constitutes a transaction-specific asset? An examination of the dimensions and types. *Journal of Business Research*, *30*, 261–270.

Mouritsen, J. (1998). Driving growth: Economic value added versus intellectual capital. *Management Accounting Research*, *9*, 461–482.

Mouritsen, J. (2003). Intellectual capital and the capital market: The circulability of intellectual capital. *Accounting, Auditing and Accountability Journal*, *16*, 18–30.

Mouritsen, J., Larsen, H. T., & Bukh, P. N. D. (2001). Intellectual capital and the 'capable firm': Narrating, visualizing and numbering for managing knowledge. *Accounting, Organizations and Society*, *26*, 735–762.

Nunnally, J. C. (1978). *Psychometric theory* (2nd edn.). New Delhi: Tata McGraw-Hill.

Roslender, R., & Fincham, R. (2001). Thinking critically about intellectual capital accounting. *Accounting, Auditing and Accountability Journal*, *14*(4), 383–398.

Skandia. (1998). *Human capital in transformation: intellectual capital prototype report*. Stockholm, Sweden: Skandia Corporation.

Snell, S. A., & Dean, J. W. Jr. (1992). Integrated manufacturing and human resource management: A human capital perspective. *Academy of Management Journal*, *35*, 467–504.

Stolowy, H., & Jeny-Cazavan, A. (2001). International accounting disharmony: The case of intangibles. *Accounting, Auditing and Accountability Journal*, *14*, 477–496.

Van der Meer-Kooistra, J., & Zijlstra, S. M. (2001). Reporting on intellectual capital. *Accounting, Auditing and Accountability Journal*, *14*, 456–476.

Watson Wyatt. (1999). The human capital index™: Linking human capital and shareholder value. Survey Report. Watson Wyatt Worldwide.

Widener, S. K. (2004). An empirical investigation of the relation between the strategic use of human capital and the design of the management control system. *Accounting, Organizations and Society*, *12*, 377–399.

Williamson, O. E. (1979). Transaction-cost economics: The governance of contractual relations. *The Journal of Law and Economics*, 233–261.

APPENDIX. SURVEY QUESTIONS

Capability	Our strategic human capital resources enable the firm to be more efficient. (1 = strongly agree, 7 = strongly disagree)
	Our strategic human capital resources enable the firm to be more effective in exploiting opportunities. (1 = strongly agree, 7 = strongly disagree)
Firm-Specific	Is the knowledge base held by your firm's strategic human capital primarily specific to your organization? (1 = knowledge base is readily transferable to other firms, 7 = knowledge base is specific to this firm)
	Indicate the extent of your agreement with the following statement: It would be easy for an experienced employee to come into your organization and contribute as part of your firm's strategic human capital without any additional firm-specific training. (1 = strongly agree, 7 = strongly disagree)
	How much time is required for a newly hired employee with experience in the industry to become familiar with firm-specific knowledge of your products and customers in order to contribute as strategic human capital? (1 = little time, 7 = much time)
	On average, how much time would it take for a replacement employee to learn the firm-specific tasks necessary to be as effective as a current employee that contributes to your firm's strategic human capital? (1 = little time, 7 = much time)
Causal Ambiguity	To what extent would you say your firm's strategic human capital perform repetitive activities? (1 = all the time, 7 = very little)
	To what extent are the tasks performed by your firm's strategic human capital the same from day to day? (1 = tasks remain the same, 7 = tasks change daily)
	Does the firm's strategic human capital perform about the same job in the same way most of the time? (1 = methods used constantly change, 7 = methods used are very stable)
	To what extent is there an understandable sequence of steps that can be followed by the firm's strategic human capital in performing tasks? (1 = sequence of steps is very understandable, 7 = sequence is not well understood)
	To what extent would you say the work of your firm's strategic human capital is routine? (1 = very routine tasks, 7 = very non-routine tasks)
	To do the work of your organization, to what extent can your firm's strategic human capital actually rely on established procedures and practices? (1 = procedures exist for all tasks, 7 = there are few procedures)

APPENDIX. (*Continued*)

	How repetitious are the duties performed by your firm's strategic human capital? (1 = duties are very repetitious, 7 = duties are very unique)
Spread	Approximately what proportion of your organization's employees would you consider to be strategic human capital (i.e., those employees critical to sustaining your firm's competitive advantage)?
	0% 1–20% 21–40% 41–60% 61–80% 81–99% 100% 1 2 3 4 5 6 7
	Are the skills used by the strategic human capital group found throughout the organization? (1 = skills are localized in one employee, 7 = skills are spread throughout organization)
	Is the knowledge possessed by the strategic human capital group found throughout the organization? (1 = knowledge is localized in one key employee, 7 = knowledge is spread throughout organization)

MEASURING PERFORMANCE OF IT INVESTMENTS: IMPLEMENTING THE IT CONTRIBUTION MODEL

Marc J. Epstein and Adriana Rejc Buhovac

ABSTRACT

The pressure to remain competitive in a dynamic, global economy forces organizations to consider the results-based approach when deciding on investments in information technology (IT). Senior IT managers are convinced that they do create value and believe that if measured properly and with adequate support, they would be significant profit centers for their organizations. However, without adequate performance evaluation systems they have difficulties proving the value-adding role of IT and find themselves continually fighting for and justifying the resources that are needed. The article provides a model and a methodology for evaluating performance in IT to help chief information officers (CIOs) better justify and evaluate their initiatives and aid CEOs and CFOs in making better resource allocation decisions. The IT Contribution Model *and the subsequent IT Payoff Methodology is illustrated by and empirically tested in Istrabenz Group, an international group engaged in food, investments, tourism, and energy. The study shows that the methodology's requirement for active employee involvement in the identification of the*

Advances in Management Accounting, Volume 17, 43–79
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017004

critical drivers of success, the expected outputs of the IT initiative, in particular, substantially facilitates the IT initiative implementation by increasing the level of understanding and acceptance.

INTRODUCTION

There have been significant discussions in both the managerial and academic literature concerning the payoffs of information technology (IT) investments. Many senior business managers have questioned IT's contribution to their bottom line (Leavitt, 1999; Schwartz, 1999; Carr, 2003, 2004). Indeed, cumulative results from the earlier studies, which examined the relationship between IT investment and firm performance, along with economics-based studies investigating IT productivity, were plagued with ambiguities and inconsistencies (Strassman, 1990; Weill, 1992; Loveman, 1994). Recent studies, however, examining the value of IT investment in two research streams, one using production economics and the other focusing on 'process-oriented' models, have been more encouraging (Barua & Mukhopadhyay, 2000). Research in both streams managed to mitigate the earlier skepticism on the IT payoffs (Barua & Lee, 1997; Mukhopadyay, Kekre, & Kalathur, 1995). In production economics, Brynjolfsson and Hitt (1996) reported positive returns on IT investment. Aral, Brynjolfsson, and Wu (2006) find evidence that the use of ERP causes performance increases rather than performance inspiring ERP purchases. But also, that success with ERP encourages adoption of extended enterprise systems, which in turn improve productivity and operational performance. Examples of similar positive results of process-oriented studies include Davies, Dehning, and Stratopoulous (2003), Love and Irani (2004), and Lee (2001). Lim, Richardson, and Roberts (2004) posit that contextual factors moderate the relationship between IT investment and firm performance. Byrd, Lewis, and Bryan (2006) indicate that there is a synergistic coupling between strategic alignment and IT investment with firm performance. Business process redesign and human capital also influence the impact of IT investment on firm performance (Davern & Kaufman, 2000). Brynjolfsson (2003), similarly, finds that the greatest IT benefits are realized when an IT investment is coupled with a specific set of complementary business investments.

Despite empirical evidence on tangible outcomes of investing in IT, so far, there has been little guidance of how to design or implement an appropriate IT performance evaluation system. On the one hand, there was a shortage of relevant metrics. On the other hand, even approaches such as the balanced scorecard and shareholder value analysis that do provide overall frameworks for analysis and management, need additional specificity and definition. Increased specificity was necessary to model, measure, and manage the organizational links that operationalize these approaches. Therefore, even financial managers that have expertise in management control and performance measurement have not focused on the benefits of IT and have not developed the appropriate measures. Consequently, the payoffs of IT are not measured, return on investment (ROI) is not calculated, and IT investments are not evaluated with the same rigor as other corporate investments. Furthermore, CEOs and CFOs lacked information to make well-informed decisions on the payoffs of these investments and, as a consequence, corporate goals seem to focus on reduction of the costs of IT rather than maximizing the IT value creation activities.

As IT managers must show the payoffs of IT investment to convince key executives that they should be strong supporters of IT efforts, a framework for evaluation of IT performance is a significant need. Few things are more convincing to top executives than measurable results. We provide a model and a methodology for evaluating performance in IT in both for-profit and non-profit organizations to help CIOs better justify and evaluate their initiatives and aid CEOs and CFOs in making better resource allocation decisions. More specifically, we develop a model of key factors for organizational success in IT integration (*IT Contribution Model*) that includes four dimensions: the critical inputs and processes that lead to success in IT outputs and ultimately to overall organizational success (outcome). The methodology further articulates each of the key factors (antecedents and consequences of IT success) as objectives and outlines the specific drivers of IT success based on these objectives. It identifies the causal relationships between the drivers and develops performance measures for improved management control. Finally, it provides the IT ROI calculation formula following the cause-and-effect relationships between the drivers of IT success. The metrics can be used for both IT project's justification prior to its start (planning) as well as for evaluation after completion (performance measurement).

This comprehensive albeit pragmatic methodology is empirically tested in Istrabenz Group, a four-division holding company. The methodology was applied in the Tourism division, which was facing the challenge of justifying the introduction of a uniform information system for supporting the operation of all the hotels in the division. The selected case provides a suitable empirical context for testing since in 2005, the company adopted guidelines on the use of information and communication technology (ICT), from which it follows that this area is one of the key factors of the Istrabenz Group for achieving its strategic business goals. The company leaders recognized the strategic role IT integration can play in the strategy implementation process, but required IT investments to be evaluated with the same rigor as other corporate investments.

The article has both academic and managerial implications. From the academic perspective, the presented model and the methodology make a twofold contribution. Firstly, the model builds on the process-oriented studies examining the value of IT investment; it upgrades the existing literature by offering an integrated model of critical drivers of IT success. Secondly, the methodology represents a more complete analytical tool for evaluating the payoffs of investing in IT based on the proposed model. The methodology includes the identification of the antecedents and consequences of IT investments, develops the cause-and-effect relationships between the drivers and outcomes, helps identify and measure marginal costs and benefits of the IT initiative to calculate the IT ROI, and provides performance measures for managerial control of the IT initiative.

From the practical perspective, with the *IT Contribution Model*, managers can implement a performance measurement system to more effectively evaluate the effectiveness of IT investments, which can lead to dramatic improvements in decision-making, corporate resource allocations, and performance. More specifically, the new methodology will help the accounting and finance professionals that deal with the challenges of performance measurement and control in IT. CIOs, CTOs, and senior IT managers will better understand how IT contributes to higher levels of corporate performance, more easily evaluate the profitability of IT investments, and make better resource allocation decisions. CEOs, CFOs, and other decision makers will be able to identify, document, measure, and communicate the short-term results and long-term impacts of IT investments. This includes both cost savings and value creation, and thus provides arguments for additional IT resources when appropriate.

The article is organized as follows. The first section provides a review of the existing methodologies used to measure performance of IT investments. In the second section, we describe the *IT Contribution Model* and the methodology to calculate IT payoff. We apply the model to the case of Istrabenz Group, the Tourism division, in the third section in an empirical test of the proposed model. Finally, we discuss the practical implications of the proposed model and the empirical testing, address the study's limitations, and point to some critical performance measurement implementation issues.

AN OVERVIEW OF THE RECENTLY DEVELOPED *IT* PERFORMANCE MEASUREMENT METHODOLOGIES

With CEOs and CFOs demanding accountability for the tremendous investment in IT, IT managers are required to ensure accountability, calculate the return on investment, develop a value-added approach, and make a bottom-line contribution. Generally, however, there has been little guidance of how to design or implement an appropriate IT performance evaluation system, i.e., how to identify and document the contribution of IT to high-performance organizations.

Total Cost of Ownership (TCO) analysis (Gartner Group, 1997), sometimes referred to as total cost of operation, ideally offers a final statement reflecting not only the cost of purchase of software and hardware but all aspects in the further use and maintenance of the equipment, device, or system considered. This includes the costs of training support personnel and the users of the system, costs associated with failure or outage (planned and unplanned), diminished performance incidents (i.e., if users are kept waiting), costs of security breaches (in loss of reputation and recovery costs), costs of disaster preparedness and recovery, floor space, electricity, development expenses, testing infrastructure and expenses, quality assurance, boot image control, marginal incremental growth, decommissioning, e-waste handling, and more. When incorporated in any financial benefit analysis (e.g., ROI, internal rate of return (IRR), EVA), TCO provides an excellent cost basis for determining the economic value of that investment. However, it is insufficient as it does not address the benefits arising from an IT investment. Also, most of what is measured in the TCO analysis is owned by the IT organization while real business benefits can only be determined and owned by the responsible parts of the organization.

Strassman developed a ratio called *Information Productivity* (IP), which is the ratio of the Economic Value-Added (EVA) to the total cost of information management (Strassman, 1999). With IT, being one of the fastest growing components of the costs of information management, this metric is designed to reflect an organization's success at converting the costs of information management into profit. As such, this approach cannot be used for determining an IT initiative payoff.

Another proposal is to expand conventional financial measurement like return on investment and payback period to an eBusiness context, which is a whole-view measurement of business performance across both internal and external constituents (Cameron, Meringer, Dawe, & Jastrzembski, 2000).

By setting weighted eBusiness objectives relating to end-customer success, hyper-partnering efficiency, and multi-organization financial performance and applying quantitative and qualitative impact metrics, organizations can track a project's impact on a given eBusiness objective.

In yet another approach, Intel has developed a *Business Value Index* (BVI) (Intel, 2003; Curley, 2004). BVI is a component index of factors that affect the value of an IT investment. It evaluates IT investments along three vectors: IT business value, impact on the IT efficiency, and the financial attractiveness of the investment. All three vectors use a predetermined set of defining criteria that includes customer need, business and technical risks, strategic fit, revenue potential, level of required investment, and the amount of innovation and learning generated. Each criterion is weighted, and project managers or program owners score their projects against these criteria to produce total scores for each of the three vectors. By graphically depicting the three indices for each project, BVI methodology provides some decision support to managers to compare and contrast investments, and then determine the investments that align best with their business priorities.

Enterprise Resource Payback (IFS Resource Payback) (EAC Report, 2005) is considered a more complete analysis of an IT investment return that the ROI as it looks at the overall payback that enterprise software can offer to a company. It includes not only quantifiable improvements in bottom and top line functionality, but also more qualitative measures – such as new business opportunities, new customer and partner relations, and improved time to market – that contribute significantly to the success of a company's enterprise software implementation and use. Increased quest for accountability in IT, however, demands measurement rather than assessments and assigning monetary value to IT outputs.

IT Value Mapping (Hajela, 2005) is considered a holistic framework that quantifies and visually depicts it capabilities of an organization. It creates diagrams, or value maps, to depict the state of key business and IT components at any given point in time. It also depicts the impact of each component's 'state' on business value. This approach is used to maximize returns on IT investments by eliminating IT investments that are not in line with business imperatives (as seen from the value maps depicting organizational and IT capabilities) and managing the remaining investments to improve returns. It is not used for calculating IT returns per se.

Total Value of Opportunity Approach (Apfel, 2002) is a methodology that measures business performance of an IT initiative by including the important factors of risk, time, and an assessment of the organization's ability to convert projected value into actual business benefit. The methodology is based on the cost/benefit analysis where the costs are done on the basis of the TCO principles, whereas benefits are modeled against all of the controllable activities of the company. The metrics are monitored before, during, and after implementation to determine how the projected value is being delivered.

The so called emerging IT valuation measures also include applied information economics that uses scientific and mathematical methods to evaluate the IT investment process, EVA, economic value sourced that quantifies the dollar value of risk and time and adds these in the valuation equation, portfolio management that manages IT assets from an investment perspective by calculating risks, yields, and benefits, and real option valuation that tracks 'assets in place' and 'growth options' to present the widest array of future possibilities (Davies et al., 2003). Not only are they difficult to apply, they also fail to shed light on how the IT value is generated.

Other approaches can be found in Tardugno, DiPasquale, and Matthews (2000), Remenyi, Money, and Sherwood-Smith (2000), Murphy (2002), Devaraj and Kohli (2002), Lutchen (2004), Weill and Ross (2004), and Schubert (2004).

Though all of these approaches are helpful, they have critical limitations as discussed above. Various approaches and methodologies fall short on providing information on how to make better IT decisions based upon the analysis. Also, in many organizations, after the business initiative was launched, the project was not monitored or benchmarked against the original projected benefits. Performance measures were not specified for subsequent managerial control. Specific tools for the identification and measurement are necessary. In this article, we attempt to provide a useful model and a methodology that will help organizations measure an IT initiative's payoff in a more comprehensive way and execute efficient management control.

THE IT CONTRIBUTION MODEL AND THE IT PAYOFF METHODOLOGY

To properly assess the payoffs of investments in IT, organizations must implement comprehensive systems to evaluate impacts of IT initiatives on financial performance. In Fig. 1, we provide the *IT Contribution Model*

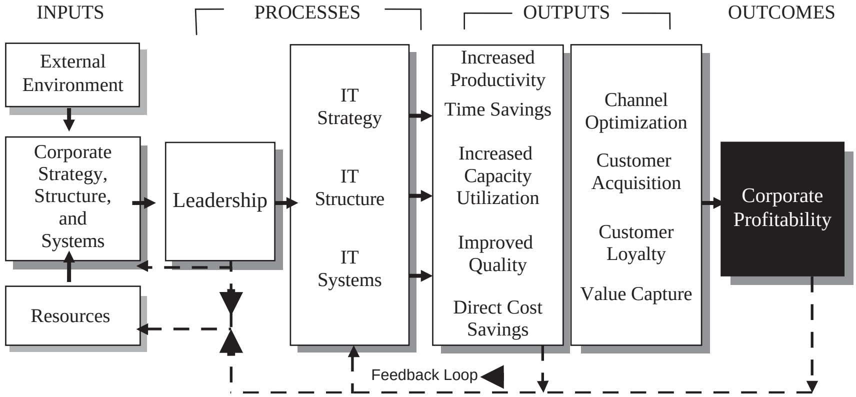


Fig. 1. *IT Contribution Model*: Antecedents and Consequences of IT Success.

(Epstein & Rejc, 2005), which describes the key factors for corporate success in IT integration. One of the basic premises of the model is that the impact of IT is realized mainly in combination with other organizational factors (Barua & Mukhopadhyay, 2000; Lim et al., 2004; Byrd et al., 2006). The model includes the critical inputs and specific processes that organizations need, which will lead to success in IT outputs (internal and external). However, as IT success ultimately must be measured by its contribution to overall organizational success (such as profitability or shareholder value) that is the ultimate outcome and measure of success, it includes outcomes as well.

The model implies that an organization's IT success is dependent on various *inputs*. This includes its existing corporate strategy, structure, and systems that provide both opportunities and constraints on IT initiatives. These, along with available resources and the external environment, are critical inputs that affect choices in the formulation and implementation of IT strategies (initiatives). Other factors, such as leadership and IT strategy, IT structure, and IT systems (*processes*) also significantly impact the performance and success of IT initiatives. Both the inputs and processes impact on various IT *outputs* that can be classified as either *internal outputs* such as improvement in productivity, time savings, increased utilization of capacities, improved quality, overall cost reduction, as well as *external outputs* such as channel optimization, customer acquisition, satisfaction, and loyalty, and overall value capture. If the IT strategy (initiative) formulation and implementation is successful, these outputs should ultimately be realized in improved overall corporate profitability (*outcomes*).

The viability of any IT initiative must therefore be estimated through proper evaluation of external environment and inputs available in an organization. Managers responsible for planning and developing IT initiatives must also consider the processes necessary to drive superior IT performance. Leadership of the organization, for example, must be knowledgeable about IT, committed to the IT initiatives, and aware of the impacts of existing organizational culture and behavioral patterns that may act as impediments to effective implementation of new IT initiatives. Top management involvement is an important factor in IT success (Armstrong & Sambamurthy, 1999). Similarly, it is essential that IT systems such as specialized HR practices for IT departments, IT training, performance measurement, and management control are part of the processes pertinent to IT. In many organizations, the gap between the rate of technology innovation and employees' skills and knowledge to use these innovations productively is growing preventing IT efforts to realize its full potential. Also critical is the alignment of the IT strategy with the corporate strategy and the establishment of appropriate IT structure.

If the IT initiatives are well designed and executed, the identified inputs and processes should lead to improved performance in outputs, and ultimately to increased corporate financial performance. The overall outputs of IT initiatives can be divided into two categories. Internal outputs relate to increased productivity, time savings, increased capacity utilization, improved quality, and direct cost savings. Increased productivity, for example, is one of the expected immediate benefits of new IT programs and projects. Improvements in IT infrastructure, for example, in terms of fully integrated application systems allow for better access to databases, faster exchange of information, reduced operating cycles, and so forth. In addition, the standardization of IT work processes, segmentation of the work, and global dispersion for greatest efficiency permit numerous improvements. These include reuse of applications and technical architectures, automation of much of the delivery process, and codification of methodologies so that they can be repeated, which all greatly increases productivity. IT can reduce the firm's fixed overhead costs, or reduce the variable costs of designing, developing, or manufacturing a product (Thatcher & Oliver, 2001). The financial consequences of improvements in internal outputs are all reflected in cost savings or, potentially, in increased sales. The external outputs, on the other hand, relate to achievements realized in the market and cover a broad array of results with respect to channel optimization, customer acquisition, loyalty, and retention, and overall value capture. Customer acquisition, for example, can significantly be increased by creating and using new channels of providing customers with products and services. Organizations, for example, that move more commerce to the web can accomplish expanded global coverage and exposure with a relatively minimal investment. For a more detailed description of all internal and external outputs, see Epstein and Rejc (2005).

For IT initiatives to be of value, the intermediate outputs must eventually payoff in increased organizational success (corporate profits). Viewed simply, increased profitability can only be achieved through reduced costs or improved revenues. Thus, in order to prove that IT investments in programs and projects were financially sound, the ultimate effect on corporate financial profitability must be determined and the payoffs clearly documented.

Following the *IT Contribution Model*, we propose a 6-step methodology that identifies critical drivers of an IT initiative success and creates causal relationships among these drivers. The visual presentation of the causality of drivers helps better understand how the inputs, processes, outputs, and outcomes of an IT initiative are interrelated. The methodology thus enables precise identification and measurement of all present and future marginal costs and benefits of IT initiatives fundamental for a comprehensive and objective calculation of IT initiative payoff. Finally, it also develops performance measures for the drivers that can be used for managerial control after an IT initiative is launched (see Fig. 2).

STEP 1: IT Initiative Overall Purpose and Goals

The IT Payoff Methodology starts by an overall description of an IT initiative purpose – Why should an IT initiative be implemented? What are the overall business outcomes of the project? – and goals – What are the expected direct results of an IT initiative? The overall purpose and goals can be stated as narratives but must clearly reflect the alignment of the IT initiative (solution) with the business. This will ensure project alignment with business imperatives and accountability for stated purpose and goals.

STEP 2: The IT Contribution Model*: Identify Relevant Inputs, Processes, Outputs, and Outcomes*

Step 2 introduces the *IT Contribution Model* that helps identify all required inputs and processes of an IT initiative, as well as expected outputs, and outcomes. Critical drivers specify more precisely the keys to IT success and

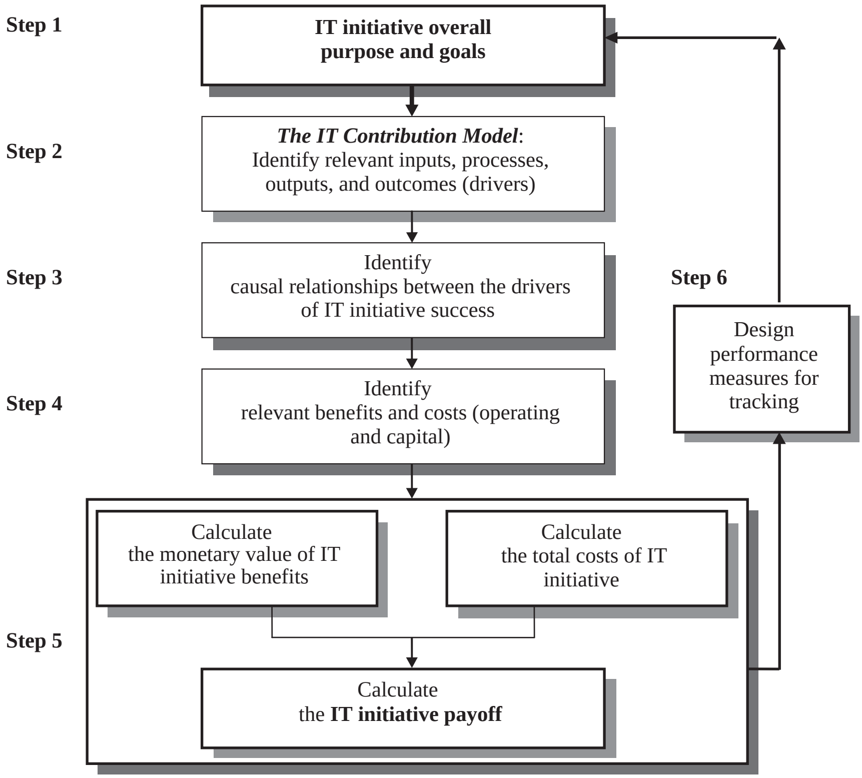


Fig. 2. The IT Payoff Methodology.

the actions that managers must take to improve the success of the IT activities that will ultimately impact on overall organizational success. With the *IT Contribution Model*, an IT initiative antecedents and consequences can be determined more comprehensively which is needed for a visual presentation of the causality of drivers, the designing of performance measures, identification of relevant benefits and cost, and the calculation of an IT payoff.

STEP 3: Identify Causal Relationships between the Drivers of IT Initiative Success

After having identified specific drivers of IT success, their causal relationships must be developed (see Fig. 3).

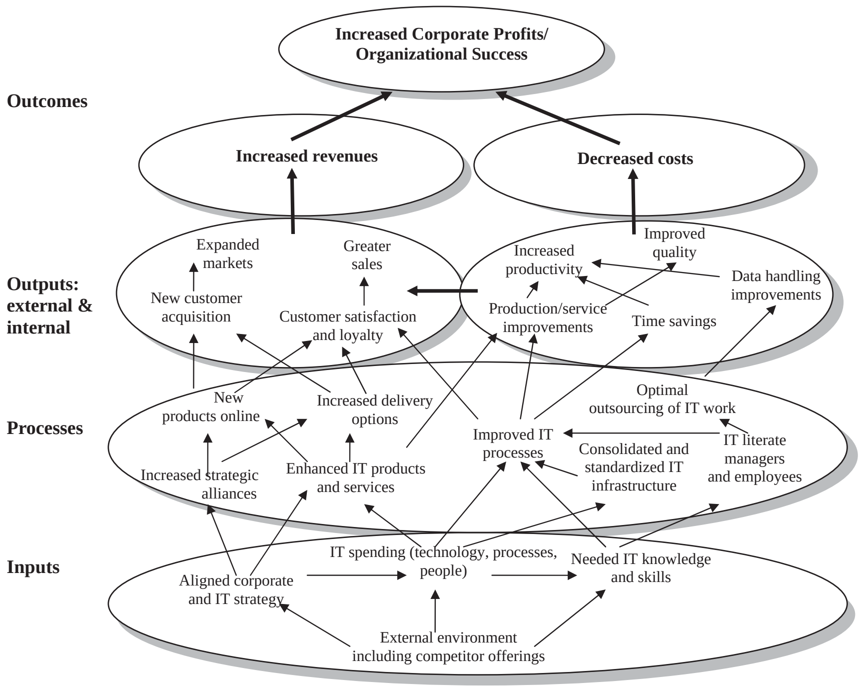


Fig. 3. Causality of IT Performance Drivers.

A clear understanding of the cause-and-effect relationships underlying the primary drivers of value in an IT initiative is one of the most important determinants of its success. In practice, there are numerous drivers of IT success and the ones outlined in Fig. 3 do not attempt to cover all choices. On the other hand, the illustrated example is comprehensive. In practice, there should be fewer critical IT performance drivers and the illustration of the causality of IT performance drivers less complex.

Fig. 3 shows, for example, that if organizations align the corporate and IT strategy, then they will potentially obtain more resources to spend on IT technology. More resources spent on IT technology can enable them to consolidate and standardize IT infrastructure leading to improved IT processes, increased productivity and quality, and decreased costs. Similarly, if organizations devote more resources to enhance IT products and services, they can increase delivery options leading to higher customer satisfaction, sales, and revenues.

Causal relationships between drivers within each of the four dimensions as well as between drivers in different dimensions are based on hypothetical assumptions of causes and effects, i.e., leading and lagging elements. In practice, the notion of leading versus lagging elements should be thought of as a continuum, as, for example, improved IT processes leads to time savings, but at the same time lags the IT spending. These hypothesized relationships need to be continuously tested and revised.

STEP 4: Identify Relevant Benefits and Costs (Operating and Capital)

Step 4 requires an exact specification of all benefits arising from the IT initiative and the capital and operational costs. The identified causal relationships between the critical drivers of an IT initiative will help determine both the costs and expected benefits of the initiative. Although benefits do not always clearly translate into short-term profits, they should ultimately lead to either cost savings or increased revenues. Sometimes, the direct relationship between a specific action or process, such as better and faster information, and the business value creation is not clear enough to provide an easy calculation of the benefit's monetary value. In such cases, additional inquiry in terms of 'How does this improvement specifically help you in your work?' should be undertaken. It may be that the system supports increased throughput per employee (increased productivity), saves time (time savings), helps optimizing the use of existing resources (increased capacity utilization), or allows fewer mistakes (improved quality). As shown in Fig. 4, the new methodology specifically recognizes the importance of measuring both the total costs of an IT initiative – including a range of different disruption costs – as well as the benefits, and additionally considers the risks associated with IT investments. It is important to note, that a precise identification and measurement of the present and future marginal costs and benefits of IT initiatives is fundamental for a comprehensive and objective calculation of IT initiative ROI. In particular, disruption costs associated with the adoption of IT initiatives require a thorough evaluation as they are typically significant.

STEP 5: Calculate the IT Payoff

In Step 5, the benefits are assigned monetary value and the costs are calculated. Finally, the IT payoff is calculated (see Fig. 4).

1 **Calculate the Monetary Value of IT Initiative Benefits**

Outputs	Benefits	Monetary Value
Increased Productivity	Increase in output (units produced, services offered)	$..................
Time Savings	Labor hours saved, machine hours saved, increased on-time deliveries reducing cost of grievances etc.	$..................
Increased Capacity Utilization	Increase in output (units produced, services offered)	$..................
Improved Quality	Labor hours saved, machine hours saved, cost of quality reduced, increased on-time deliveries reducing cost of grievances etc.	$.................. $..................
Direct Cost Savings	Reduced IT expenses, reduced direct administrative and operating costs, reduced fraud incidence, reduced hours of IS downtime	$.................. $..................
Channel Optimization, Customer Acquisition and Loyalty, Value Creation	Increased sales from existing and new customers	$.................. $..................

2 **Calculate the Total Costs of IT Initiative**

Costs		Value
Front-end Direct Costs of IT initiative	Hardware, software, installation and configuration costs, overhead, training costs	$.................. $..................
Disruption Costs Related to Human Factors	Decline in labor productivity, hours lost because of IT training, decline in product and service quality, absenteeism, revenues lost	$.................. $..................
Disruption Costs Related to Organizational Factors	Technical disruptions, breakdowns in service, costs of system support from vendors, organizational restructuring	$.................. $..................
Costs of Risk Mitigation	Development and implementation of IT performance framework	$..................
	Total Capital Costs	$..................
Operating Costs of IT Initiative	Direct IT operation costs, maintenance costs	$..................
	Total Operating Costs	$..................

3 **Calculate the IT Initiative ROI**

$$\mathbf{ROI} = \frac{\textbf{Total Benefits – Operating Costs}}{\textbf{Capital Costs (Investment)}} * 100$$

Fig. 4. Calculation of an IT Initiative Payoff.

IT ROI but also other financial performance indicators can be calculated such as the anticipated net present value of investment cash flows, the internal rate of return, and the period of investment payback; a flexibility analysis can also be carried out to determine the most critical factors of an IT initiative success. The IT payoff methodology can be used for both IT project justification prior to its start (planning) as well as for evaluation after completion (performance measurement).

STEP 6: Design Performance Measures for Tracking

To successfully attain the IT payoff goals, the cause-and-effect relationships in the causality map need to be monitored closely. For that purpose, appropriate metrics must be developed, consistent with and supporting the drivers of success, and milestones determined. Metrics should be used to foster an understanding of the IT initiative purpose and goals and performance drivers that will enhance cooperation between business units and stimulate a forward-thinking approach to achieving relevant objectives. The role of performance measures in motivating and coordinating employee behavior is fundamental as they – when properly designed and communicated – focus employee attention to the critical drivers of success. Performance measures and their targeted values also enable efficient managerial control of the IT initiative overall success.

The starting point for developing the appropriate metrics is the causal relationships of the IT initiative drivers. Attempts should be made to measure as many drivers as possible with monetary values. For example, improvements in quality may well be measured by the percentage of high-quality products, but it is more important to measure the dollars saved on less rework. Both the non-financial and financial measures, as long as they are expressed quantitatively, i.e., either in absolute or percentage terms, are useful, allow comparability, and target setting. However, financial measurement is especially important as managers want to calculate ROI and demonstrate IT payoff. Table 1 presents examples of performance measures that can be used for tracking an IT initiative progress and success.

It is important to focus on the key indicators rather than introducing indicators for everything that can be measured. Prior to the implementation of an IT initiative, baseline indicators for the specified performance measures need to be established. A lack of information of the initial status of the critical drivers of IT success prevents drawing conclusions about the actual benefits from IT initiatives after their completion. Even more

Table 1. Examples of Performance Measures for Tracking an IT Initiative.

Inputs	*Performance Measures*
Corporate strategy	% of planned change in annual IT budget
Corporate structure	Level of empowerment to strategic business unit (SBU) and functional managers
Corporate systems	% of employees compensated based on individual or group performance
Resources	Growth rate of IT spend per growth rate of direct total spend
External	Assessment of competitor IT investments
Environment	Assessment of customer and supplier needs and capabilities
Processes	*Performance Measures*
Leadership	% of CIO's and IT managers' bonus linked to IT profitability
Create and execute appropriate IT strategies	% of discretionary spending decisions aligned with corporate and business unit strategy
	Planned costs, benefits, and profitability of IT projects
Design and institute proper IT structure	% of systems developed/maintained outside the organization
	% of standardized hardware, databases, communications and applications systems
Develop and implement appropriate IT systems	% of IT employee turnover
	% of IT staff with pay for performance compensation
	Break/fix maintenance response/resolution time
Internal Outputs	*Performance Measures*
Increased productivity	% increase in production output per employee
	Dollar increase in sales based on productivity improvements
Time savings	Reduction in on-line response time
	Dollars saved based on time savings
Increased capacity utilization	% increase in capacity utilization
	% of utilization of databases
Improved quality	Dollars saved on prevention and appraisal cost of quality
Direct cost savings	% reduction in IT mandatory expenses
	Time saved on disaster recovery/business continuity
External Outputs	*Performance Measures*
Channel optimization	Dollar value of activities completed through web sites
	Hours of web site downtime (in a year)
Customer acquisition	% of customers using web sites exclusively
	% of visitors to web site who are also buyers (reach)
Customer loyalty	Sales from retained customers versus new customers
	% of customer attrition

Table 1. (*Continued*)

Value capture	Profitability of IT projects Number of new IT products and services introduced
Outcomes	*Performance Measures*
Long-term corporate profitability/ organizational success	% change in stock price attributable to IT initiatives EVA, ROI, ROA Earnings growth
Short-term corporate profitability/ organizational success	Cash flow growth Revenue growth % in overall cost reduction

importantly, target values (milestones) must be set for performance measures to establish benchmarks and to motivate.

The IT Payoff Methodology with its underlying *IT Contribution Model* has several advantages over other IT performance measurement approaches and IT valuation metrics. Firstly, the *IT Contribution Model* incorporates all important drivers of IT success as identified in various empirical and case studies. The model specifically underlines the role of strategic alignment and leadership in realizing the full potential of IT investment. The alignment of IT strategy with business strategy has been touted as a critical element in IT management and as a moderator between IT investment and firm performance (Byrd et al., 2006). Along with other impacts, the alignment of these two strategies increases the involvement of business managers in IT activities. The inclusion of senior IT managers in top management teams and their informal interactions, in particular, enhance IT managers' business knowledge (Armstrong & Sambamurthy, 1999). These, in turn, are better able to utilize their IT investment. The model also assumes the critical role of structural alignment (corporate structures aligned with IT structures) and, even more importantly, the alignment of corporate systems with IT systems. Brynjolfsson (2003) specifically points to the role of redesigned incentive systems and decentralized decision-making to achieve productivity gains. The model is thus based on valid assumptions and contributes to the existing literature by integrating the critical antecedents and consequences of IT success.

Secondly, the IT Payoff Methodology requires a careful consideration of all critical inputs and processes. When planning an IT investment, it is not only financial, human, and material resources that are considered, strategic alignment, potentially changed organizational structures and systems, as well as committed and knowledgeable leadership is also taken into account. When an evaluation of an IT investment's outputs and outcomes takes place, they are not analyzed in isolation but judged in the light of the model's antecedents' status. This is an important managerial contribution that distinguishes this methodology from extant ones.

Also, the visual presentation of performance drivers' causality helps better understand the cause-and-effect relationships among the various drivers of success. It points to the many areas that need improvement today in order to reap benefits in the output and outcome areas later. As such, it provides managers with timely information to make better IT decisions.

Further, the methodology requires a specification of performance measures and their target values along the time horizon to monitor the drivers' progress and to benchmark the final results against the initially projected net benefits. IT managers can thus execute effective managerial control over the milestones and the ultimate outcomes. Performance metrics is particularly important as it focuses attention on the critical drivers and stimulates a forward-thinking approach to achieving relevant objectives.

Finally, the methodology provides practical guidance on how to calculate the monetary value of IT benefits, which is often one of the major concerns of those responsible for calculating an IT ROI. The formula for IT ROI is provided along with the specification of all relevant capital and operational costs. None of the existing IT performance measurement methodologies incorporates all these characteristics.

APPLICATION OF THE IT PAYOFF METHODOLOGY IN ISTRABENZ GROUP

The Istrabenz Group is an international group of affiliated companies managed by the Istrabenz Holding Company. Its activity is organized into four divisions comprising energy, tourism, investments, and food, as well as IT support as an accompanying activity. In 2005, the company prepared guidelines on the use of ICT, from which it follows that this area is one of the key factors of the Istrabenz Group for achieving its strategic business goals; this is why the ICT strategy must be in line with business goals. Among other things, the guidelines regulate information system operation and the exploitation of synergies in ICT. This primarily involves the method of performing IT services and the efficient use of common resources such as the use of technological solutions that make possible the long-term stable operation and development of the Istrabenz Group (Istrabenz Group, 2005, Guidelines on the Use of Information and Communication Technology).

In 2005, the Istrabenz Holding Company was considering the introduction of the return on sales (ROS) information system as a uniform information system for supporting the operation of all the hotels in the Istrabenz Group tourism division. For the Morje Hotels, this would replace the FIDELIO information system, whereas for the Palace Hotels, the ROS information system had already been introduced in 2001. The company top management required exact calculations of the investment payoff. The *IT Contribution Model* and the IT Payoff Methodology have been selected as analytical tools and permission was granted for empirical testing of the model. As researchers, we were able to observe the effects of the implementation and the effectiveness of the model.

The company initiated its own project group for the IT initiative valuation. It included representatives of the Istrabenz Group tourism division and the ROS Company. The group comprised a project council (composed of the representatives of the tourism division and ROS), the project head, working group coordinators, the module head, key information system users, and advisers to ROS. The role of the key information system users was especially important as they are the ones that know best how the existing business operations function and what changes the new system is intended to yield. Key information system users were determined by the IT head at the Istrabenz Hotels; they included the managing director of the Morje Hotels, the managing director of the Palace Hotels, the representatives of the invoice and material accounting departments, the head of Food Supervision, and the sales manager at the Istrabenz Hotels.

The project group used the IT Payoff Methodology to determine the expected benefits from the information system unification and to set up the tools for subsequent managerial control if the project is approved. With the help of structured interviews, data were gathered on how the information system would change the operations. Starting points for discussions were prepared and sent to each interviewee in advance. The purpose was to determine the business process before and after the introduction of ROS and, on the basis of this, to determine the potential effects with the help of key users. In evaluating the effects of the ROS information system, secondary data from the Istrabenz Hotels were also used.

The subject of the research project is the IT initiative to introduce the ROS information system as a uniform information system for supporting the operation of the Morje and Palace Hotels. The study relates to one part of the tourism division of Istrabenz only but the project could be easily expanded to other parts of the company. The practical example of the IT Payoff Methodology application presented below demonstrates how methodology can be used for project justification prior to its start and for subsequent managerial control of the project and its applicability to other companies.

STEP 1: IT Initiative Purpose and Goals – The ROS Information System

The ROS information system comprises various modules specified for various business areas. It includes ROS HIS, an information system for hotel and convention services; ROS GIS, an information system for catering services; ROS ZIS, a health resort information system; ROS Wellness, a wellness information system; ROS FRS, a financial and accounting system; and ROS WEB extras, an online hotel reservation system with an integrated payment system (ROS company internal publication, 2003). During the project to introduce the ROS information system, the HIS, GIS, ZIS, and WELLNESS modules were implemented.

The main goals of introducing the ROS system as a uniform information system were the restructuring and unification of information solutions and processes in the tourism division of the Istrabenz Group with the purpose to ensure timely information for the needs of the companies' management, and information solutions that enable high-quality support for implementing tourism business processes. In reality, the introduction of the ROS information system was expected to have dual effects. On the one hand, these involve the effects of the replacement of the FIDELIO information system for the Morje Hotels and, on the other, the effects of the unification of both systems into a uniform information system.

STEP 2: The IT Contribution Model*: Inputs, Processes, Outputs, and Outcomes for the ROS Information System Implementation*

In accordance with the methodology, individual elements in the *IT Contribution Model* were defined in terms of their content (inputs, processes, outputs, and outcomes). Each element was carefully described by the project members, particularly by the selected main users of the ROS information system. In terms of the needed inputs, for example, all required resources were determined. The cost of the needed capital investment was calculated; there was no need to hire additional employees; the ROS Company offered IT training for employees that would be using the new information system with costs being incorporated in the capital investments numbers. The IT support and system maintenance costs were considered as operational costs. Corporate strategy was found to be supportive of the ICT strategy. Similarly, other inputs and required processes were determined. After the conversation with the managing directors of the Morje and Palace Hotels, the dimensions of the effects (outputs) of the ROS information system were divided in the following areas for the needs of investment evaluation: sales, reception office, catering and wellness, support staff, material accounting, invoices, and general effects.

Some of the expected results of introducing the ROS information system were not completely definitive. In the evaluation of results, various hypotheses about cause-and-effect relations between the ROS and operations were used, which were defined on the basis of key users' experiences. In defining the hypotheses, there was a certain extent of uncertainty regarding their accuracy. The project group tried to eliminate this with the help of sensitivity analysis. Another possibility is the probability theory, where several scenarios are created for a specific fuzzy hypothesis and then probability is attributed to them (Anandarajan & Wen, 1999, p. 329).

STEP 3: Identify Causal Relationships Between the Drivers of IT Initiative Success

Causal relationships between inputs, processes, outputs, and outcomes must be result of a unified understanding of how the project is expected to evolve towards its goals and purposes. The visual presentation of the cause-and-effect relationships between antecedents and consequences of the ROS information system implementation is shown in Fig. 5. The most interesting areas in the figure are the processes and outputs area with descriptions of critical changes in the hotel operations and subsequent effects on the customers. As can be seen, the final outcomes are manifested as lower operating expenses or as an increase in sales revenues, which results in an increase in the company's profit.

The project group used the comprehensive causal relationships scheme as the basis for laying out the processes expected to change and for describing all the expected cost and benefits.

STEP 4: Relevant Costs (Operating and Capital) and Benefits of the ROS Information System Implementation

The changes caused by the ROS information system and the subsequent costs and benefits will be presented in more detail in the area of catering and

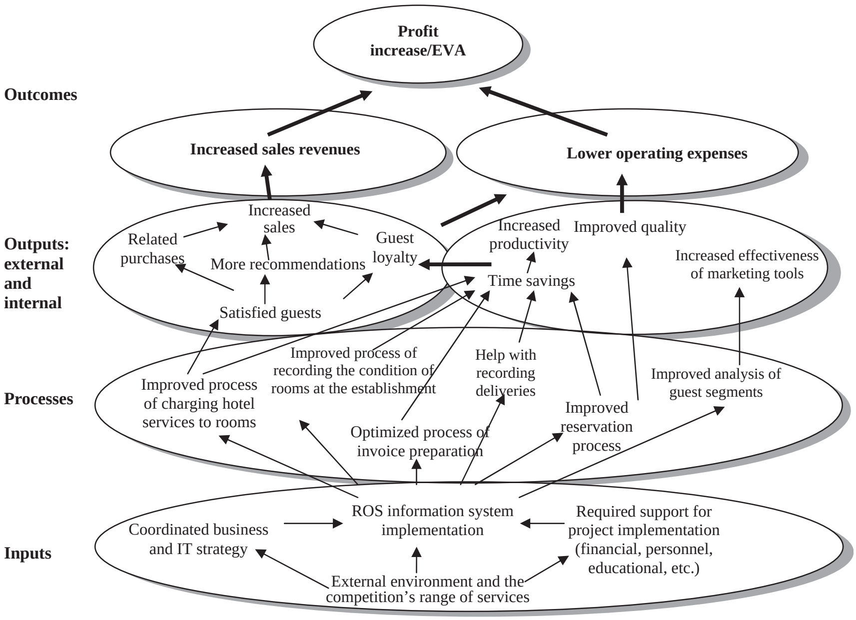


Fig. 5. Causal Relationships between the Drivers of Implementing the ROS Information System. *Source:* Moze (2006).

wellness; other areas were analyzed in the same way. Catering at Istrabenz Hotels includes service on all the premises, both in bars and restaurants. In service, the ROS information system helped mitigate the process of calculating hotel credit or the charging of guests' hotel services to their rooms. In wellness, the process is similar, with the only exception that time savings are a little different.

Fig. 6 shows the process of charging services to rooms before and after the introduction of the ROS information system at the Morje Hotels.

Prior the introduction of the ROS, the process starts with the guest's order, for which the waiter prepares a check. The guest signs it, by which he confirms that he has used the service. The waiter has to enter the check into the account book and then take it to the reception office, where the receptionist checks if the guest is really staying in the room he has stated. Then he confirms the copy of the check, in which he assumes responsibility for any potential non-payment. The waiter takes the check back to the reception office, while the receptionist has to put the data on the guest's room and insert the original check in the room's pigeonhole. The checks

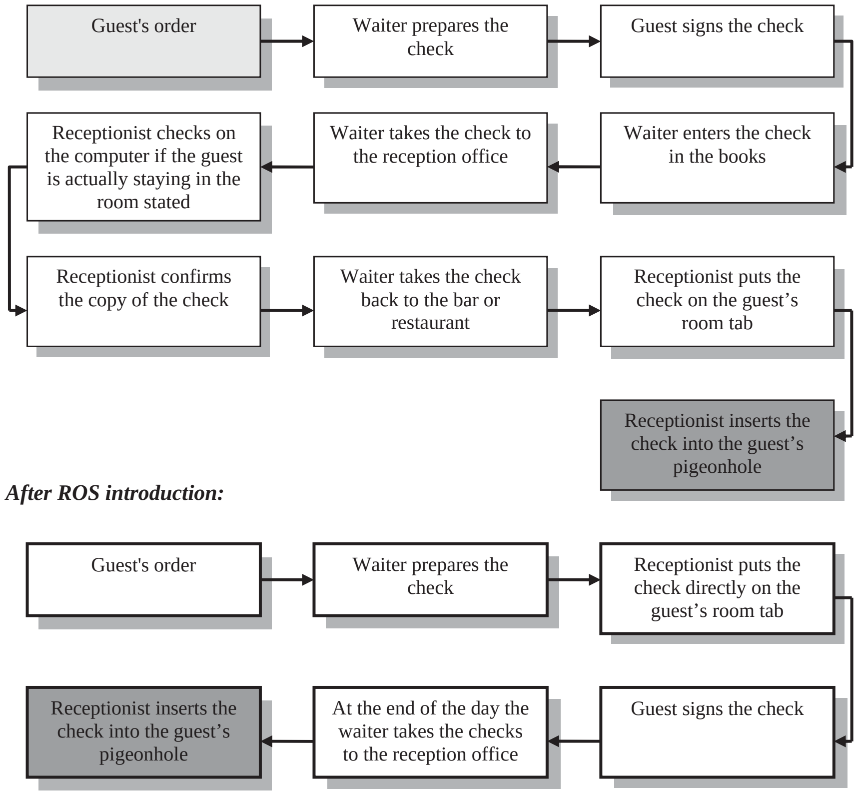


Fig. 6. The Process of Charging Hotel Bar and Restaurant Services to the Room before and after the Introduction of the ROS Information System.

prove that the guest has really used the service charged. After the introduction of the ROS, the process of charging services to rooms will change. The check will be automatically put on the guest's room tab as soon as the waiter prepares it and this is also recorded on the receptionist's computer. At the end of the day, the waiter only has to take the checks to the reception office for recordkeeping if guests demand proof that they really used the service charged.

The ROS information system was expected to optimize the process of charging hotel services to rooms. Through direct transfer of the check from the bar, restaurant, or wellness center to the guest's room, savings in the time used for the process are created for both the waiter as well as the hotel or wellness center receptionist. After the conversation with the managing director of the Morje Hotels, time savings that the ROS information system makes possible were determined. Thus, it has been determined that a waiter saves 2.30 min for each process of charging to the room. A more accurate division of the waiter's time saved is as follows: (1) time saved walking to the reception office: 1.30 min, (2) waiting at the reception office to confirm the check copy: 45 sec, and (3) time used to enter the check into the books: 15 sec.

The division of time saved for the wellness center receptionist is similar to that for the waiter. By using the ROS information system, a receptionist would be no longer required to check if the guest is actually staying in the room stated and to confirm the check copies. He would only insert the checks brought to them from the catering or wellness center into the pigeonholes. This saves a receptionist 45 sec per process. These effects of time savings are direct; but the effects can also be indirect. By saving time, the ROS relieves employees' (waiters' and receptionists') workload, and so they can devote more time to the guests, which increases their satisfaction. Greater guest satisfaction results in increased use of hotel services. The assumption is that a satisfied guest will be happy to return and/or will recommend the hotel to friends and acquaintances. The effects of greater guest satisfaction can thus be summarized as follows: (1) related purchases: guests will use more hotel services, (2) repeated purchases: guests will be happy to return, and (3) recommendations: satisfied guests will recommend the hotel to their friends.

For a better overview of all direct results of introducing the ROS information system, a table is provided in Appendix.

STEP 5: Calculation of Total Costs and Benefits of the ROS Information System Implementation

The costs associated with the investment in the ROS information system comprise software purchase, hardware purchase, training and education costs, opportunity costs of employees' time, and annual maintenance costs. The costs in the first four groups are one-time costs that are created at the project's beginning (capital costs), whereas maintenance represents an annual (operating) cost. Training and education include costs of training and education for all modules implemented in the project. The opportunity costs of the employees' time are based on the time used for training and education, during which their normal work was interrupted. From the ROS time schedule, the information was retrieved on the planned number

Table 2. Expected Total Costs of Investment in the ROS Information System.

Cost	Calculation	Item Value (€)
Software	Istrabenz Turizem internal sources	95,438
Hardware	Istrabenz Turizem internal sources	19,629
Training and education	Istrabenz Turizem internal sources	9,114
Opportunity costs of employees' time	Labor hours planned for training and education × labor cost per hour	3,990
Total capital costs	Sum of all items above	128,171
Annual maintenance	Istrabenz Turizem internal sources	2,304

of hours for training and education by every employee. This time was then multiplied by employees' average hourly payment, which equaled the expected opportunity costs of the employees' time (see Table 2).

Table 3 shows the calculation of anticipated financial effects of time savings in catering and wellness after the introduction of the new information system for a five-year period of time (2006–2010). At the Istrabenz tourism division, the payback period for information systems is five years, which can be considered the period of the system's duration and thus the period of investment evaluation. The estimations for the first two years of the ROS being in place and in use also include the so called implementation factors which are used to substitute for suboptimal use of the information system and the subsequent disruption costs (0.6 and 0.8, respectively).

The same procedure was used for other areas (sales, reception office, support staff, material accounting, and invoices). Table 4 shows the calculation of selected anticipated financial benefits from general effects, such as increased guest satisfaction. An increase in the base of regular guests and recommendations by satisfied guests are included. Other effects include increased efficiency of marketing tools, related purchases, as well as decreased number of claims, and they were all included in the final calculations.

Altogether, total expected benefits for the first year sum up to € 55,073, the total costs are in the amount of € 130,474; in the second year, the benefits increase to € 73,882 and continue to rise up to the final year, while the costs fall to € 2,303 and remain at the maintenance level.

The investment in the ROS information system was evaluated by calculating the anticipated net present value of investment cash flows, the internal return rate, and the period of investment payback; at the same time,

Table 3. Calculation of Expected Financial Effects of Time Saving in Catering and Wellness after the Introduction of ROS.

Area	Result	Calculation	Financial Effects (€)				
			2006	2007	2008	2009	2010
Catering and wellness	Saving 2.5 min in waiter's work per each process of charging to room	*Time savings*: number of processes per year × 2.5 min = 7,500 × 2.5 min = 687.5 h *Time value*: € 5 per hour *2006 annual savings*: 687.5 h × € 5 × 0.6 = € 2,063	2,063	2,754	3,505	3,567	3,630
	Saving 1.75 min of wellness receptionist's work per every process of charging to room	*Time savings*: number of processes per year × 1.75 min = 8,250 × 1.75 = 239.3 h *Time value*: € 4.2 per hour *2006 annual savings*: 239.3 × € 4.2 × 0.6 = € 603	603	799	1,016	1,035	1,053
	Saving 45 sec of receptionist's work at Morje Hotels per every process of charging to room	*Time savings*: number of processes per year × 45 sec = 24,750 × 45 = 309.375 h *Time value*: € 6.25 per hour *2006 annual savings*: 309.375 × € 6.25 × 0.6 = € 1,160	1,160	1,549	1,972	2,007	2,042
	Saving 30 sec of the receptionist's work at Palace Hotels per day	*Time savings*: 3 h per year *Time value*: € 6.25 per hour *2006 annual savings*: 3 × € 6.25 per hour × 0.6 = € 11.25	11.25	15.02	18.78	18.77	18.77

Table 4. Calculation of Expected Financial Effects of Increased Base of Regular Guests and Recommendations by Satisfied Guests.

Effect	Result	Calculation	Financial Effects (€)				
			2006	2007	2008	2009	2010
Increased base of regular guests	Lower promotion costs	*Formula*: Δ of average use × annual promotion costs × training factor *2006 annual savings*: 1% × € 1,070,694 × 0.6 = € 6,424	6,424	8,724	11,103	11,302	11,500
Recommendations by satisfied guests	Lower promotion costs	*Formula*: Δ of the number of overnight stays × advertising costs per overnight stay × training factor Δ *of number of overnight stays with recommendations*: 0.05% × 64,800 = 32 *Promotion costs per overnight stay*: € 6.6 *2006 annual savings*: 32 × € 6.6 × 0.6 = € 126.72	126.72	174.48	222.10	226.03	230

a sensitivity analysis was also carried out. Investment cash flows represent the calculated financial effects of the ROS information system and the costs of investment. The discount rate by which the investment cash flows are discounted is the required return rate of the investor – in this case, the owner of the Istrabenz Group – according to the investment risk and returns of alternative investments with comparable risk. The required return rate of the investment in the ROS information system was 8.5%. The calculations show the net present value of investment cash flows is € 220,068, the internal rate of return is 139%, and the period of investment payback is 2 years and 15 days.

Sensitivity analysis tested the sensitivity of the investment's evaluation to the change in the evaluations of savings on marketing costs made by the sales manager at the Istrabenz tourism division. Results show that in the case of an evaluation of savings on marketing costs decreased by 1%, the net present value decreases by 8%, whereas the internal return rate decreases by 17 percentage points or by 12%. In the case of an evaluation of savings on marketing costs increased by 1%, the net present value of the investment increases by 9%, whereas the internal return rate increases by 19 percentage points or by 14%. With the $\pm 1\%$ change in evaluation of savings on marketing costs, the period of investment payback extends by 29 or shortens by 30 days. The sensitivity analysis presented above offers an example of great investment sensitivity to the evaluation of savings on marketing costs. This was taken into account in the final evaluation of the effectiveness of the investment in the ROS information system.

STEP 6: Design Performance Measures for Tracking

In the final step, after the project was approved, performance measures were developed to foster the anticipated changes as foreseen in the initial calculations. Performance measures were drawn from the causality of drivers' scheme; for outputs, however, a more detailed look at the changes in various processes was needed to design appropriate indicators. Table 5 provides selected performance measures for tracking the outputs of catering and wellness.

For many performance measures that have not been tracked before baseline indicators were determined. Specific measurements and evaluations took place as separate activities in the project to determine these baseline values. Then, target values or milestones were set for performance measures across the expected period of investment payback. These values were determined by

Table 5. Selected Performance Measures for Tracking the Outputs of Catering and Wellness.

Catering and Wellness Internal Outputs	Performance Measures
Time savings	% of time saved in the waiter's work
	% of time saved in the hotel receptionist's work
	% of time saved in the wellness receptionist's work
	Dollars saved based on time savings
Improved quality	% decrease in customer complaints related to waiters'/receptionists' work
	Dollars saved based on fewer customer complaints
External Outputs	*Performance Measures*
Customer acquisition	% of guests being recommended by friends and acquaintances
Customer loyalty	% of guests returning to the hotel
	Sales from retained customers versus new customers
Value capture	Dollars earned on related purchases

the project group members but with a consent of those employees who were responsible for the processes and activities under evaluation.

DISCUSSION

The article addresses an important methodological question that has been addressed in both the IT and management control literature, namely the question of identifying, measuring, and managing the IT's contribution to the bottom line. While empirical research work, recent studies, in particular, provides evidence of the IT value, there is little practical guidance on how to design and implement an appropriate IT performance measurement system. Various approaches and IT valuation measures fall short on providing a comprehensive overview of all critical drivers of IT success, their interrelations, the way they can be measured, and how to make better IT decisions based upon the analysis. In this article, we provide an integrated model (the *IT Contribution Model*) and a methodology (the IT Payoff Methodology) that bridge this methodological gap and help organizations measure an IT initiative's payoff in a more comprehensive way and execute efficient management control.

The academic contribution of the article is twofold. On the one hand, we present and empirically test, the *IT Contribution Model*, which upgrades the existing literature by offering an integrated model of critical drivers of

IT success. The model was designed based on empirically tested assumptions about the cause-and-effect relationships between the antecedents and consequences of IT success provided in existing empirical studies. With a model incorporating IT inputs, processes, and outputs that lead to overall IT payoff and improved corporate profitability, organizations will less likely rely on a reactive approach to their adoption of new technologies or risk making costly, personality-driven choices.

On the other hand, we also present and implement the IT Payoff Methodology, which represents a more complete analytical tool for evaluating the payoffs of investing in IT based on the proposed model. The methodology is more straightforward than existing tools as it relies on six well-defined steps, applies standard methods and analytical tools, and does not require complex calculations. It includes the identification of critical drivers of an IT investment's success, develops the cause-and-effect relationships between the drivers and outcomes, helps identify and measure all important costs and benefits of the IT initiative to calculate the IT initiative ROI, and provides performance measures for tracking the IT initiative. All these steps are necessary to properly value and manage an IT investment.

Apart from the academic dimension, the article also has several practical implications. The new methodology for valuing IT investments offers practical insights into how to identify, measure, and manage the critical drivers of IT success. More specifically, the IT Payoff Methodology helped decision makers at Istrabenz in several ways

- Firstly, it provided exact calculations of the expected investment payoff and enabled well-informed resource allocation decision, which was the initial purpose of the project. The methodology specifically recognizes the importance of measuring both the total costs of an IT initiative – including a range of different disruption costs – as well as the benefits, and additionally considers the risks associated with IT investments. Since most organizations have little experience in assigning monetary values to IT outputs and the measurement of IT payoffs, the methodology's specific instructions on these questions helped resolve many dilemmas.
- Secondly, by having a clear picture of the IT cause-and-effect relationships, IT managers can monitor how the IT initiatives are progressing and more fairly evaluate their intermediate results. The causal linkage map of drivers is useful and important as it helps ensure that all actions that are necessary to achieve success are taken, that unnecessary actions are not taken, and that all employees understand their critical role in the success of the IT activities.

- Thirdly, the project group members specifically acknowledged the importance of steps 2–5 for a precise and objective calculation of the IT initiative payoff. The visual representation of the causality of critical drivers of success was considered as particularly helpful for projecting the monetary benefits and costs of the IT initiative. The financial calculations have further shown how important it is to understand the influence that the hypotheses about the cause-and-effect relations and various subjective evaluations have on the investment net present value, internal return rate, and payback period. The sensitivity analysis has revealed great investment sensitivity to the evaluation of savings on marketing costs.
- Fourthly, the IT Payoff Methodology requires the active participation of all important holders of processes under investigation, their managers, and subordinate employees. The active involvement in the identification of the critical drivers of success, expected internal and external outputs, in particular, however, substantially facilitates the IT initiative implementation by increasing the level of understanding and acceptance. In Istrabenz tourism division, this benefit has been widely acknowledged.
- Finally, the Istrabenz project team found the methodological tool to be very pragmatic, simple, and with feasible implementation costs. The methodology can be performed internally without the need to hire external consultants.

The project group also listed potential challenges of the methodology.

- Firstly, the methodology can best be applied when extant business processes are already identified and described thus allowing for the establishment of baseline indicators of performance. In the opposite case, baseline measurements and evaluations need to take place, which takes time but is crucial for subsequent comparisons.
- Secondly, the methodology necessarily requires various assumptions about expected savings from improved processes. The objectivity of these assumptions is best attained when they are set by those who perform these processes, the so-called key users of the new IT. Still, sensitivity analyses or probability scenarios are required to mitigate some of the uncertainty.
- Thirdly, one of the most vexing problems in estimating performance impacts of IT investment is simultaneity bias. If companies undertake technology implementations when demand for their products is high or when they expect to perform well, estimates of the impact of IT adoption on performance may be biased upward creating indeterminacy in causal interpretations (Brynjolfsson & Hitt, 2003).

- Fourthly, the case study supports Brynjolfsson's finding (Brynjolfsson, 2003), namely, that companies do not simply plug in computers and telecommunications equipment and achieve productivity gains. Discussions with the project group members revealed that without efforts to improve employee IT literacy, understanding, and ability to use these innovations, further, actually measure improvements based on implemented innovations, and, finally, establish proper compensation policies to stimulate employees to deploy the use of IT, the projected benefits will not be realized. To realize full potential of IT investment, organizations must often go through a process of organizational redesign. Brynjolfsson refers to a cluster of related innovations, such as automation of numerous routine tasks, highly skilled labor, more decentralized decision-making, improved information flow vertically and laterally, strong performance-based incentives, and increased emphasis on recruiting and training (Brynjolfsson, 2003, p. 42). Earlier research and case studies have also proven that IT investments complement other long-term performance-enhancing investments, including innovations in business methods and organization, human capital investments, and supply chain management systems, which are carried out over a period of several years (Brynjolfsson & Hitt, 2003; Bresnahan, Brynjolfsson, & Hitt, 2002; Davenport & Short, 1990; Short & Venkatramen, 1992).
- Finally, it has also been agreed in the project group that, to attain business value from an IT initiative, a structured and ongoing careful examination of costs, benefits, and risks from the initial feasibility through post-implementation is needed. Even when business value is achieved, there is no guarantee that this value will be maintained unless there is an ongoing attention to IT performance measurement.

The implementation of an IT payoff measurement system should by no means be seen as a threat to or imposition on staff, rather as a mechanism to enhance performance and corporate learning. A properly developed and implemented measurement system promotes productivity by focusing attention on the most important issues, tasks, and objectives of the project.

CONCLUSION

The article presents the empirical testing of a new model and the subsequent methodology for identifying and measuring the IT investment payoff in the case of Istrabenz Group. While the use of a singular study limits the generalization of the findings, the research shows that with a properly implemented *IT Contribution Model* and the IT Payoff Methodology IT managers are able to demonstrate the impact on corporate profitability and value creation from IT.

The new methodology can assist IT managers as they evaluate the trade-offs and decide which IT project provides the largest net benefit to both short-term financial performance as well as the overall long-term success of the organization. It can help CIOs, CTOs, CFOs and other senior corporate and financial managers as they develop an IT strategy to make overall corporate resource allocations to support that strategy. They can rely on convincing evidence based on formal measurement and evaluation when making recommendations on these allocations. Also, the IT staff will know better how well they are performing, correct any deficiencies, and by seeing the results of their work develop an important sense of personal satisfaction.

The *IT Contribution Model* and the IT Payoff Methodology can be adapted into any management system that an organization utilizes. It is compatible with measurement and management frameworks such as the balanced scorecard and shareholder value analysis that focus on a better understanding of the causal relationships and linkages within organizations and the actions managers can implement to improve both customer and corporate profitability and drive increased value.

The proposed methodology could be further improved, although methodological refinements should not jeopardise its pragmatism and comprehensiveness, which are two of its greatest benefits. It would be of great theoretical and practical importance to be able to test the methodology along all six steps, including the role performance measures can play in coordinating employee efforts, both as metrics in the performance measurement system as well as reward triggers in the compensation system. The *IT Contribution Model* should also be further tested and validated to provide additional empirical evidence of the causal relationships stated in the model.

ACKNOWLEDGMENTS

We are grateful to the Istrabenz Group, Inc., for providing a suitable context for empirical testing of the model and Andrej Moze for conducting the research and granting access to the data.

REFERENCES

Anandarajan, A., & Wen, J. H. (1999). Evaluation of information technology investment. *Management Decision, MCB University Press*, *37*(4), 329–337.

Apfel, A. (2002). The total value of opportunity approach. Decision Framework, DF-17-0235.

Aral, S., Brynjolfsson, E., & Wu, D. J. (2006). Which came first: IT or productivity? The virtuous cycle of investment and use in enterprise systems. Paper presented at 27th International Conference on Information Systems, Milwaukee.

Armstrong, C. P., & Sambamurthy, V. (1999). Information technology assimilation in firms: The influence of senior leadership and IT infrastructures. *Information Systems Research*, *10*(4), 304–327.

Barua, A., & Lee, B. (1997). The IT productivity paradox revisited: A theoretical and empirical investigation in the manufacturing sector. *The International Journal of Flexible Manufacturing Systems*, *9*, 145–166.

Barua, A., & Mukhopadhyay, T. (2000). Information technology and business performance: Past, present, and future. In: R. M. Zmud (Ed.), *Framing the domains of IT management: Projecting the future through the past* (pp. 65–84). Cincinnati, OH: Pinnflex Education Resources.

Bresnahan, T., Brynjolfsson, E., & Hitt, L. (2002). Information technology, workplace organization and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics*, *117*(1), 339–376.

Brynjolfsson, E. (2003). The IT productivity gap. *Optimize*, ABI/Inform GLOBAL, July, pp. 26–43.

Brynjolfsson, E., & Hitt, L. M. (1996). Paradox lost? Firm-level evidence on the returns to information systems spending. *Management Science*, *42*(4), 541–558.

Brynjolfsson, E., & Hitt, L. M. (2003). Computing productivity: Firm-level evidence. *Review of Economics and Statistics*, *85*(4), 793–808.

Byrd, T. A., Lewis, B. R., & Bryan, R. W. (2006). The leveraging influence of strategic alignment on IT Investment: An empirical examination. *Information and Management*, *43*, 308–321.

Cameron, B., Meringer, J., Dawe, C., & Jastrzembski, E. (2000). Measuring eBusiness success. *The Forrester Report*, September.

Carr, N. G. (2003). IT doesn't matter. *Harvard Business Review*, *81*(5), 41–49.

Carr, N. G. (2004). *Does IT matter? Information technology and the corrosion of competitive advantage*. Boston: Harvard Business School Press.

Curley, M. (2004). *Managing information technology for business value: Practical strategies for IT and business managers*. Hillsboro: Intel Press.

Davenport, T., & Short, J. (1990). The new industrial engineering: Information technology and business process redesign. *Sloan Management Review*, *31*(4), 11–27.

Davern, M. J., & Kaufman, R. J. (2000). Discovering potential and realizing value from information technology investments. *Journal of Management Information Systems*, *16*(4), 121–143.

Davies, L., Dehning, B., & Stratopoulous, T. (2003). Does the market recognize IT-enabled competitive advantage? *Information and Management*, *40*, 705–716.

Devaraj, S., & Kohli, R. (2002). *The IT payoff: Measuring the business value of information technology investments*. Upper Saddle River: Prentice-Hall.

EAC Report. (2005). *Beyond ROI: Enterprise payback* (Enterprise Applications Consulting, Joshua Greenbaum, Principal).

Epstein, M. J., & Rejc, A. (2005). *Evaluating performance in information technology. Management accounting guideline*. Hamilton: The Society of Management Accountants of Canada.

Gartner Group. (1997). *Total cost of ownership analyst: A white paper on GartnerGroup's next generation total cost of ownership methodology*. Stamford: GartnerConsulting.

Hajela, S. (2005). IT value mapping: A quantitative approach to maximize returns on it investment. Available at http://wwwstartsmarts.com

Intel Information Technology White Paper. (2003). Managing IT investments: Intel's IT business value metrics program, August.

Istrabenz Group. (2005). Guidelines on the use of information and communication technology.

Leavitt, W. (1999). Profit talking. *Fleet Owner*, *94*(1), 57–61.

Lee, C. S. (2001). Modeling the business value of information technology. *Information and Management*, *39*, 191–210.

Lim, J. H., Richardson, V. J., & Roberts, T. L. (2004). Information technology investment and firm performance: A meta-analysis. In: E. R. Dennis (Ed.), *Proceedings of the 37th Hawaii International Conference on System Sciences*, Hawaii.

Love, P. E. D., & Irani, Z. (2004). An exploratory study of information technology evaluation and benefits: Management practices of SMEs in the construction industry. *Information and Management*, *42*, 227–242.

Loveman, G. W. (1994). An assessment of the productivity impact of information technologies. In: T. J. Allen & M. S. Scott-Morton (Eds), *Information technology and the corporation of the 1990s: Research studies*. Cambridge, MA: MIT Press.

Lutchen, M. D. (2004). *Managing IT as a business: A survival guide for CEOs*. Hoboken: Wiley.

Moze, A. (2006). *The application of a new methodology for evaluating IT investments – A case study*. Master's thesis, University of Ljubljana, Faculty of Economics.

Mukhopadyay, T., Kekre, S., & Kalathur, S. (1995). Business value of information technology: A study of electronic data interchange. *MIS Quarterly*, *19*(2), 137–156.

Murphy, T. (2002). *Achieving business value from technology: A practical guide for today's executive*. Hoboken: Wiley.

Remenyi, D., Money, A., & Sherwood-Smith, M. (2000). *The effective measurement and management of IT costs and benefits*. Oxford: Butterworth-Heinemann.

Schubert, K. D. (2004). *CIO survival guide: The roles and responsibilities of the chief information officer*. Hoboken: Wiley.

Schwartz, E. (1999). CEOs diss IT investments. *InfoWorld*, *21*(1), 32.

Short, J. E., & Venkatramen, N. (1992). Beyond business process redesign: Redefining Baxter's business network. *Sloan Management Review*, *34*(1), 7–21.

Strassman, P. A. (1990). *The business value of computers*. New Canaan: Information Economics Press.

Strassman, P. A. (1999). *Information productivity: Assessing the information management costs of US industrial corporations*. New Canaan: Information Economics Press.

Tardugno, A. F., DiPasquale, T. R., & Matthews, R. E. (2000). *IT services: Costs, metrics, benchmarking, and marketing*. Upper Saddle River: Prentice Hall.

Thatcher, M. E., & Oliver, J. R. (2001). The impact of technology investments on a firm's production efficiency, product quality, and productivity. *Journal of Management Information Systems*, *18*(2), 17–45.

Weill, P. (1992). The relationship between investment in information technology and firm performance: A study of the valve manufacturing sector. *Information Systems Research*, *3*(4), 307–333.

Weill, P., & Ross, J. W. (2004). *IT governance: How top performers manage IT decision rights for superior results*. Boston: Harvard Business School Press.

APPENDIX. RESULTS OF INTRODUCING THE ROS INFORMATION SYSTEM PER INFLUENCE AREA

Area	Result	Result Specification
Sales	Time savings	ROS accelerates the response to the demand and thus creates savings in the work time of • Sales personnel: 10% of work time, • Heads of reservation department: 15% of work time.
	More efficient management tools	ROS enables better analysis of marketing segments and more target-oriented marketing. Thus, the company saves on marketing expenses.
	Fewer complaints	ROS decreases the number of reservation errors and thus the number of complaints. In this way, certain costs concerning claims are saved.
	Fewer guests lost	Unsatisfied guests may not complain, but they will never return.
Catering and Wellness	Time savings	ROS saves time for the • Waiter: 2.5 min per process, • Hotel receptionist: 45 sec per process, • Wellness receptionist: 1.75 min per process.
	More satisfied guests	ROS relieves the waiter's and receptionist's workload, which is why they can devote more time to guests. Thus, ROS results in • Related purchases, • Guests returning to the hotel, • Recommendations to guests' friends and acquaintances.

APPENDIX. (*Continued*)

Area	Result	Result Specification
Support staff	Time savings	Savings in the work time of • Cleaning staff: 1 min per day, • Receptionist: 30 sec per day.
Material Accounting	Time savings	ROS aids in recording deliveries and thus saves two bookkeepers 1 h of work per day.
Invoice Department	Time savings	Time savings will not be visible until next year; they will amount to 25% of the work of two invoice clerks.
General Effects	Optimization of human resources	ROS enables detailed analyses that help organize work.
	Economization of human resources	ROS standardizes working processes and thus simplifies transfers of employees from one hotel to another.
	Greater guest satisfaction and loyalty	ROS enables the use of a uniform guest database, creating a basis for keeping records of regular guests, and can also help direct the relationship with the customer. This increases guest loyalty and satisfaction, which results in • Related purchases, • Guests returning to the hotel, • Recommendations to guests' friends and acquaintances.

EFFECTS OF REFLECTIVE THINKING AND PROFESSIONAL EXPERIENCE ON PURCHASING DECISIONS WITH INACCURATE COST INFORMATION

Sebastiaan Morssinkhof, Marc Wouters and Luk Warlop

ABSTRACT

This article addresses purchasing decisions and the use of total cost of ownership (TCO) information. TCO is based on a monetary quantification of nonfinancial attributes and aggregation into a summary measure (such as cost per hour, per wafer, or per kilometer). From an accounting point-of-view, one intricate issue is the accuracy of the monetary quantification and how this affects decision-making. We distinguish three different kinds of inaccurate monetary quantification, and we investigate the weight that decision makers attach to attributes that are inaccurately monetarily quantified and subsequently included in TCO information. Specifically, we investigate whether this weight depends on reflective thinking and experience. This question is relevant beyond TCO, for all decision-making situations that involve monetary quantification of attributes and subsequent aggregation, such as in activity-based costing,

Advances in Management Accounting, Volume 17, 81–112
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017005

net present value calculations for capital budgeting decisions, or cost-benefit analyses in public administration.

We found support for the hypothesis that reflective thinking increases the weight decision makers attach to the attribute that is included as a minimum cost in the TCO-numbers, but not for the hypothesis that reflective thinking would reduce the weight attached to the attribute that is included as a maximum cost in the TCO-numbers. Students and practitioners differed significantly in the weight they attached to an attribute that was excluded from the TCO-numbers, and practitioners gave less weight to such attributes. Together these results suggest that TCO-numbers should be provided with care and possible inaccuracies should be clarified.

INTRODUCTION

This article focuses on comparing alternative supplier offerings when making purchasing decisions, and it investigates how such comparisons can be supported by accounting information. Many different considerations come into play when comparing alternative supplier offerings, and various kinds of data may be available: qualitative information (such as lists of reference customers provided by different suppliers), quantitative non-financial information (such as lead times), and financial information (such as purchase price and discounts). There is no shortage of data available in organizations, but we know that human decision makers have limited mental capacity for parallel processing of all this information, especially when qualitatively different kinds of information are involved (Bettman, Luce, & Payne, 1998; Bonner, 1994). To simplify the task, they may limit their focus to the subset of information that is easily expressed with a common (dollar) metric (Kadous, Koonce, & Towry, 2005). This strategy will reduce the mental burden, but at the cost of disregarding factors that may be important, but cannot be monetarily quantified. In this study, we investigate the challenge for people to consider both types of information in a purchasing decision.

The term "total cost of ownership" (TCO) is used in accounting for the monetary quantification of all costs associated with acquiring and using a particular purchasing alternative, such as transaction costs related to purchasing activities (e.g., ordering, freight, quality control), costs resulting from poor quality (e.g., rejection, rework, and warranties), or costs related to performance (e.g., inventory, transportation costs) (Carr & Ittner, 1992; Ellram, 1995; Wouters, Anderson, & Wynstra, 2005). Degraeve, Labro, and Roodhooft (2005) demonstrated the application of TCO in a case study of a manufacturer of electronic products. They compared alternative suppliers of standard electrical components such as resistors, transformers, and printed circuit boards. The most significant cost savings resulted from making better use of prices and discounts; furthermore, costs related to inventory, quality, ordering could also be reduced.

Monetary quantification is the key element of TCO. Attributes that are initially not expressed in a financial metric are "translated" into financial numbers, and financial data are aggregated into a summary measure (such as the cost per hour, cost per unit, or the net present value). Cost drivers can be at various levels, such as unit level (e.g., purchasing price, quality control cost when each item must be inspected), batch level (e.g., cost of creating a purchasing order, inspecting an order received), supplier sustaining level (e.g., cost of identification and certification of a supplier), and product or part sustaining level (e.g., cost of maintaining technical product information).

The setting of monetary quantification and aggregation of attributes is not limited to TCO. For example, investment appraisal techniques intend to capture the impact of alternative investment options in terms of their associated cash flows for the firm, and more advanced techniques such as real options can be used to model uncertainty (Haka, 2006; Verbeeten, 2006). Life-cycle costing aims to quantify and aggregate the diverse costs throughout the value chain and the product life cycle (Anderson & Sedatole, 1998; Dunk, 2004). Similarly, monetary quantification of the impact of decisions (including externalities) is the purpose of cost-benefit analyses supporting public policy decisions (Boardman, Greenberg, Vining, & Weimer, 2001).

However, monetary quantification of nonfinancial attributes can be problematic, at least for some attributes, and some of the "richness" of the nonfinancial information gets lost in translation (Galbraith, 1973; Israelsen, 1994; Chambers, 1996; Chapman, 1997; Lillis, 2002). But monetary quantification and aggregation are persuasive and likely to have a significant impact on the decision-making processes (Kadous et al., 2005). A financial metric draws attention, suggests accuracy, is easy to communicate, and makes comparisons of alternatives easy. Monetary quantification suggests a greater competence of the manager who prepared it and a higher subjective plausibility of a favorable outcome (Kadous et al., 2005). And because of this persuasiveness, decision makers could easily loose sight of nuances and even consider the nonfinancial information less credible. For example, activity-based costing may easily become a system that denies and annihilates anything which is not a routine activity and therefore less easy to standardize and count. But that is just inherent to many staff activities in, for example, human resources, R&D, and purchasing. Routine activities of the purchasing department (such as ordering items and making payments) are much easier to include in a costing system than more strategic activities of purchasing (such as supplier selection, developing relationships with important suppliers, participating in new product development). The introduction of costing systems may render such less-easy-to-quantify activities less credible (Armstrong, 2002). In the public sector, Skærbæk and Thisted (2004) describe some of the related problems that governmental agencies face when calculating unit costs.

The foregoing discussion suggests that there could be unintended effects of providing better – but not perfect – costing information. Nonetheless, little is known about decision-making processes that involve the combination of accounting information and other decision-relevant information. How much importance will decision makers attach to available and relevant information that is, however, inaccurately monetarily quantified and included in aggregated financial information, such as TCO? This question is relevant for many decision-making situations in which some form of monetary quantification and aggregation is used that, however, cannot encompass every important consideration.

Previous empirical studies in accounting have more generally demonstrated that characteristics of accounting information, the decision context, and decision makers have effects on decision-making processes. For example, it was found that complexity and amount of information provided impact decision-making strategies (Shields, 1980), decision quality (Isselin, 1988), and information selection and judgment accuracy (Shields, 1983). Groups may process information better compared to individuals (Stocks & Harrell, 1995). Financial information tends to carry more weight compared to nonfinancial information (Reck, 2001; Schiff & Hoffman, 1996). Recent studies showed that unique measures may receive less weight compared to common measures when comparing the performance of organizational units (Lipe & Salterio, 2000). This effect may also depend on the presentation of performance measures in categories (Lipe & Salterio, 2002), subjectivity of the performance measurement system (Ittner, Larcker, & Meyer, 2003), accountability (Libby, Salterio, & Webb, 2004), and on links between measures and the unit's strategy (Banker, Chang, & Pizzini, 2004).

Other studies have investigated factors contributing to the adoption of more accurate cost systems such as activity-based costing (Fullerton

& McWatters, 2004; Snead, Johnson, & Ndede-Amadi, 2005), and how more accurate cost information improves decision-making (Briers, Luckett, & Chow, 1997; Cardinaels, Roodhooft, & Warlop, 2004; Gupta & King, 1997; Harrison & Killough, 2006; Maiga & Jacobs, 2007). In sum, the accounting literature has shown that judgment and decision-making processes are central to understanding the impact of accounting information on the outcomes of decision-making processes. However, such processes are not well understood when it comes to multi-attribute decisions and the role of imperfect monetary quantification of attribute information that is initially nonfinancial.

We conduct several experiments to investigate the impact of reflective thinking and experience on the weight decision makers attach to attribute information that is not or inaccurately included in TCO-numbers (see Table 1). We expect that induced reflective thinking will stimulate decision makers to process the presented inaccurate cost information more carefully. As a result, we expect that decision makers will accordingly adjust the weight they attach to the inaccurately quantified attributes. This hypothesis is tested in experiments 1A and 1B. Experiment 2 was conducted with both students and practitioners to investigate the influence of professional experience.

We contribute to research in several ways. Firstly, we conceptualize three ways in which monetary quantification can be inaccurate because it ignores decision-relevant information. Secondly, we empirically investigate how reflective thinking affects the weight of the attribute that is inaccurately quantified. Thirdly, we investigate how this effect is moderated by experience of the decision maker.

Table 1. Experiments 1A and 1B Compared to Experiment 2.

	Experiments 1A and 1B	Experiment 2
Participants	Students	Students and practitioners
Purchasing information	Three alternatives, described by six attributes each	Two alternatives, described by four attributes each
Nature of the *Quantification* of the nonfinancial attribute	Minimum or maximum estimation	Excluded
Manipulation of *Reflection*	Participants are requested to think thoroughly before choosing	Participants have to justify their decision strategy after choosing

The remainder of this article is structured as follows. In the next section, we discuss various ways in which TCO-numbers may be inaccurate. The section ''Reflective Thinking'' describes the hypotheses, design, and results for the first two experiments (1A and 1B) and about the effect of reflection on the weight of an attribute that is inaccurately quantified as part of provided TCO-numbers. The section ''Moderating Effect of Experience'' describes experiment 2 and about the effects of experience and reflection on the decision weight of an attribute that is excluded from the provided TCO-numbers. As mentioned above, both experiments differ in several ways, and therefore we present the experiments subsequently, including the accompanying hypotheses and results. Concluding remarks are in the final section.

SOURCES OF ERROR IN MONETARY QUANTIFICATION

In some cases decision-relevant information about choice options can unambiguously be translated to its monetary cost equivalent. The TCO transformation will then be unbiased with respect to the original information relevant for the decision. However, often there is only a vague or probabilistic relationship between an attribute and its financial implications. The information provider may then choose one of three options, each resulting in an imperfect match between the given information and its TCO equivalent. They are summarized in Table 2.

One possibility is to consider only the subset of the information that is unambiguously translatable in cost terms, and ignore all other information. The more ambiguous implications, for which it is hard to include costs objectively, will be left out. Here, the monetary quantification of a nonfinancial attribute is kind of minimum estimate, which will be downwardly biased: *the financial implications of that attribute are only partially quantified and subsequently included in the aggregate financial TCO-number*.

Table 2. Monetary Quantification of the Nonfinancial Attribute.

Nature of the *Quantification* of the Nonfinancial Attribute	Effect of *Reflection*	Effect of *Experience*	Interactive Effect of *Reflection* and *Experience*
Minimum estimation (experiment 1A)	H1	–	–
Maximum estimation (experiment 1B)	H2	–	–
Excluded (experiment 2)	H3	H4	H5

For example, the downtime of production equipment may lead to costs for repair and extra labor costs, and these costs may be accurately quantifiable. Downtime may also lead to production orders being too late and, as a consequence, disappointed customers. Such costs may be much more difficult to estimate. This example has also been used in the experimental tasks (see Fig. 1). Note that we use uptime – the opposite of downtime – in this example (e.g., a downtime of 4.0% equals an uptime of 96.0%). Uptime is one of six attributes provided for alternative brands of a machine. Except for the uptime attribute, costs per hour numbers can be calculated

Experimental Task in Experiment 1 A
(Condition: *Uptime C* = 99%; no-reflection)

You are the manager in a company. One machine has to be renewed. You can choose between three brands, all three meet all specifications.

In the table below information is provided about:

- The lifetime of a new machine;
- The uptime percentage of the machine;
- The time an employee needs to check the machine;
- The purchasing price of the four component new machine;
- Energy consumption per hour;
- The amount of chemicals necessary.

The 'uptime (%)' is the percentage of the time the machine is available for production. 100% – 'up time (%)' = 'downtime (%)'. Downtime is caused by machine failure, maintenance, etc. that cause rescheduling of production, not meeting delivery times, etc. Some downtime cost, like failure cost, maintenance cost,etc. can be calculated objectively. Objectively calculable costs are included in the table below. Other downtime cost, like cost of not reaching times of delivery or additional transports are highly uncertain and therefore not included in the table below.

Labor cost per operator (€ /hour)	€ 30,00
Electricity price (€/Kwh)	€ 0,13
Chemicals (€/liter)	€ 10,00

		Brand A	Cost per hour	Brand B	Cost per hour	Brand C	Cost per hour
Life time (hours)		2.900		2.700		2.800	
Uptime (%)		99,0%	€ 1,01*	96,0%	€ 6,25*	99,0%	€ 1,41*
Inspection time (minutes/hour)		9,0	€ 4,50	11,0	€ 5,50	10,0	€ 5,00
Purchasing price (€):	Component A	€ 2.800	€ 0,97	€ 2.300	€ 0,85	€ 2.100	€ 0,75
	Component B	€ 1.700	€ 0,59	€ 1.400	€ 0,52	€ 1.300	€ 0,46
	Component C	€ 2.000	€ 0,69	€ 1.600	€ 0,59	€ 1.600	€ 0,57
	Component D	€ 1.300	€ 0,45	€ 1.100	€ 0,41	€ 1.000	€ 0,36
Energy (kwh/hour)		7,0	€ 0,91	6,0	€ 0,78	5,0	€ 0,65
Chemicals (liter/hour)		0,09	€ 0,90	0,09	€ 0,90	0,08	€ 0,80
Total cost per hour			€ 10,01		€ 15,80		€ 10,01

* Minimum downtime cost estimation

Indicate which brand you will buy: _____

Fig. 1. Experimental Task in Experiment 1A. (Condition: *Uptime* $C = 99\%$; No-Reflection.)

straightforwardly. The costs of downtime are only partially included, and the estimated cost of downtime is a minimum cost estimation, thereby downwardly biased. The effect of increasing reflection on the impact of this type of inaccuracy of monetary quantification is investigated in experiment 1A.

Second, the monetary quantification of a nonfinancial attribute can be a kind of maximum estimation, which is upwardly biased: *all possible financial implications of that attribute – also the ones that are ambiguous – are quantified and subsequently included in the aggregate financial TCO-number*. In the example, besides the more straightforward costs of downtime of production equipment (such as repair and extra labor) some more uncertain – but possible – costs associated with downtime (such as production orders being too late and, as a consequence, disappointed customers) can also be quantified and included in the TCO information. The monetary quantification of this attribute represents an estimation of the maximum cost of this attribute and is, therefore, upwardly biased. The effect of increasing reflection on the impact of this type of inaccuracy of monetary quantification is investigated in experiment 1B.

Third, when the monetary quantification of a nonfinancial attribute is difficult and inherently inaccurate, such an attribute can also be *not translated in monetary terms at all*. The relevant information is then completely left out of the aggregate financial number, and should be traded off by the decision maker against the information that is included in the TCO-number. In experiment 2, costs for downtime are not quantified and not included in the presented TCO-numbers. Fig. 2 shows an example of the experimental task used in experiment 2.

In summary, we investigate three forms of inaccurate quantification of the nonfinancial attribute within TCO information: a nonfinancial attribute might be (1) included as minimum costs, (2) included as maximum costs, or (3) excluded from the costs numbers. These forms are indicated in Table 2. Each of these inaccurate quantifications will reduce the quality of the decision, but we propose that their impact will be reduced if decision makers are motivated or are able to think more reflectively about the decision problem.

REFLECTIVE THINKING

This section contains the motivation of hypotheses, experimental design, and results of first experiment. On the basis of these results, the second set of hypotheses and findings will be presented in the next section.

Experimental Task in Experiment 2

(Condition: *Uptime B* = 96%; no-reflection)

You are the manager of a production department. One machine has to be renewed. You can choose between two brands, both meet all specifications.

In the table below information is provided about: the purchasing price of a new machine, the lifetime of a new machine, energy consumption per hour, and the uptime percentage of the machine. The 'uptime (%)' is the percentage of the time the machine is available for production. 100% – 'uptime (%)'= 'downtime (%)'. Downtime is caused by machine failure, maintenance, etc. that cause rescheduling of production, not meeting delivery times, etc. Both machines only differ on the four attributes mentioned below; the machines are equal on all other attributes.

Energy cost (Kwh) € 0.13

	Brand A	Cost per hour	Brand B	Cost per hour
Life time (hours)	2,900		2,800	
Uptime (%)	99.0%		96.0%	
Purchasing price (€)	€ 1,280	€ 0.44	€ 1,020	€ 0.36
Energy per hour (Kwh)	2.0	€ 0.26	2.2	€ 0.29
Total cost per hour		€ 0.70		€ 0.65

Indicate which brand you will buy: _____

Fig. 2. Experimental Task in Experiment 2. (Condition: *Uptime* $B = 96\%$; No-Reflection.)

Hypotheses Development

Human decision makers often adopt a strategy of least effort, and they may not be willing or not able to adopt a demanding cognitive strategy such as specified in normative models (Dewey, 1933; Simon, 1979; Tetlock, 1985). Reflective thinking may enhance the quality of decisions (Baron, 1981; Langer, 1978). It may help decision makers to judge the meaning, relevance, and quality of the available information. Such judgments may help decision makers to get more insight in the data, and bring together facts that seem to be incoherent and disconnected (Dewey, 1933). Empirical studies in accounting and auditing (e.g., Johnson & Kaplan, 1991; Kennedy, 1993; Libby et al., 2004) and in psychology (e.g., Pryor, Gibbons, Wicklund, Fazio, & Hood, 1977; Koriat, Lichtenstein, & Fischhoff, 1980; Snyder & Kendzierski, 1982) found that reflective thinking may improve the quality of decision-making. The quality of decision-making has been measured in various ways, for example, by less discrepancy between answers and actions (Pryor et al., 1977), greater predicted correctness of answers (Koriat et al., 1980), higher correspondence between attitudes and behavior (Snyder & Kendzierski, 1982), and reduction of recency effects (Kennedy, 1993).

We expect that reflection will increase the awareness of the inaccuracy of monetary quantification included within TCO information. Decision makers who are motivated to reflect, will be more conscious of the different attributes involved in the purchasing decision-problem, and they are more likely to think through and weight the information carefully. The implication is that including nonfinancial attributes as minimum or maximum costs in TCO information would result in opposite effects of reflection. Reflective decision makers may attach *more* weight to attributes that are monetarily quantified within the TCO-number as minimum costs, while they may attach *less* weight to attributes that are monetarily quantified as part of the TCO-number as maximum costs. Minimum cost estimations are likely to be underestimations, so the relative weight of the inaccurately quantified attribute will increase if a decision maker considers that the costs can actually be higher. The opposite will be the case if costs are estimated as maximum costs; these numbers are likely to be overestimations, and so the relative weight of the attribute may decrease when the decision maker takes this estimation error into account.

H1. When the monetary quantification of an attribute is a minimum estimation (downwardly biased), reflection will *increase* the weight of that attribute included in a TCO-number.

H2. When the monetary quantification of an attribute is a maximum estimation (upwardly biased), reflection will *decrease* the weight of that attribute included in a TCO-number.

H1 will be tested in experiment 1A, and H2 will be tested in experiment 1B (Fig. 3).

General Procedure Description

We first discuss the common procedural aspects of all experiments in the article. In all experiments, the participants assume the role of a decision maker within a production company, who has the task of purchasing one of several production machines. One option (Option C in experiments 1A and 1B, Option B in experiment 2) was the target product, and two levels of the target attribute "uptime" were experimentally manipulated in the target option. For half of the participants, the uptime was presented as 99%, for

Summary of the Hypotheses

H1-H2 concern monetary quantification of all attributes, but for one attribute this quantification is either an estimation of minimum costs (H1), or an estimation of maximum costs (H2):

H1, 2: Reflection → Attribute weight

H3-H5 concern monetary quantification whereby one attribute is not quantified at all and therefore not included in the TCO information:

H3: Reflection → Attribute weight

H4: Experience → Attribute weight

$$\text{H5:}\quad \begin{array}{ccc} & \text{Experience} & \\ & \downarrow & \\ \text{Reflection} & \longrightarrow & \text{Attribute weight} \end{array}$$

Fig. 3. Summary of the Hypotheses.

the other half it was presented as 96%. Obviously a machine with 99% uptime carries lower cost than a machine with 96% uptime, and this should be reflected in a higher likelihood of choosing machine C. The difference between the two choice probabilities can be considered as an indication of the weight of that attribute level difference in the choice, akin to the common research method of conjoint analysis. Operationally, we therefore predict a statistical interaction between the variables *Uptime* (96% versus 99%) and *Reflection*.

Experiment 1A

Participants were undergraduate students who were familiar with the basic concepts of cost accounting. They were provided with three alternative purchasing options from which they had to select one. These options concerned three alternative brands for production equipment. Information was presented on the same attributes for each brand. Some information was financial (i.e., expressed in Euros) and other information was quantitative nonfinancial. The dependent variable was the choice for the target alternative (*Choice C*). Fig. 1 shows the text introducing the experimental task for one of the conditions. The uptime percentage of Brand C (*Uptime C*) was an independent variable at two levels (99% and 96%), manipulated between subjects.

The second independent variable was reflection on the information provided (*Reflection*), which was manipulated at two levels (no-reflection, reflection), using a classic "accountability" manipulation (Tetlock, 1983). Participants in the no-reflection conditions were only asked to choose one of three alternatives. Participants who were encouraged to reflect were asked to answer the following two questions *before* choosing one alternative:

Please, explain below how you make a choice for one of the three brands and which trade-offs you make.
Please, explain below to what extent you are able, on the basis of the information provided above, to make a choice without further analysis.

In experiment 1A, TCO information was provided as an estimate of the *minimum costs associated with either 96 or 99% of uptime*. Participants learned that TCO-numbers for downtime were difficult to estimate. The note below the table read "Minimum downtime cost estimation", and the introductory text included the following explanation:

> Some downtime cost, like failure cost, maintenance cost, etc. can be calculated objectively. Objectively calculable costs are included in the table below. Other downtime costs, for example costs of not reaching delivery times or additional transports, are highly uncertain and are therefore not included in the table below.

The interaction of the two independent variables *Uptime C* and *Reflection* allows testing H1 based on a logistic regression with the following specification:

$$\mathit{Choice\ C} = \beta_1 + \beta_2\ \mathit{Uptime\ C} + \beta_3\ \mathit{Reflection} + \beta_4(\mathit{Uptime\ C} \times \mathit{Reflection})$$

If coefficient β_4 is significant and positive, then the effect of *Uptime C* on the dependent variable will be stronger if participants have to reflect compared to when participants do not have to reflect. A negative coefficient indicates the opposite: a lower weight.

This between-subjects experiment was conducted in a laboratory setting. Experimental conditions were randomly assigned to participants. The experimental task (each printed on one A4-sheet of paper) was distributed on a sheet of a paper to each participant. Participants were not compensated for participating in the experiment, and each participant participated only once in the experiment.

Table 3. Total Number of Respondents and the Choices for the Target Option per Condition in Experiments 1A and 1B.

		Experiment 1A		Experiment 1B	
Quantification		Minimum estimation		Maximum estimation	
Reflection		No-reflection	Reflection	No-reflection	Reflection
Uptime C	96%	15 (4)[a]	14 (1)	35 (7)[b]	35 (4)[c]
	99%	14 (8)	14 (12)	37 (27)	38 (24)
Total		29	28	72	73

[a]Between parentheses: number of respondents choosing Brand C, from Brands A, B or C.
[b]Two participants chose Brand B in this condition (not included in the numbers in the table). Note that Brand B is inferior to Brands A and C, and it was only included to increase the number of alternatives to three. As expected, almost no participants chose Brand B.
[c]One participant chose Brand B in this condition (not included in the numbers in the table).

Results and Discussion, Experiment 1A

Table 3 shows the number of participants and their choices in the various conditions of the first experiment.

Results related to H1 are in Fig. 4a and in Table 4. To reiterate, we expected that if, uptime was estimated as a *minimum* costs, the impact of uptime on the preference for the target Brand C would be *larger* if participants are encouraged to reflect, compared to when they only have to choose. Fig. 4a shows that, consistent with this hypothesis, the effect of *Uptime C* on *Choice C* was moderated by *Reflection*. The coefficient for the interaction term was significant ($p = .043$).[1]

Experiment 1B

In experiment 1B, the procedure paralleled the procedure for experiment 1A. Participants were again undergraduate students from the same subject pool as in experiment 1A. Choices were again among three machines, and Option C was again the critical option. The target attribute was again uptime, manipulated at 99 and 96%. In this experiment TOC information is a maximum estimate, by including all information that might potentially be relevant. 145 students participated for course credit. The experimental task

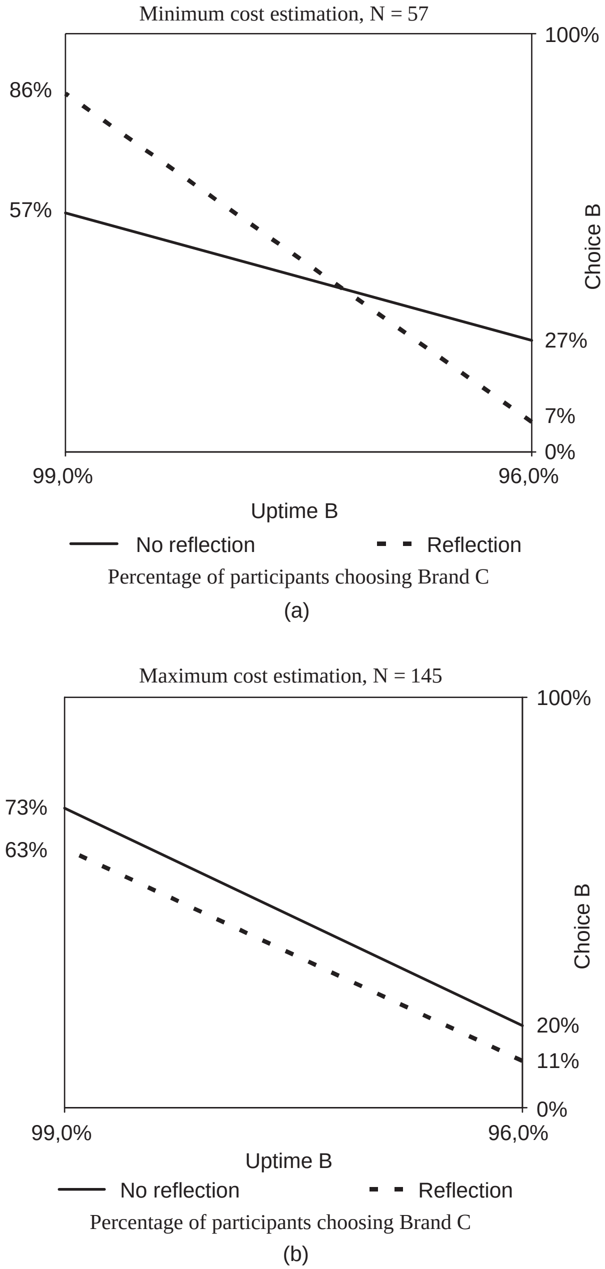


Fig. 4. Interaction of *Reflection* by *Uptime C*, Experiments 1A and 1B. (a) Minimum Cost Estimation, $N = 57$. (b) Minimum Cost Estimation, $N = 145$.

Table 4. Logistic Regression Results Experiments 1A and 1B.

Choice C is the Dependent Variable	Experiment 1A	Experiment 1B
	$N = 57$	$N = 145$
Constant	−.373[a]	−.155
	(.349)	(.578)
Uptime C	.015	.028
	(.102)	(.000)
Reflection	−.051	−.554
	(.946)	(.184)
Uptime C × *Reflection*	.036	.002
	(.043)	(.806)

[a]Values of coefficients (and *p*-values for two-tailed Wald test).

was the same as in experiment 1A, except for the note placed below the table, which was formulated as "Maximum downtime costs estimation". Furthermore, the sentence on the estimation of downtime cost was different

> Other downtime costs, for example costs of not reaching delivery times or additional transports, are highly uncertain. The number in the table below is an estimation of the maximum downtime costs that might occur, however these might be lower.

Reflection should reduce the effects of the upward bias in the choice situation (H2).

Results for H2 are in Fig. 4b and Table 4. This hypothesis is not supported. We expected that if, as part of the TCO information, uptime was provided as an estimation of *maximum* costs, the impact of uptime on the preference for Brand C would be *smaller* if participants are encouraged to reflect compared to when they only have to choose. Fig. 4b suggests that the weight was the same whether or not they were encouraged to reflect. The coefficient for the interaction term between *Reflection* and *Uptime C* was not significant ($p = .806$).

Discussion

The difference between results under the minimum versus maximum conditions is relevant to discuss further. Representing costs as a minimum or maximum implies that decision makers have to make a decision under uncertainty; "real" costs might deviate from the numbers provided in the tables. Prospect theory (Kahneman & Tversky, 1979) describes how human decision makers attach more weight to negative consequence (losses) than to the equivalent positive consequences (gains). Reflective decision makers considering a minimum cost may realize that costs could also be higher, which may be experienced as a relative loss, and this information is likely to carry a high weight in the decision. On the other hand, a maximum cost representation indicates that "it can only get better", which may be experienced as a relative gain, and this information is likely to carry a smaller weight compared to losses. Therefore, decision makers could have attached relatively more weight to a cost attribute represented as a minimum compared to costs represented as a maximum.

Additionally, the differential results for the two experiments may be understood in terms of the implied ranges of possible outcomes. Formulating costs as a minimum implies a larger range of possible outcomes compared to formulating costs as a maximum. Decision makers tend to give a higher weight to larger ranges (Fischer, 1995). For example, the downtime cost of Brand C in Fig. 1 is € 1.41. When this is an estimation of the minimum cost, the range is, in principle unlimited: € 1.41 to infinity. When this is an estimation of the maximum cost, the range is € 1.41 (at most € 1.41 and at least € 0). Decision makers are therefore likely to attach more weight to the downtime cost in case these costs are represented as a minimum compared to when these costs are represented as a maximum.

Experiments 1A and 1B suggest that reflection matters for the way in which TCO-numbers are used to support purchasing decisions, and also the type of inaccuracy of such TCO-numbers may have an impact. However, the findings are based on experiments pertaining to business decisions taken by students. Therefore, a logical next step is to involve more experienced research participants. This also raises the question of how reflective thinking may be achieved in real-life situations, and in the next experiment we will create more business-like conditions to invoke more careful and reflective decision-making.

MODERATING EFFECT OF EXPERIENCE

In experiment 2, we investigate the effect of reflective thinking also for experienced participants. We first held roundtable discussions with practitioners to discuss the experimental task and its administration. From these discussion emerged that a commonly experienced inaccuracy by practitioners was the exclusion of a subset of the information from the TCO-number.

Also because of the involvement of experienced participants, the manipulation of reflective thinking in the experiment was changed and more reflective of practical demands on accountability. In the previous experiment, *Reflection* was manipulated by simply asking participants to explicate how they arrived at their choice (beforehand), and this was expected to induce sufficient reflective thinking to improve their understanding of the information provided. However, in real business settings, decision makers may need "stronger" incentives to encourage reflective thinking, such as the expectation that they will have to justify their decisions. Therefore, we manipulated *Reflection* by creating a setting of decision justification (Curley, Yates, & Abrams, 1986; Tetlock, 1992). Organizational buyers have to be able to provide analysis and give a formal justification, which is quite different from how consumers are accountable for their decisions (Sherlock, 1991). Being able to give justifications is also important for managers, and managerial success is related to being able to convince others (Tetlock, 1985).

Hypotheses Development

The three hypotheses that will be discussed in this section are also represented in Fig. 3. We first motivate a hypothesis for a main effect of reflective thinking on the weight of the attribute that is not included in the TCO-numbers provided to decision makers (H3). Decision makers may want to avoid "loosing face" when they feel they have to justify their decisions, and this may change their decision processes (Carnevale, 1985). This may be even more so in threatening situations; if the decision maker's explanation is monitored by others and mistakes may result in sanctions (Fox & Staw, 1979). Therefore, decision justification may motivate decision makers to adopt a strategy that will result in more profound analyses and evaluations (Tetlock, 1983). Empirical studies found that decision justification may lead to greater effort (Kennedy, 1995; Koonce, Anderson, & Marchant, 1995; Doney & Armstrong, 1996) and better decisions (Kennedy, 1993; Johnson & Kaplan, 1991).

Therefore, decision makers who are encouraged to take a closer look at the aggregate cost information and reflect on their decision-making may attach more weight to the information that is not included in the TCO-numbers. These decision makers may realize that attributes excluded from the TCO-numbers may raise the total costs of an alternative, and should be traded-off against the costs that are included within the TCO-numbers.

However, reflective thinking in a social setting might also result in the opposite effect – it is a double-edged sword. Instead of thinking through the issues more carefully in order to arrive at a justifiable decision, a decision maker might simply take the decision that will satisfy the person to whom he has to justify his decision, without much thinking (Tetlock, 1983; Lerner & Tetlock, 1999). This requires that a decision maker knows (or can readily infer) the beliefs and preferences of his evaluators. Empirical studies in auditing have found such effects (e.g., Peecher, 1996; Haynes, Jenkins, & Nutt, 1998; Wilks, 2002): auditors who were informed about the preferences of their supervisors or the interests of their clients before making a decision, shifted in their decision making towards decisions which satisfy these stakeholders.

Therefore, reflective thinking may result in focusing mainly on the aggregate cost numbers and a tendency to ignore attributes which are not included in these cost numbers. Financial information may be seen as particularly important in business (Brierley, Cowton, & Drury, 2006), and decision makers will realize that TCO-numbers are not "automatically" produced. TCO information is presumably produced (when top) management wants it to be used for decisions – and it probably includes what top management thinks it is important to consider. In addition, decision makers may anticipate the difficulty of explaining a subtle multi-attribute trade-off between different attributes to their superiors, which may lead them to follow a decision strategy that is more easily explainable. As a result, decision makers may ignore attributes for which no monetary quantification is included in the TCO-numbers provided.

In sum, we posit that reflective thinking influences the weight attached to the attribute not included in the TCO-numbers. However, because of the contradictory findings we predict no directional effect. Thus, we hypothesize

H3. When an attribute is not monetarily quantified and, hence, not included in a TCO-number, reflection will influence the weight of that attribute.

Experience may moderate the effect of *Reflection* (H5), but first we will hypothesize a main effect of experience (H4). Experience in this study refers to general business experience or professional experience, and not to specific experience with and knowledge of accounting or purchasing tasks. We assume that professionally experienced people have a better understanding of operations, technology, and the usage of cost information in organizations; the kind of knowledge related to general business experience (Bonner & Lewis, 1990). For example, knowledge obtained during college, working in business, reading, and other kinds of individual life experiences. However, the interacting effects of reflective thinking and experience may be subtle (e.g., Tan & Kao, 1999; Tan, Ng, & Mak, 2002).

Experienced decision makers are likely to be more mindful of the inaccuracies in the provided data. Experience may result in expertise, defined as a general ability to solve domain-specific problems (Alba & Hutchinson, 1987). For example, more experienced decision makers appear better able to distinguish relevant from irrelevant information (Shelton, 1999). Experience may help a decision maker to distinguish what information is relevant and what information can be excluded from the decision-making process (i.e., whether or not the attribute that is not included in the cost numbers should be taken into account). Broad domain knowledge (e.g., knowledge of basic accounting principles that accounting professionals have who are working in specialized accounting areas), may help them to recognize problems and to further investigate relevant information, even if they are not directly familiar to a specific topic (Vera-Muñoz, Kinney, & Bonner, 2001). Experience is especially helpful if tasks become more complex; and general training (such as education in business administration) and experience helps decision makers to outperform less experienced decision makers (Chang, Ho, & Liao, 1997). Experience may help decision makers to assess the completeness of the provided cost information. More experienced decision makers should better understand that some attributes are not included in these cost numbers, and they may better realize that it is important to make a deliberate trade-off between the TCO-numbers and the nonincluded attributes. As a result, experienced decision makers could be expected to attach *more* weight to the attribute not included in the TCO-numbers, compared to less experienced decision makers.

However, experience may also have the opposite effect. It may be associated with habit, routine, automaticity, and superficial thinking. Several studies found unfavorable effects of experience. Experienced decision makers may find it difficult to adopt another strategy, and they may unconsciously adopt the same decision strategy over and over again unless something stops them from doing this (Alba & Hutchinson, 1987). For example, Marchant, Robinson, Anderson, and Schadewold (1991) found dysfunctional effects of high-accounting knowledge as a result of inflexibility to change decision strategies. Also, a task that deviates from the task structure experts are used to might make it difficult for experts to excel

(Nelson, Libby, & Bonner, 1995). In another study (Vera-Muñoz, 1998), it was found that information that is highly relevant according to economic theory was ignored by more experienced decision makers, because they were used to analyze information by decision rules that attached less importance to this particular type of information. These studies indicate that expert knowledge might result in inflexible decision strategies. These studies make us expect that decision makers could mainly focus on the TCO-numbers and show a tendency to ignore information not included in TCO-numbers. After all, financial information is important and caries a lot of weight in decisions (Schiff & Hoffman, 1996; Reck, 2001).

No directional effect is predicted, because these contradictory empirical findings do not indicate whether experienced decision makers are likely to attach more or less weight to information not included in TCO-numbers. Therefore, we formally hypothesize

H4. When an attribute is not monetarily quantified and, hence, not included in a TCO-number, experience will influence the weight of that attribute.

The idea behind H4 is that more experienced decision makers have a more "outspoken" decision strategy. This leads them to either pay less attention to attributes not included in the TCO info, or to pay more attention to these (hence, we made no directional prediction for H4). However, compared to students, we expect practitioners to change their decision strategy (and the weight they give to uptime) less as a result of reflective thinking: *Experience* may moderate the effect of *Reflection* (H5) (see Fig. 3). The proposed result is similar to findings reported by Kennedy (1993). She found that reflective thinking reduced an inexperienced decision maker's bias, whereas she did not find such bias in experienced decision makers' decisions.

So, suppose the main effect of experience is such that these decision makers give *more* weight to the attribute not included in TCO, because decision makers with business experience may have learned that not all relevant information is accurately included in cost numbers, and they are reluctant to ignore information that is not included in cost numbers. For these decision makers, we would expect the effect of reflective thinking to be smaller than for students. And suppose the main effect of experience is such that decision makers with business experience give *less* weight to the attribute not included in TCO, because they realize that top management wants purchasers to use TCO-numbers, and these numbers are then also likely to include what management thinks is important. If reflective thinking of these decision makers is stimulated, they are less likely to change their decision strategy and give another weight to excluded attributes, compared to students. To summarize, we formally posit the following hypothesis:

> **H5.** When an attribute is not monetarily quantified and, hence, not included in a TCO-number, experience will negatively moderate the impact of reflection on the weight of that attribute.

Research Method

As in experiments 1A and 1B, participants received information on several purchasing alternatives from which they had to select one. Inaccuracy of cost information in this experiment meant that for the attribute uptime no cost estimation was included in the TCO information (see Fig. 2), instead of the minimum cost and maximum cost scenarios in 1A and 1B.

The dependent variable was again the choice for a particular alternative, for Brand B in this case (*Choice B*). A 2×2 between-subjects design was used, and *Experience* was added as a measured variable (students versus practitioners). The first independent variable was the uptime percentage of Brand B (*Uptime B*), set at two levels (99.5% and 96.0%).

The level of reflection (*Reflection*) was the second independent variable, and it was manipulated at two levels (no-reflection versus reflection). Participants in the no-reflection condition participated anonymously and were only asked to choose one alternative. These participants received a one-sided printed sheet on which the experimental task was printed (the other side was blank). Participants in the reflection condition received a double-sided printed sheet. Participants had to fill-out their name, e-mail address, and phone number on the front page (the page that was blank in the no-reflection condition). Students could read on the front page that we needed those data to organize a meeting where we would discuss their choices. Practitioners could read on the front page that we needed their contact information to call them for additional questions concerning their choice. Practitioners in the reflection condition were also asked to think aloud while conducting the experimental task. Participants where randomly assigned to reflection or no-reflection condition.

Experience was included as a third independent variable, measured at two levels. Some participants ($N = 94$) were undergraduate students similar to the first experiment, and another group of participants ($N = 60$) consisted of practitioners who were visitors and exhibitors at a large international maritime trade fair in Rotterdam. The median years of working experience was 18.5 years, and practitioners were on average 42 years old. Each participant participated only once in the experiment.

175 students were asked to voluntarily participate in this experiment, after finishing an exam in the same room. They were randomly assigned to one of the conditions, but the willingness to participate turned out to be lower in the reflection conditions, yielding 68 and 26 participants, respectively. Practitioners were contacted at a trade fair. Visitors of these trade fairs were asked if they were willing to answer a short question for a research project. They received a pen and the experimental task on one sheet of paper clamped on a clipboard. The practitioner completed the task in the presence of the researcher, at a quiet spot in the same room in which the participant had been contacted. If requested, the researcher explained the research project after completing the task. The researcher kept searching for a practitioner who was willing to answer a randomly selected condition before changing to a new randomly selected condition, yielding 34 and 26 participants, respectively. Some practitioners who were initially willing to participate refused to continue the experiment after learning that they had to provide contact details and possible had to answer questions afterwards (in the reflection condition).

Hypotheses H3–H5 are tested using logistic regression with the following specification:

$$\begin{aligned}\textit{Choice B} = {} & \beta_0 + \beta_1\ \textit{Uptime B} + \beta_2\ \textit{Reflection} + \beta_3\ \textit{Experience} \\ & + \beta_4(\textit{Uptime B} \times \textit{Reflection}) + \beta_5(\textit{Uptime B} \times \textit{Experience}) \\ & + \beta_6(\textit{Reflection} \times \textit{Experience}) \\ & + \beta_7(\textit{Uptime B} \times \textit{Reflection} \times \textit{Experience})\end{aligned}$$

The coefficient β_4 of the interaction term of *Uptime B* $\times$ *Reflection* is estimated for testing H3: A significant coefficient β_4 indicates that decision makers attach a different weight to *Uptime B* if they are encouraged to reflect, compared to decision makers who are not encouraged to reflect. The coefficient β_5 of the interaction term of *Uptime B* $\times$ *Experience* is estimated for testing H4: A significant coefficient β_5 indicates that professionally experienced decision makers attach a different weight to *Uptime B*, compared to students. Finally, the coefficient of the interaction term *Uptime B* $\times$ *Reflection* $\times$ *Experience*, β_7, is estimated for testing H5, because keeping *Reflection* constant, a significant value β_7 would indicate that *Reflection* has a significant influence on the interaction of *Uptime B* $\times$ *Experience*. That is, the weight students and practitioners attach to the attribute not included in the TCO-numbers is different as a result of reflective thinking.

Results and Discussion

Table 5 shows the number of participants and their choices in the various conditions of the second experiment.

Results for H3–H5 are shown in Table 6. In column 1, the coefficient of *Uptime B* $\times$ *Reflection* is not significant ($p = .950$), hence, we find no support for our hypothesis H3. In column 2, the coefficient for the two-way interaction term *Uptime B* $\times$ *Experience* was significant ($p = .020$), hence, these results support hypothesis H4. As Fig. 5 shows, students put a higher weight on the attribute uptime, which was not included in the TCO-number in this experiment, compared to practitioners. In column 3, the coefficient for the interaction term *Uptime B* $\times$ *Reflection* $\times$ *Experience* was not significant ($p = .759$), hence, we find no support for hypothesis H5.

The lack of support for the effect of reflective thinking in this experiment is surprising. We assumed that reflective thinking would change the weight of the attribute that was not included in the TCO-numbers (H3). One possible explanation could be that the experimental task may have been too simple. The inaccuracy may have been too obvious to ignore, regardless of the manipulation of reflection. Maybe participants made the same trade off in all conditions between total costs and the attribute that was not included in this TCO information. A higher task complexity, for example, through more attributes per alternative, might have been required to see an effect of reflection. For a more difficult task, decision makers who are not pushed to think more

Table 5. Total Number of Respondents and the Choices for the Target Option per Condition in Experiment 2.

Quantification		Excluded			
Reflection		No-reflection		Reflection	
Experience		Students	Practitioners	Students	Practitioners
Uptime B	96.0%	35 (11)[a]	16 (7)	11 (4)	12 (4)
	99.5%	33 (30)	18 (12)	15 (14)	14 (10)
Total		68	34	26	26

[a]Between parentheses: number of respondents choosing Brand B.

Table 6. Logistic Regression Results Experiment 2.

Choice B is the Dependent Variable	Inaccurate TCO		
	$N = 154$		
	H3[a]	H4[b]	H5
Constant	.509[c]	.897	.239
	(.032)	(.004)	(.500)
Uptime B	.022	.032	.010
	(.000)	(.000)	(.183)
Experience		−.696	.582
		(.091)	(.250)
Reflection	.004		−.096
	(.992)		(.862)
Uptime B × *Experience*		−.019	.022
		(.020)	(.033)
Uptime B × *Reflection*	.001		.007
	(.950)		(.549)
Experience × *Reflection*			.377
			(.676)
Uptime B × *Experience* × *Reflection*			−.006
			(.759)

[a]A specification including *Experience* as an independent variable was also estimated. Because the χ^2 difference was not significant ($\chi^2_d = 7.817$, df $= 4$, $p = .099$), we estimated the coefficient for *Uptime B* × *Reflection* using the empirical specification provided in the first column in the table (cf. Pampel, 2000).
[b]A specification including *Reflection* as an independent variable was also estimated. Because the χ^2 difference was not significant ($\chi^2_d = .572$, df $= 4$, $p = .966$), we estimated the coefficient for *Uptime B* × *Experience* using the empirical specification provided in the second column in the table.
[c]Values of coefficients (and *p*-values for two-tailed Wald test).

carefully might ignore the inaccuracy in the TCO-numbers and attach less weight to TCO-numbers, compared to decision makers who do not reflect.

Another explanation could be that selection bias in the experimental design may have countered the effect of reflective thinking. Participants were asked to provide their contact details, as explained above. Their responses were excluded if they refused to provide such information. As can be seen from Table 5, the number of participants in the reflection conditions is smaller than in the no-reflection conditions. As a result of self-selection, it could be that the resulting participants in the reflection conditions were more confident or were otherwise not particularly concerned about how their decisions would be judged. In other words, for those participants we

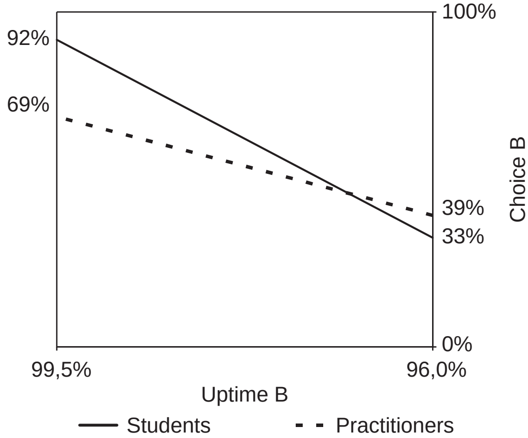


Fig. 5. Interaction of *Experience* by *Uptime B*, Experiment 2, $N = 154$.

may have failed to actually create reflective thinking and anxiety for losing face. However, this limitation is not exclusive to our study, although few studies of reflective thinking discuss it in great detail. For example, in reflective thinking studies conducted by Antonioni (1994) and Schwartz, Chapman, Brewer, and Bergus (2004), the number of participants who were stimulated to reflect was lower compared to the number of participants who were not stimulated to reflect. As in our study, participants could drop out of the experiment at any time, and they did more so if they were motivated to reflect because of being accountable for their decisions. Antonioni (1994) asked subordinates in an insurance company to evaluate their managers. Participants admitted that they preferred anonymous appraisals because reprisal might be costly. Participants in the study by Schwartz et al. (2004) were family physicians who selected a patient treatment. Reflection was stimulated by asking the physicians for a written justification on their decision, which also disclosed the research participant's name. In both studies, it is possible that self-selection factors were stronger in the reflection condition than in the no-reflection condition.

CONCLUDING REMARKS

We investigated decisions that are informed by information on the TCO of purchasing alternatives. TCO aims to provide a monetary quantification of points of difference of such alternatives, and this enables comparisons between noncommensurate differences in alternatives. However, a "perfectly" accurate monetary quantification of all attributes is unrealistic. We considered three ways in which an attribute may be inaccurately included in the TCO information: as a minimum cost, a maximum cost, or not at all. We investigated the weight of such an attribute in purchasing decisions, and how this depended on reflective thinking and experience.

Purchasing decisions such as those described in this article are cognitively challenging. The alternatives are described by several attributes; some are financial, whereas other attributes are nonfinancial and describe, for example, quality. Humans adapt their decision-strategy to the specific situation and try to minimize cognitive effort (Payne, Bettman, & Johnson, 1993). Relying on TCO info is tempting, because monetary quantification facilitates trade-offs among attributes that are noncomparable, as the measurement units of these attributes were initially more difficult to compare. However, monetary quantification may often not be capable to fully capture the operational considerations, and "decisions can no longer be taken and actions can no longer be chosen at a distance via the abstract language of numbers. Operational considerations must to some extent be involved in the decision process, not coded, but in their original form" (Chapman, 1997, p. 202). We investigated whether reflective thinking and experience affected how decision makers considered the inaccuracy of the TCO info, and, hence, how much weight they gave to the inaccurately quantified attribute.

Participants in our experiments had to choose between alternative offerings for replacing a production machine, each characterized by several attributes. One attribute (downtime of the machine) was problematic in terms of providing a monetary quantification. In some conditions in the first experiment, the cost associated with downtime was included in the TCO-numbers as minimum cost estimation. As hypothesized, it was found that the weight attached to the inaccurate included TCO-numbers increased as a result of reflective thinking. In other conditions, the maximum cost associated with downtime was estimated and included the TCO-numbers. We found no support for the hypothesis that the weight attached to this attribute would be reduced as a result of reflective thinking. In the second experiment, downtime was completely excluded from the TCO calculation. We found no support for the hypothesis that reflection would affect the weight of this attribute. We found strong support for the hypothesis that professional experience would affect that weight of this attribute. No support was found for an interactive effect of experience and reflection.

The results of this study may have managerial implications for the introduction and use of TCO information, and other costing information that aims to better capture the indirect costs of cost objects, but at the same time leaves out some important characteristics of these objects. Improved costing information that captures financial impact and aggregates financially quantified attributes may be helpful for the decision maker, but care should be given to unintended affects – that those elements that are not included in the new costing information get less attention as a result of providing the information.

The results suggest that when introducing cost information, it is important that the limitations should be explained very clearly to managers, so that they will take a closer look at the accuracy of the cost information. This seems to be especially important in case TCO-numbers are likely to be underestimated; without additional information, decision makers may tend to ignore the inaccuracies in these cost numbers. It is probably also important to give a prominent position in costing reports to the excluded attributes. This might attract attention and result in judgment and decision-making where decision makers take these results into account.

Several limitations have to be mentioned. Conducting experiments on computers is preferable to running experiments by paper. A computer would have provided the opportunity to measure process variables such as decision time. Although we are aware of this limitation, we choose to conduct experiments on paper for several reasons. We wanted to include practitioners in our experiments. Although it is not too difficult to run experiments with students in a laboratory, it is much harder too find a group of practitioners. For that reason, we did not ask practitioners to come to our laboratory, but decided to go to locations where we would find many practitioners from many different companies. We used paper because we were not able to use computers on the trade fair. As we wanted to minimize the impact of different methods, we choose to conduct the experiments with students and practitioners on paper. Despite the fact that we had the intention to run all experiments with both students and practitioners, the difficulty to find enough practitioners (i.e., get access to trade fairs), made us decide to conduct only the second experiment with practitioners.

Future research could help to get a better understanding of the impact of TCO information on in judgment and decision-making. The impact of a social context on judgment and decision-making is an interesting topic for a new study. In this study, individual in judgment and decision-making has been investigated. However, humans do not live in an isolated world, they live and work in setting where they interact with each others (Katz & Kahn, 1978); future research could therefore research the use of inaccurate costing information in relation to reflective thinking in a social context. We could also conceptualize inaccuracy of TCO-numbers differently: the monetary quantification of some attributes may be provided as a range, which may affect judgment and decision-making differently than investigated here (Fischer, 1995).

NOTE

1. We did not include a manipulation check on the accountability/reflection manipulation. We can therefore not be certain that a failure to find a predicted effect of reflection would not be due to a weak or inappropriate manipulation. However, we did obtain the predicted effect of reflection in this condition, which parsimoniously can only be explained by an effective manipulation. See also Sawyer, Lynch, and Brinberg (1995) and O'Keefe (2003) for a discussion about the value of including manipulation checks in experiments.

ACKNOWLEDGMENTS

The authors would like to thank Eddy Cardinaels, the Reviewers, participants at the AAA Management Accounting Section Midyear Meeting, 2007 and at the University of Arkansas for their useful suggestions.

REFERENCES

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research*, *13*(4), 411–454.

Anderson, S. W., & Sedatole, K. (1998). Designing quality design into products: The use of accounting data in new product development. *Accounting Horizons*, *12*(3), 213–233.

Antonioni, D. (1994). The effects of feedback accountability on upward appraisal ratings. *Personnel Psychology*, *47*(2), 349–356.

Armstrong, P. (2002). The costs of activity-based management. *Accounting, Organizations and Society*, *27*(1–2), 99–120.

Banker, R. D., Chang, H., & Pizzini, M. J. (2004). The balanced scorecard: Judgmental effects of performance measures linked to strategy. *The Accounting Review*, *79*(1), 1–23.

Baron, J. (1981). Reflective thinking as a goal of education. *Intelligence*, *5*(4), 291–309.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, *25*(3), 187–217.

Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2001). *Cost-benefit analysis*. *Concepts and practice*. Upper Saddle River, NJ: Prentice Hall.

Bonner, S. E. (1994). A model of the effects of audit task complexity. *Accounting, Organizations and Society*, *19*(3), 213–234.

Bonner, S. E., & Lewis, B. L. (1990). Determinants of auditor expertise. *Journal of Accounting Research*, *28*(Suppl.), 1–20.

Brierley, J. A., Cowton, C. J., & Drury, C. (2006). A note on the importance of product costs in decision-making. *Advances in Management Accounting*, *15*, 249–265.

Briers, M., Luckett, P., & Chow, C. (1997). Data fixation and the use of traditional versus activity-based costing systems. *Abacus*, *33*(1), 1–20.

Cardinaels, E., Roodhooft, F., & Warlop, L. (2004). Customer profitability analysis reports for resource allocation: The role of complex marketing environments. *Abacus*, *40*(2), 238–258.

Carnevale, P. J. D. (1985). Accountability of group representatives and intergroup relations. In: E. J. Lawler (Ed.), *Advances in group processes: Theory and research* (Vol. 2, pp. 227–248). Greenwich, CT: JAI Press Inc.

Carr, L. P., & Ittner, C. D. (1992). Measuring the cost of ownership. *Journal of Cost Management*, *6*(3), 42–51.

Chambers, R. J. (1996). Ends, ways, means and conceptual frameworks. *Abacus*, *32*(2), 119–132.

Chang, C. J., Ho, J. L. Y., & Liao, W. M. (1997). The effects of justification, task complexity and experience/training on problem-solving performance. *Behavioral Research in Accounting*, *9*(Supplement Conference), 98–116.

Chapman, C. S. (1997). Reflections on a contingent view of accounting. *Accounting, Organizations and Society*, *22*(2), 189–205.

Curley, S. P., Yates, F. J., & Abrams, R. A. (1986). Psychological sources of ambiguity avoidance. *Organizational Behavior and Human Decision Processes*, *38*(2), 230–256.

Degraeve, Z., Labro, E., & Roodhooft, F. (2005). Constructing a total cost of ownership supplier selection methodology based on activity-based costing and mathematical programming. *Accounting and Business Research*, *35*(1), 3–27.

Dewey, J. (1933). *How we think. A restatement of the relation of reflective thinking to the educative process*. Boston, MA: D. C. Heath and Company.

Doney, P. M., & Armstrong, G. M. (1996). Effects of accountability on symbolic information search and information analysis by organizational buyers. *Academy of Marketing Science*, *24*(1), 57–65.

Dunk, A. S. (2004). Product life cycle cost analysis: The impact of customer profiling, competitive advantage, and quality of IS information. *Management Accounting Research*, *15*(4), 401–414.

Ellram, L. M. (1995). Total cost of ownership. An analysis approach for purchasing. *International Journal of Physical Distribution and Logistics Management*, *25*(8), 4–23.

Fischer, G. W. (1995). Range sensitivity of attribute weights in multiattribute value models. *Organizational Behavior and Human Decision Processes*, *62*(3), 252–266.

Fox, F. V., & Staw, B. M. (1979). The trapped administrator: Effects of job insecurity and policy resistance upon commitment to a course of action. *Administrative Science Quarterly*, *24*(3), 449–471.

Fullerton, R. R., & McWatters, C. S. (2004). An empirical examination of cost accounting practices used in advanced manufacturing environments. *Advances in Management Accounting*, *12*, 85–113.

Galbraith, J. R. (1973). *Designing complex organizations*. New York: Addison-Wesley Publishing Company.

Gupta, M., & King, R. R. (1997). An experimental investigation of the effect of cost information and feedback on product cost decisions. *Contemporary Accounting Research*, *14*(1), 99–127.

Haka, S. (2006). A review of the literature on capital budgeting and investment appraisal: Past, present and future musings. In: C. S. Chapman, A. G. Hopwood & M. D. Shields (Eds), *Handbook of management accounting research* (Vol. 2). Oxford: Elsevier.

Harrison, D. S., & Killough, L. N. (2006). Decision outcomes under activity-based costing: Presentation and decision commitment interactions. *Advances in Management Accounting*, *15*, 169–193.

Haynes, C. M., Jenkins, J. G., & Nutt, S. R. (1998). The relationship between client advocacy and audit experience: An exploratory analysis. *Auditing: A Journal of Practice and Theory*, *17*(2), 88–104.

Israelsen, P. (1994). ABC and variability accounting differences and potential benefits of integration. *European Accounting Review*, *3*(1), 15–48.

Isselin, E. R. (1988). The effects of information load and information diversity on decision quality in a structured decision task. *Accounting, Organizations and Society*, *13*(2), 147–164.

Ittner, C. D., Larcker, D. F., & Meyer, M. W. (2003). Subjectivity and the weighting of performance measures: Evidence from balanced scorecard. *The Accounting Review*, *78*(3), 725–758.

Johnson, V. E., & Kaplan, S. E. (1991). Experimental evidence on the effects of accountability on auditor judgments. *Auditing: A Journal of Practice and Theory*, *10*(Suppl.), 96–107.

Kadous, K., Koonce, L., & Towry, K. L. (2005). Quantification and persuasion in managerial judgement. *Contemporary Accounting Research*, *22*(3), 643–686.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations*. Chichester: Wiley.

Kennedy, J. (1993). Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research*, *31*(2), 231–245.

Kennedy, J. (1995). Debiasing the curse of knowledge in audit judgment. *The Accounting Review*, *70*(2), 249–273.

Koonce, L., Anderson, U., & Marchant, G. (1995). Justification of decisions in auditing. *Journal of Accounting Research*, *33*(2), 369–384.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 107–118.

Langer, E. J. (1978). Rethinking the role of thought in social interaction. In: J. H. Harvey (Ed.), *New directions in attribution research* (Vol. 2, pp. 35–58). Erlbaum: Hillsdale, NJ.

Lerner, J.S., & Tetlock, P.E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*(2), 255–275.

Libby, T., Salterio, S. E., & Webb, A. (2004). The balanced scorecard: The effects of assurance and process accountability on managerial judgment. *The Accounting Review*, *79*(4), 1075–1094.

Lillis, A. M. (2002). Managing multiple dimensions of manufacturing performance: An exploratory study. *Accounting, Organizations and Society*, *27*(6), 497–529.

Lipe, M. G., & Salterio, S. (2000). The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review*, *75*(3), 283–298.

Lipe, M. G., & Salterio, S. (2002). A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations, and Society*, *27*(6), 513–540.

Maiga, A. S., & Jacobs, F. A. (2007). Activity-based cost management and manufacturing, operational and financial performance: A structural equation modeling approach. *Advances in Management Accounting*, *16*, 217–260.

Marchant, G., Robinson, J., Anderson, U., & Schadewold, M. (1991). Analogical transfer and expertise in legal reasoning. *Organizational Behavior and Human Decision Processes*, *48*(2), 272–290.

Nelson, M. W., Libby, R., & Bonner, S. E. (1995). Knowledge structure and the estimation of conditional probabilities in audit planning. *The Accounting Review*, *70*(1), 27–47.

O'Keefe, D. J. (2003). Message properties, mediating states, and manipulation checks: Claims, evidence, and data analysis in experimental persuasive message effects research. *Communication Theory*, *13*(3), 251–274.

Pampel, F. C. (2000). *Logistic regression – A primer*. Thousand Oaks, CA: Sage.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.

Peecher, M. E. (1996). The influence of auditors' justification processes on their decisions: A cognitive model and experimental evidence. *Journal of Accounting Research*, *34*(1), 125–140.

Pryor, J. B., Gibbons, F. X., Wicklund, R. A., Fazio, R. H., & Hood, R. (1977). Self-focused attention and self-report validity. *Journal of Personality*, *45*(4), 513–527.

Reck, J. L. (2001). The usefulness of financial and nonfinancial performance information in resource allocation decisions. *Journal of accounting and public policy*, *20*(1), 45–71.

Sawyer, A. G., Lynch, J. G. Jr., & Brinberg, D. L. (1995). A Bayesian analysis of the information value of manipulation and confounding checks in theory tests. *Journal of Consumer Research*, *21*(4), 581–595.

Schiff, A. D., & Hoffman, L. R. (1996). An exploration of the use of financial and nonfinancial measures of performance by executives in a service organization. *Behavioral Research in Accounting*, *8*, 134–153.

Schwartz, J., Chapman, G., Brewer, N., & Bergus, G. (2004). The effects of accountability on bias in physician decision making: Going from bad to worse. *Psychonomic Bulletin and Review*, *11*(1), 173–178.

Shelton, S. W. (1999). The effect of experience on the use of irrelevant evidence in auditor judgment. *The Accounting Review*, *74*(2), 217–224.

Sherlock, P. (1991). *Rethinking business to business marketing*. New York: The Free Press.

Shields, M. D. (1980). Some effects of information load on search patterns used to analyze performance reports. *Accounting, Organizations and Society*, *5*(4), 429–442.

Shields, M. D. (1983). Effects of information supply and demand on judgment accuracy: Evidence from corporate managers. *The Accounting Review*, *58*(2), 284–303.

Simon, H. A. (1979). Rational decision-making in business organizations. *Economic Sciences*, *69*(4), 493–513.

Skærbæk, P., & Thisted, J. A. (2004). Unit costs in central government annual reports: A critical appraisal of the practices developed. *European Accounting Review*, *13*(1), 7–38.

Snead, K. C. Jr., Johnson, W. A., & Ndede-Amadi, A. A. (2005). Expectancy theory as the basis for activity-based costing systems implementation by managers. *Advances in Management Accounting*, *14*, 253–275.

Snyder, M., & Kendzierski, D. (1982). Acting on one's attitudes: Procedures for linking attitude and behavior. *Journal of Experimental Social Psychology*, *18*(2), 165–183.

Stocks, M. H., & Harrell, A. (1995). The impact of an increase in accounting information level on the judgment of quality of individuals and groups. *Accounting, Organizations and Society*, *20*(7/8), 685–700.

Tan, H.-T., & Kao, A. (1999). Accountability effects on auditors' performance: The influence of knowledge, problem-solving ability, and task complexity. *Journal of Accounting Research*, *37*(1), 209–223.

Tan, H.-T., Ng, T. B.-P., & Mak, B. W.-Y. (2002). The effects of task complexity on auditors' performance: The impact of accountability and knowledge. *Auditing: A Journal of Practice and Theory*, *21*(2), 81–95.

Tetlock, P. E. (1983). Accountability and complexity of thought. *Journal of Personality and Social Psychology*, *45*(1), 74–83.

Tetlock, P. E. (1985). Accountability: The neglected social context of judgment and choice. *Research in Organizational Behavior*, *7*, 297–332.

Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In: M. P. Zanna (Ed.), *Advances in experimental social psychology*, (Vol. 25, pp. 331–376).

Vera-Muñoz, S. C. (1998). The effects of accounting knowledge and context on the omission of opportunity costs in resource allocation decisions. *The Accounting Review*, *73*(1), 47–72.

Vera-Muñoz, S. C., Kinney, W. R., & Bonner, S. E. (2001). The effects of domain experience and task presentation format on accountants' information relevance assurance. *The Accounting Review*, *76*(3), 405–429.

Verbeeten, F. H. M. (2006). Do organizations adopt sophisticated capital begetting practices to deal with uncertainty in the investment decision? A research note. *Management Accounting Research*, *17*(1), 106–120.

Wilks, T. J. (2002). Predecisional distortion of evidence as a consequence of real-time audit review. *The Accounting Review*, *77*(1), 51–71.

Wouters, M., Anderson, J. C., & Wynstra, F. (2005). The adoption of total cost of ownership for sourcing decisions – A structural equations analysis. *Accounting, Organizations and Society*, *30*(2), 167–191.

GROUP DECISION-MAKING AND LEADERSHIP: AN EXPERIMENTAL EXAMINATION IN AN EXECUTIVE COMPENSATION SCENARIO

Arron Scott Fleming

ABSTRACT

There has been concern expressed in the financial press and focus established in the accounting literature over rising levels of executive compensation. Individuals on the compensation committee, a sub-committee of the board of directors, collectively determine executive compensation and are responsible for maintaining the pay-for-performance standard, a concept that warrants further attention. This study examines the process of exaggeration of a group decision over individual beliefs and the impact of leadership upon a committee's outcome when making compensation awards. In an experiment with 98 subjects role-playing as compensation committee members, results show that in a committee of individuals where a coterie and a majority belief is present, group polarization occurs and the compensation results are exaggerated as compared to individual beliefs. The findings also suggest, though, that the appointment of a leader as chair of the committee, either in the majority or minority view, has a moderating effect on the group outcome. These results highlight and add to the literature the potential for

Advances in Management Accounting, Volume 17, 113–149
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017006

agency costs in the group decision process that may be found in the executive compensation-setting environment.

1. INTRODUCTION

There has been concern expressed in the financial press and focus established in the accounting literature during the past few years over rising levels of executive compensation.[1] An important compensation concept that repeatedly draws attention is pay-for-performance, i.e. increased compensation when firm performance increases and reduced compensation when firm performance declines. However, executive pay and performance do not always move in tandem. A CEO may maintain a high level of compensation or compensation increases even when his or her firm is producing losses or performance declines.[2] The pay-for-performance concept may be moderated or mediated by factors in the environment or in the decision process that affect the decision-makers. While past research has focused on factors such as board composition, board leadership structure, and firm size to name a few; inconsistencies in associative findings have led researchers to suggest exploring alternative theories (Daily, Johnson, Ellstrand, & Dalton, 1998) and moderating variables (Tosi, Werner, Katz, & Gomez-Mejia, 2000) as determinants of CEO compensation to gain further insights. This research experimentally examines potential factors involved in the decision process of the group responsible for determining executive compensation, the compensation committee. This research examines the process of exaggeration of a group decision over individual beliefs, and the impact of leadership type, in-group versus out-group, upon a committee's outcome.

One factor that is theorized to influence the compensation decision-making process is the relationship between the board and the CEO. The board of directors provides oversight and guidance to the executive management of a publicly traded corporation and has a fiduciary duty to the shareholders and investors they represent. However, the majority membership on the board is often more similar in status and peerage to the CEO than to the shareholders and investors. While the board often has outside members such as academicians and retired government officials, boards often also have other CEOs as the majority of outside members. It is the social relationship of status and power between the board and the CEO that may influence the compensation decision process

(Belliveau, O'Reilly, & Wade, 1996). Perel (2003) also suggests that complex interactions between the board of directors and CEOs can compromise the rational decision-making process surrounding CEO compensation. He notes that many corporate boards are closely allied with the CEO and "have become a part [sic] of the very culture they are supposed to challenge" (Perel, 2003, p. 383).

Though while the board of directors is ultimately responsible for ratification of CEO pay, a sub-committee of the board, the compensation committee, is the group that determines the compensation of the executive management and is responsible for maintaining the pay-for-performance standard. According to Nell Minow, editor of The Corporate Library, a corporate governance research firm, the best predictor of CEO overpay is the number of chief executives on the compensation committee (Burns, 2003). Ms. Minow suggests that a primary reason for excess pay is the composition of the compensation committee. Academic research is mixed is this area. Research has shown the presence of CEOs on the compensation committee allows the focal CEO to exert influence (Lorsch & MacIver, 1989), leads to compensation packages more in line with those preferred by CEOs than shareholders (O'Reilly, Main, & Crystal, 1988), but also that the proportion of CEOs on the compensation committee may be associated with lower pay (Daily et al., 1998).[3] Therefore, the relationship between the constitution of a compensation committee and the remuneration of the CEO is still an open, and important, question. This research attempts to address this question experimentally by performing a limited simulation of a compensation committee with individuals role-playing as outside CEOs or outside non-CEOs.

As individual decision-makers, members of the compensation committee formulate beliefs about the appropriate compensation award for a CEO. When these individuals gather and meet as a group, though, their individual beliefs may not be averaged to form a group decision (Sniezek & Henry, 1989), but may instead be exacerbated into an extreme position, depending upon the majority position of the group. For example, a compensation committee with three highly compensated CEOs and one significantly less compensated non-CEO member may have differing opinions as to a particular compensation award. If each individual CEO member supports a relatively high wage award and the non-CEO supports a more modest award, the group decision is not likely to be an averaging of these beliefs, but rather in this case, something larger. This could be due to multiple group decision processes. For example, if there is a dominant argument (three high pays versus one modest), the dominant argument may win and become the group belief (Pruitt, 1971). Or, for example, when an individual argues a position he or she may be inclined to exaggerate that belief. If the individual wins over the group, this exacerbated belief then becomes supported (Moscovici & Zavalloni, 1969; Myers & Arenson, 1972).

The general group process consists of the committee gathering as a group, reviewing various aspects of the performance of the CEO, possibly enlisting compensation consultants to gauge industry standards, and then determining the compensation of the executive management. Each committee member has their own personal pay-for-performance beliefs, compensation anchors, and social schemas. Membership of this committee is comprised entirely of outside directors[4] and may or may not have a publicly disclosed chair. The majority membership and leadership of this committee are often CEOs from other companies.[5] So, for example, the committee reviewing the compensation of the focal CEO may be comprised of chief executives from other corporations with possibly a member who is not a CEO. The leader of the committee to determine CEO pay may or may not be another CEO.

In scenarios role-played in this experiment, results indicate that a group decision outcome differs from the average of the individual decision outcomes. That is, at certain performance levels, the group awards greater compensation than the average of the individual members, and at lower levels, the group awards less compensation than the average of the individual members. Additionally, this research examines a potential influencing factor in the group decision-making process, leadership, which may have an impact on the compensation-setting process. This research finds support for moderating effects between firm performance and compensation committee leadership type (in-group versus out-group), but non-significant results for leadership type alone.

This research adds to the literature in two primary ways. First, it examines additional factors in the decision-making process that may moderate the impact of accounting performance measures and adds to the greater understanding of social and leadership dynamics of groups involved in the assimilation of accounting performance data. It does this by examining group outcomes under differing leadership type with identical performance data. Second, by utilizing an experimental approach in an effort to re-create a condition that may not occur frequently in the business world, the research helps isolate a cause and effect factor previously untested in past associative studies in accounting. This research experimentally examines the leadership factor in a scenario that is not commonly found in practice: compensation committee leadership by a non-CEO director. By identifying a moderating effect of leadership type and performance on compensation in an experimental setting, it highlights a potential agency cost in the authentic environment. Infrequently are compensation committee chairs non-CEOs, thus making archival observations difficult to obtain or non-existent. Empirically, it would be difficult to test the role of leadership in an archival manner if very few non-CEO or out-group directors exist. By experimentally creating and examining this phenomenon, this research sheds light on a decision process that relates to a corporate governance process, which further explains and predicts executive compensation awards.

The research proceeds as follows. First, prior research relating to executive compensation is reviewed, followed by group decision-making and social comparison phenomena. This is followed by the research methodology, statistical analysis, and summary.

2. RELATED LITERATURE AND HYPOTHESIS DEVELOPMENT

2.1. Executive Compensation

Investors and shareholders expect executive compensation and firm performance to be correlated, but associative studies on CEO compensation and performance have yielded mixed results. Research has indicated both positive relationships (e.g., Akhigbe, Madura, & Ryan, 1997; Duru & Iyengar, 1999; Gomez-Mejia, Tosi, & Hinkin, 1987; Lambert & Larcker, 1987; Natarajan, 1996; Newman & Mozes, 1999; Young & Buchholtz, 2002) and weak, poor, or negative relationships (e.g., O'Reilly et al., 1988; Fosberg, 1999; Iyengar, 2003). When associations are identified, often other interesting factors or confounds are likewise linked. For example, Newman and Mozes (1999) find that the relationship between performance and pay is positive, yet more favorable to the CEO when the firm has insiders on the compensation committee.[6] O'Reilly et al. (1988) note that return on equity is related to executive cash compensation, but also related to executive cash compensation is the relationship of compensation of the outside members on the compensation committee to CEO pay. These latter two studies suggest compensation committee membership may play a part in compensation awards.

The relationship between the board and the CEO can also compromise the compensation decision-making process (Perel, 2003). Confounding factors and multiple determinants of CEO compensation may exist, suggesting that researchers explore alternative theories (Daily et al., 1998) and moderator variables (Tosi et al., 2000). Additional associative studies on the composition and outcomes of the compensation committee or board of directors include Dechow, Huson, and Sloan (1994), Zajac and Westphal (1996), Gaver and Gaver (1998), Young and Buchholtz (2002), and Adut, Cready, and Lopez (2003), to name a few. However, to date there have been no experimental studies published on the same topics known to the author.

2.2. Group Formation and Decision Making

The psychology literature suggests people in general have a need to belong to groups, and there is constant pressure towards uniformity within groups (e.g., Greenberg, Solomon, & Pyszczynski, 1997). Research on groups has given attention to the natural formation of groups, uniformity within groups, and a normalization of behavior (e.g., Festinger, 1950, 1954). Greenberg et al. (1997) suggest that individuals belong to at least some group and that people identify themselves with a particular worldview. Baumeister and Leary (1995) indicate that group interactions take place in a stable environment and endure due to a concern for other group members' welfare. Additionally, Baumeister and Leary (1995) note that groups exist beyond the need for basic survival and may be shaped by economic need and opportunity. Festinger (1950, 1954) describes a social comparison tension within groups as a force that persuades members to strive for the uniformity necessary for group goal achievement. Individuals within the group share information, socially compare, and corroborate beliefs. Groups perform decision making on the basis of consensus after thorough discussion (Whyte, 1993).

In this research groups are formed in a role-playing experiment. The subjects are either in-group members, role-playing as CEO directors, or they are out-group members, role-playing as non-CEO directors. This role, along with a description of their background and personal compensation, forms the basis of their individual worldviews. This creates a sub-group of individuals within a committee with similar, but not exact, views. This slight difference in views is a catalyst that leads to differing beliefs and individual decisions. The group, though, strives for uniformity and seeks a consensus in the decision-making process.

In this research teams are formed for a specific purpose, to evaluate and determine the pay for a fictitious CEO given a certain level of performance. The individual team members have different personal compensation levels, which facilitate the formation of individual beliefs. Through face-to-face discussions, the members share individual information, espouse individual beliefs, consolidate the beliefs through discussion, and, as a group, reach a decision. The appointed committee chair records and communicates the decision.

One way in which individual decision-making differs from group decision-making is that the latter is more prone to social influences. When a group gathers to make a decision, the process may be influenced by members' ideas, communications, and perceived group norms. There is pressure to conform. In making a decision, group members may rely upon the views of other members, thus superficially processing the information; if the decision or outcome is considered highly important, the members may systematically process the information while considering the views of the other members (Smith & Mackie, 2000). When a group makes a decision, the outcome may be an average of individual decisions or something more extreme. If group members are fairly evenly divided on an issue yet reach a decision, the group may compromise and converge to a moderate or average position. If a majority of group members favor a particular view, though, the discussion may become polarized and move the group to a more extreme position (Smith & Mackie, 2000). Originally described as Risky Shift (Myers & Bishop, 1970), this phenomenon explains the group behavior in terms of shared risk. An individual making a decision may be more conservative than the same person making a decision within a group, since the risk is shared. Thus, individuals make less risky decisions as compared to group decisions. It was noted, though, that the direction of the shift was dependent upon the initial dominant viewpoint of the group. This phenomenon later became the Group-Polarization Hypothesis (Myers & Lamm, 1976), which suggests that groups make extreme decisions in the direction of the majority members' beliefs.

Group judgments are not merely an averaging process (Sniezek & Henry, 1989) but may result in exacerbated or extreme positions, particularly if initial individual positions differ. In many groups, members have differing opinions or views. In this experiment, members are provided with personal information that is different from the person next to them to facilitate this event. In a four-member group, three will have personal compensation levels that are relatively high (outside CEOs) and one will have a personal compensation level that is relatively low (outside non-CEO). The belief is that the group will assimilate this information and, through a group decision-making process, will determine an outcome that may not be an averaging of individual beliefs.

There are compelling explanations for group polarization, this non-averaging outcome. Persuasive Argument Theory (Pruitt, 1971), an instance of group polarization, suggests that individual members of a group develop arguments in support of their personal positions. Information is systematically processed, and, in order to convince other group members, extreme arguments are espoused. Members exaggerate a position to convince other members. As a result, one argument will dominate, and the group will buttress an extreme decision (Pruitt, 1971). Once the position is won, the member feels obligated to support the position, even though it may feel excessive (Moscovici & Zavalloni, 1969; Myers & Arenson, 1972). In this research, subjects are induced into personal positions through role assignment and personal pay anchors. In-group members are assigned the role of CEO director and have higher personal compensation, whereas out-group members are assigned the role of academic director and have lower personal compensation. In this experiment, the number of in-group to out-group subjects is approximately 3-to-1.

Another explanation for group polarization is Escalation of Commitment (Staw, 1976), which suggests that as members strive for and begin to achieve group uniformity, they become more committed to the actions and outcomes of the group. Members process the information superficially and employ heuristics: they incorporate and rely upon the views of other group members in their decision-making. This is a social process of alignment of individual opinions and decision movement within a group (Friedkin, 1999). Whyte (1993) cites Staw (1980) and Sandelands, Brockner, and Glynn (1988) in explaining the primary theoretical underpinnings of escalation commitment: individuals are involved in self-justification to prove to themselves and others they are competent and rational, often at the expense of error persistence. Even if individuals within the committee disagree at some level with the group decision, they will support the decision. The individual will not want to be at odds with the group. If the group is perceived as deriving the correct answer, any individual not agreeing may be perceived as incompetent.

Escalation is also explained within the group setting as a decline in personal responsibility over individual actions and decisions (Darley & Latane, 1968; Zimbardo, 1970; Mynatt & Sherman, 1975). This group mechanism can develop not only confidence in decision-making but also opinion extremes (Baron et al., 1996). In this experiment, each individual subject is asked to determine and record the compensation award to the CEO. Since there is no particular correct answer, subjects should approach a decision with some conservatism. The same individuals are then introduced into groups where they then again must determine and record the compensation award to the CEO. This time the decision is made by the group. If the decision is somehow "incorrect", the group is responsible, not the individual. Therefore, less risk may be viewed in a group decision by individuals in supporting a solution that is more risky or aggressive.

A group may even be more concerned with reaching a consensus rather than making the correct decision, thus potentially leading to *Groupthink* (Smith & Mackie, 2000). Groupthink may be an issue in cohesive groups (Janis, 1972). In a team that is striving for harmony and thus shying away from conflict, members may either remain silent or endorse an outcome to remain in accord with the group. As a result of groupthink, individuals may agree with a group that is making a decision to provide high pay for a CEO who has low performance. Since in this experiment there are two to three individuals with higher personal compensation and one with lower compensation, the less compensated individual may succumb to groupthink. That is, in order not to conflict with the group, the lower-paid out-group member may stay silent in an attempt to avoid conflict with the other group members. It is this phenomenon that can induce groups to make bad decisions in certain situations. A possible mitigative action to this phenomenon is through the introduction of a minority or dissenting viewpoint. By having a minority viewpoint expressed, conformity pressure may be relieved. This may be achieved by appointing the member with the minority viewpoint, presumably the non-CEO director, as the leader of the group.

In summation, people may form associations or coteries that align and marshal opinions or goals. This can lead to group confidence with potentially exaggerated, exacerbated, or extreme decision outcomes. This experiment is set up in such a way as to induce such an effect. The general theoretical outcome is illustrated in Fig. 1.

2.3. Corporate Governance Committees as Groups

Groups and group processes exist in business and corporate governance contexts. Both the board of directors and the compensation committee may represent a group not only in name, but also in the pursuit of common goals, purpose, and demographic make-up. Bainbridge (2002) notes that corporations are hierarchical in nature, whereas the board of directors is a collegial group working towards consensus. Zajac and Westphal (1996) find evidence suggesting that board members tend to pick someone demographically similar to their own profiles in choosing an outside successor. Young and Buchholtz (2002) find that the compensation committee is

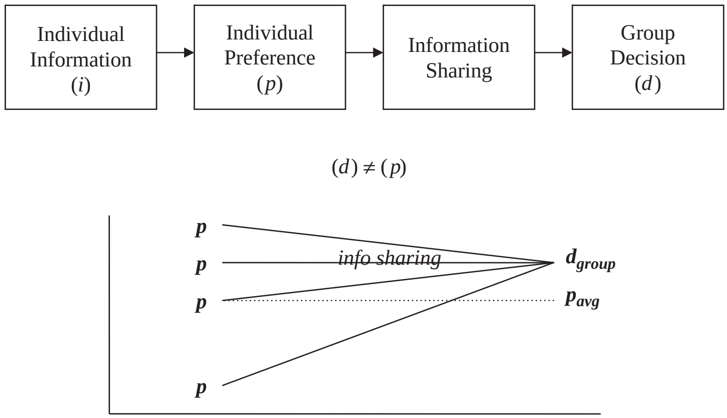


Fig. 1. Group Decision Escalation Model.

influenced by demographic similarities to the CEO. Belliveau et al. (1996) find that social capital (social status and network ties) of the CEO is associated with higher compensation. O'Reilly et al. (1988) find evidence supporting social comparison theory, noting an association between the compensation level of the outside directors on the compensation committee and that of the CEO. The theory suggests that individuals make comparisons to those they perceive as similar (Festinger, 1954). In short, theory and evidence suggests the development of associations or coteries among those with similar characteristics, including status, position, and wealth.

The compensation committee may be one such coterie, particularly if many of the individuals within the group are similar either in position or status. Associative findings indicate that the compensation committee, often filled with outside CEOs, protects or shields the chief executive's compensation against negative performance (Dechow et al., 1994; Gaver & Gaver, 1998; Duru, Iyengar, & Thevaranjan, 2002; Adut et al., 2003). That is, when performance is down, the compensation committee does not modify compensation in a similar fashion. When performance is up, though, CEOs are often rewarded. This is a different approach from the pay-for-performance paradigm. This indicates that a coterie, an exclusive group of people who associate together, may moderate the pay-for-performance ideal.

The conjecture is that a group decision of like individuals with different informational beliefs will overpay or underpay when compared to individual decisions. This relates to the compensation-setting process in that a compensation committee often has members similar in status, but with differing reference points and beliefs about compensation. As individual members review information, they will form individual opinions and preferences. When members congregate, though, they will share information and create an exacerbated group preference. That is, the group will choose an outcome that is more extreme than the simple average of the group. This group polarization, whether explained by the Persuasive Arguments Theory or Escalation of Commitment, yields the same results. The first hypothesis, based on the social/psychological underpinnings of Fig. 1, is stated in the alternative form as such

H1. Groups compensation levels will not equal individual compensation levels.

2.4. Leadership in Groups

On the basis of the nature and membership of the group, the inference exists that a group led by someone outside the coterie may serve to mitigate the group exaggeration process. This research is not concerned with the identification, style, or personality of leadership of individual committee chairs, but rather is interested in the group-decision impact of team leadership where the leader is or is not a member of the majority group coterie. If a leader is appointed, he or she may have influence over the team by the simple virtue of having the express authority to voice their personal belief.

This will be particularly true if the leader is a member of the in-group and shares similar status with the team. If, however, the leader is a member of the out-group, the leader is less likely to completely indulge the team's suggestion and the leader will exert some influence. It is in this situation that an out-group leader will serve to moderate the escalation of the in-group members. A reason behind this is that groups tend to *choose* leaders with similar worldviews and norms (Foti, Fraser, & Lord, 1982; Hogg, 1996). Group leaders with similar dispositions or positions to the group are rated as more effective leaders (Hains, Hogg, & Duck, 1997), particularly if they identify with the group. Thus, an out-group individual *appointed* as a leader will be less effective, resulting in a group decision that is less likely to follow the "normal" group polarization process in decision-making. In a compensation-decision setting relating to corporate governance, this would be synonymous with a compensation committee with an appointed non-CEO board member as leader or chair. For example, given a compensation committee of four members where three are outside CEOs and one is an academic, when the academic leads the committee the compensation award level should be closer to the individual average than when the committee is led by one of the outside CEO members. The academic or non-CEO may serve to inhibit the process of choice shift, since the non-CEO may not be within the same status or social group as the CEOs. This leadership factor may serve to explain some variance and inequities in the pay-for-performance paradigm. Fig. 2 serves as an illustration of this effect when the choice shift is upwards.

My second hypothesis, stated in the alternative form, is given below

H2. Groups led by out-group subjects will award compensation levels that differ from groups led by in-group subjects.

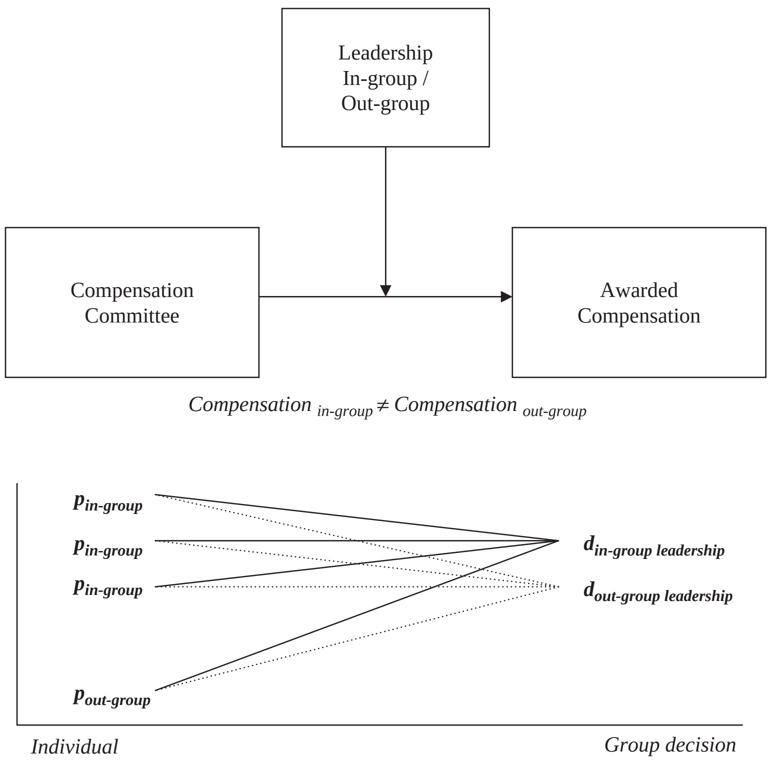


Fig. 2. Leadership Impact within Group Model.

3. RESEARCH METHODOLOGY

3.1. Subjects

Subjects in the study are from two cohorts of the executive masters' of business administration program and from the masters' of professional accountancy program (hereafter noted as eMBA and MPA, respectively) of a large eastern U.S. university. Participation was voluntary and took place during class time. Subjects were not compensated, but rather they were informed by their administrating faculty that participation in-and-of-itself would enhance their educational experience. The subject pool contained 59 eMBA and 39 MPA students, for a total of 98 individuals. The eMBA group contained 35 male and 24 female subjects, whereas the MPA group contained 23 male and 16 female subjects. The subjects averaged 10.7 and 3.24 years in experience, respectively, for the eMBA and MPA groups. The average age was 33.4 years for the eMBA group and 26.2 years for the MPA group.

The experiment had a repeated measure aspect making the total gross observations 196. Out of the 196 observations, 23 eMBA and 15 MPA observations were eliminated, due to either missing data or outlier analysis,[7] for a total of 158 individual and 50 group observations. The removal of these observations did not appreciably change either the gender make-up, age, or experience of the subject pool.

3.2. Research Overview

This research is a role-playing experiment in which subjects act as members of the compensation committee of a fictitious company. For the experiment, subjects individually read information about their role, the company and its performance, and the industry performance benchmark as background (see Appendix C for a sample vignette). For every subject role-playing as a non-CEO member of the compensation committee, three others are role-playing as CEO members of the compensation committee. From the provided material, each subject determines and records the compensation of a fictitious CEO. Then, maintaining the same roles, subjects are assigned into groups comprised of three to four members. In each group there is one non-CEO member with the remainder role-playing as CEO members. Subjects wear name tags denoting the vocation of the group members, e.g., "Professor of Business" or "CEO of ABC Company". The name tags serve as a role reminder to individual subject and to those in the group. One member, either a CEO director or a non-CEO director, is appointed as committee chair. The appointment of the chair is done in the presence of the other group members, as are the instructions to the chair in regard to the recording of the compensation award. The group, given the same company performance level as previously reviewed as individuals, then determines the compensation award and the chair records the same. The entire experiment is then repeated with the same vocational roles, but with a different company performance level, with different committee members, and with a different director type as leader. The order of leadership and company performance is counter-balanced.

3.3. Variables

The dependent variable is compensation awarded for individuals and groups, or the absolute value of the difference between individuals and groups. To alleviate potential personal confounds, the compensation is in a fictitious monetary unit (Qwert). Confounds such as preconceived beliefs about the actual dollars a CEO may or may not make, salaries of contemporary athletes and pop culture icons, or even the salaries of close friends or colleagues could confound the experiment if it is conducted in dollars. Using a fictitious monetary unit as an experimental device guides the subject into using the provided references and anchors.

The primary independent variable is director type (non-CEO director or CEO director). This is operationalized through two mechanisms, the role-played and the salary of the role. The roles and titles assigned include "Professor of Business" for the non-CEO director, "CEO of ABC Company", "CEO of DEF Company", or "CEO of GHI Company" for the CEO directors. The manipulation is strengthened by using name tags with the subject's profession as a visual cue. The name tag serves to make group affiliation more salient to the subjects in individual decision-making and to the other subjects in group decision-making. The salary for each role is 30, 70, 90, and 110 Qwert, respectively. This is not the compensation they receive as a director, but the compensation they make in their normal vocational position.

The other independent variable for the individual decision process is firm performance (−2% or 6% growth rates). Performance is expressed in earnings per share and growth rates with clear indications that the firm is performing below industry standards (−2% and 6% versus 10% industry growth rates). The performance variable is introduced to limit the time-related maturation effect of the repeated measure aspect. If, during the second attempt of the individual and group decision process the firm performance is identical, then there exists the strong possibility of carryover effects. By changing firm performance, though, carryover effects should be lessened. For the group decision process, the additional independent variable is the leader type (CEO or non-CEO leader).

3.4. Procedures

Approximately 3–5 days before the experiment, subjects read, as a primer, a short overview on corporate governance and the importance of the compensation-setting process (Appendix A). This was sent via electronic mail along with a cover letter to each subject. Prior to the start of the experiment, subjects read instructions and completed a short demographic questionnaire (Appendix B).

For the individual component of the experiment, subjects, wearing their name tags, read a vignette (Appendix C), recorded their private compensation award decision, and completed a short questionnaire (Appendix D). Individuals were then assigned to groups, leaders were assigned, and a form for the committee answer was given to the appointed chair. Following the committee decision, the individual non-chair members and the chair completed a group-related questionnaire (Appendix E). Excluding the individual demographic portion, the entire procedure was repeated. Individuals kept the same role, but the performance of the company was changed, groups were changed, and leadership was changed. Roles were randomized, and firm performance and group leadership were counterbalanced. The entire experiment took approximately 30–45 min. Table 1 has the breakdown of individual observations by role and performance level.

4. RESULTS

The results show support for hypothesis H1, that groups will award compensation levels that differ from individual compensation awards. Results show mixed support for hypothesis H2, that groups led by out-group subjects will award compensation levels that differ from groups led by in-group subjects. This is affected by the performance variable. Given that the experimental procedures involved a repeated measure between- and

Table 1. Individual Observations by Role and Performance Levels.

Director Type	Personal Pay in Qwert[a]	Company Performance Growth % (Industry Average 10%)	Number of Observations
Non-CEO	30	−2	22
Non-CEO	30	6	14
		Total	36
CEO	90[b]	−2	54
CEO	90[b]	6	68
		Total	122
		Total Observations	158

[a]Qwert is a fictional currency.
[b]Each CEO director has a personal compensation level of either 70, 90, or 110 Qwert with each group having a CEO member average of 90 Qwert.

within-subjects design, and the time between the experiments may not allow for a washout effect, an examination of sequence effects using a *t*-test for Equality of Means is employed. A test for sequence effects shows no significant difference between attempts 1 and 2, either individually or in the group setting.[8]

To analyze hypothesis H1, that groups will award compensation levels that differ from individuals, the group scores are compared to the average of the individual scores. For example, the four individual scores for the members ultimately to be assigned into group 1 are averaged (QwertIavg) and compared to the group 1 decision (QwertG). The overall results indicate an individual average of 71.97 versus a group average of 73.46, for a two-tailed *p*-value of 0.284. Upon closer examination, though, findings indicate a difference when dividing the sample by the firm performance variable. Given a stated industry average growth rate of 10%, groups awarded *lower* compensation levels than individuals when it was indicated that firm performance was −2%, but significantly *higher* compensation levels when it was indicated that firm performance was 6% (see Fig. 3). At the −2% level, the individual average is 67.42 and the group average is 65.65, for a *p*-value of 0.360. At the 6% level, though, the individual average is 76.54 and the group average is 81.27, for a *p*-value of 0.015. This suggests that the group exaggeration process works to the negative side when the performance reaches some minimal point, and that the process works to the positive side above some level. Table 2 contains the results of this analysis and Fig. 3 graphically illustrates the results.

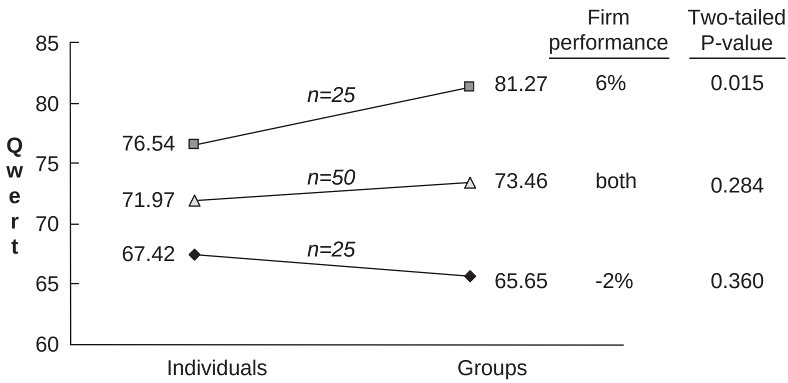


Fig. 3. Individual versus Group.

Table 2. Individual versus Group Means *t*-Test.

Performance	Pair	Mean	*N*	Standard Deviation	Standard Error	*t*	*p*-value
All	QwertIavg	71.97	50	10.258	1.45	−1.084	0.284
	QwertG	73.46	50	13.778	1.95		
−2%	QwertIavg	67.41	25	9.184	1.84	0.934	0.360
	QwertG	65.65	25	13.226	2.65		
6%	QwertIavg	76.54	25	9.335	1.87	−2.631	0.015
	QwertG	81.27	25	9.252	1.85		

Note: QwertIavg is the average of the individuals within the group. QwertG is the group award. The *p*-value provided is using a two-tailed test.

Additional analyses of the results indicate significant support for the group exaggeration process. Given that group results show an exaggeration above the individual average for the 6% performance level, and an exaggeration below the individual average for the −2% performance level, an additional analysis is performed using the absolute value of the exaggeration from the individual average to the group average. In comparing the absolute value of the difference of individual decisions from the individual average to the absolute value of the difference of group decisions from the group average, results indicate that the group decision difference is significantly greater than the individual decision difference ($t = 3.932$, $p < 0.001$). That is, group differences were further from the mean than were individual differences, suggesting that the groups suffered a decision exaggeration. With an individual average of 71.97, the absolute value average of difference to the mean is 8.42, with a standard deviation of 5.73. The group average is 73.46 with an absolute value average of difference to the mean of 11.95, with a standard deviation of 6.64. Thus, while the nature of a group is to align individuals and arrive at a group consensus, in this scenario the group outcome itself is exaggerated in comparison to the individual outcome alone.

To analyze hypothesis H2, that groups led by out-group subjects will award compensation levels that differ from groups led by in-group subjects, regression analysis is employed. The variables may be expressed in the following relationship for hypothesis H2:

$$\begin{aligned} QwertG_i = \alpha + \beta_1 Attempt_i + \beta_2 Perf_i + \beta_3 Leadtype_i + \beta_4 Attempt_i * Perf_i \\ + \beta_5 Attempt_i * Leadtype_i + \beta_6 Perf_i * Leadtype_i \\ + \beta_7 Attempt_i * Perf_i * Leadtype_i \end{aligned}$$

where, for each group *i*, *QwertG* is the compensation awarded; *Attempt* is either the first or second group attempt; *Perf* the firm performance, either −2% or 6%; and *Leadtype* the leader of the group, either a CEO or non-CEO director.

The results indicate that there are main effects for *Perf* only ($F = 34.692$, p-value < 0.001), whereas the variables *Attempt* and *Leadtype* are not significant (p-values of 0.095 and 0.717, respectively). Results also indicate that there is a two-way interaction effect for *Attempt* * *Perf* and *Perf* * *Leadtype* ($F = 23.316$, p-value < 0.001; $F = 6.259$, p-value 0.016; respectively) suggesting the existence of a moderator variable (Table 3).

In an examination of the interaction of performance and attempt, results indicate an exaggeration of group results, perhaps through group learning or momentum. For the −2% case, the group mean for the first attempt is 73.50 versus a mean of 57.15 for the second attempt, yielding a two-tailed p-value of 0.001. For the 6% case, the mean for the first attempt is 77.64 and 84.62 for the second attempt, yielding a two-tailed p-value of 0.058. Perhaps just as importantly, though, is the standard deviation of the group decisions from the first attempt to the second. For both cases, the standard deviations are cut by 4–5 Qwert, suggesting a tightening, along with the exaggeration process, of group decisions as learning, practice, and information symmetry spread. It is important to note that learning is not occurring within each performance level, since each member and each group see each performance level only once. Fig. 4 graphically illustrates the results contained within Table 4.

Table 3. Test of Effects in Group Compensation Awards (QwertG).

Source	Sum of Squares	df	Mean Square	F	p-value
Corrected Model (adjusted $R^2 = 0.539$)	5628.51	7	807.07	9.195	0.000
Intercept	245113.47	1	245113.47	2803.03	0.000
Attempt	255.62	1	255.62	2.923	0.095
Perf	3033.64	1	3033.64	34.692	0.000
Leadtype	11.687	1	11.687	0.134	0.717
Attempt ∗ Perf	2038.90	1	2038.90	23.316	0.000
Attempt ∗ Leadtype	84.86	1	84.86	0.970	0.330
Perf ∗ Leadtype	547.34	1	547.34	6.259	0.016
Attempt ∗ Perf ∗ Leadtype	0.15	1	0.15	0.002	0.967
Error	3672.72	42	87.45		
Total	279121.28	50			
Corrected Total	9301.24	49			

Note: Attempt is the group attempt in the repeated measure. *Perf* is the performance of the firm under evaluation. *Leadtype* is the leader of the group making the decision, either a CEO or non-CEO director.

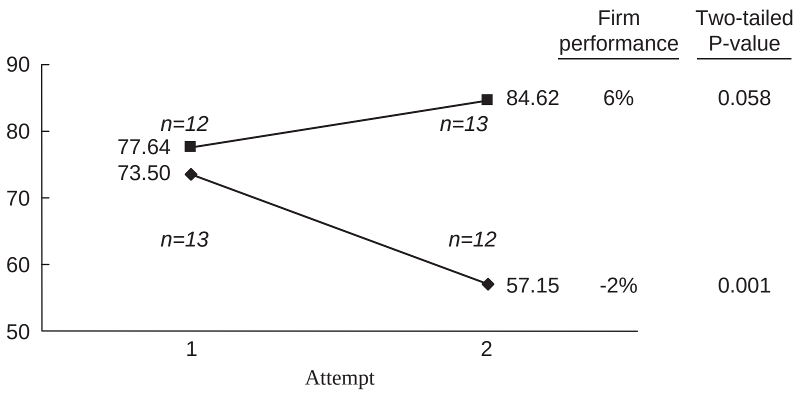


Fig. 4. Group Means by Attempt and Performance.

In an examination of the interaction of performance and lead type, results indicate an exaggeration of group results in the expected direction, but not significantly. For the −2% case, the group mean for the non-CEO led groups is 67.77 versus a mean of 62.48 for the CEO led groups, yielding a two-tailed p-value of 0.338. For the 6% case, the mean for the non-CEO led

Table 4. Group Means *t*-Test by Performance and Attempt.

Performance (%)	Attempt	Mean	*N*	Standard Deviation	Standard Error	*t*	*p*-value
−2	1	73.50	13	12.503	3.47	3.894	0.001
	2	57.15	12	7.712	2.23		
6	1	77.64	12	10.832	3.13	−1.996	0.058
	2	84.62	13	6.195	6.20		

Note: The Mean is in Qwert for the group. The *p*-value provided is using a two-tailed test.

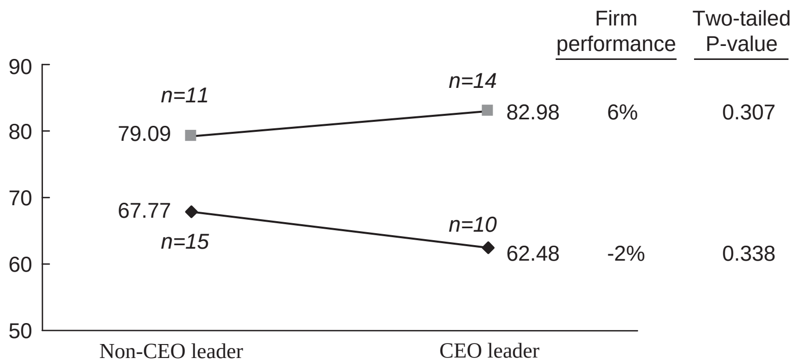


Fig. 5. Group Means by Lead Type and Performance.

groups is 79.09 and 82.98 for the CEO led groups, yielding a two-tailed *p*-value of 0.307. The results, although not significant, do provide some insight into the group process as to the direction the leader may have upon the group's decision. From the individual decisions, results show that for the −2% performance level, the CEO directors significantly reduce pay below that of non-CEO directors, and for the 6% performance level, the CEO directors significantly increase pay above the non-CEO directors. In the group results, we see the same direction of award by the same type of group leader. That is, in the −2% performance level, groups led by CEO directors award less than non-CEO directors, and in the 6% performance level, the reverse is true. This shows that the out-group leader, in this case the non-CEO director, may be a mitigator to the group exaggeration process. Fig. 5 illustrates the results contained in Table 5.

Table 5. Group Means *t*-Test by Performance and Director Type.

Performance (%)	Leader	Mean	*N*	Standard Deviation	Standard Error	*t*	*p*-value
−2	Non-CEO	67.77	15	12.420	3.21	0.978	0.338
	CEO	62.48	10	14.419	4.56		
6	Non-CEO	79.09	11	9.635	2.91	−1.045	0.307
	CEO	82.98	14	8.914	2.38		

Note: The Mean is in Qwert for the group. The *p*-value provided is using a two-tailed test.

5. CONCLUSIONS

This research studies decision outcomes by comparing the individual and group decision-making preferences in an experimental scenario. This relates to accounting and corporate governance in that the compensation committee, a group with multiple member types, weighs accounting performance in determining CEO pay. This pay-for-performance research attempts to re-create a group decision-making scenario not commonly found in practice, an appointed out-group leader. Specifically, through a role-playing experiment, the research examines the effects of group decision-making and the potential for group decision exaggeration over individual decision. Support for group exaggeration is found. Additionally, this research examines the influence of leadership type on the decision-making process. While leadership type alone is not significant, results indicate moderating effects through the interaction of leadership type and firm performance, suggesting at some levels leadership type may represent an agency cost.

An examination of this decision-making process is important since chief executive compensation continues to rise and is a social and political topic of discussion. This study adds greater understanding about the social and psychological factors underpinning the compensation decision-making process by experimentally studying the effects of group membership and committee leadership versus individual compensation awards.

Limitations to the study, though, do exist. In examining the validity of the experiment, both the internal and external validity must be weighed. Internal validity is strengthened through the application of random assignment, variable control, and variable precision in the examination of cause and effect. Extraneous variations are minimized while, it is believed, systematic variation is enhanced through the procedures employed.

External validity may be somewhat weaker, though. This research acknowledges that ecological validity, the realism of the setting and situation in which compensation committee decisions are made, is greatly reduced, but it can be argued that many of the factors that relate to an authentic setting may be mundane in nature. This *mundane realism* (Aronson & Carlsmith, 1968) relates to events or circumstances that may exist in a field setting, but in actuality may be unimportant to the population and phenomenon. *Experimental realism* (Aronson & Carlsmith, 1968) is the level in which the experimental events' mimics the real world decision process. While the exact process in determining executive compensation may differ between firms (and within firms from year to year), it can be assumed with a general assurance that a committee of individuals meet, review results, make comparisons, and then determine appropriate pay levels. Therefore, while generalizability is limited given the fact that the experiment does not use actual board of director members, the experimental realism is reasonably intact while minimizing mundane realism. In one aspect, generalizability may not be possible, since compensation committee leadership by a non-CEO does not appear to be common, and a sizeable population may not exist. This is not to say, though, that in the future through political and social influences, that this circumstance will remain.

Future research might focus on experimental factors that are beyond the scope of this study. This study is performed in the performance domain that is below average, and may be conducted in the positive performance domain, with multiple levels of performance. An additional factor that could tie into this study is the tenure of compensation members and whether they are appointed during the tenure of the CEO. This may affect the committee member's allegiance and either strengthen or reduce the member's group inclusion.

Finally, the experimental methodology used allows for a more in-depth study of cause and effect factors in a decision-making setting. Associative studies have yielded a wealth of information and have directed intellectual curiosities into certain areas. This research expands upon previous findings to isolate potential variances in this type of decision-making and helps to illuminate factors that have not previously been researched comprehensively.

NOTES

1. For example, Lublin (2005, February 25, *The Wall Street Journal*), Anonymous (2004, December 11, *The Economist*), Perel (2003), and Sheikholeslami (2001).

2. For example, Iyengar (2003), Adut et al. (2003), and Tosi et al. (2000).
3. Daily et al. (1998) note this relationship is the opposite of that hypothesized.
4. SEC rule 303A.
5. As per the SEC proxy filings (DEF 14A) for the year ended December 31, 2003, General Electric, Walmart, Citigroup, and ExxonMobil in total have 17 members on the compensation committee. Of these members, seven are active chairmen or CEOs, seven are retired chairmen or CEOs, and three are non-CEOs. Two of the companies list the leadership of the compensation committee, of which one is an active CEO/Chairman and one is retired. These companies, as per the *Forbes* 2004 rankings, represent the world's largest conglomerate, retailer, bank, and oil and gas concerns in terms of market capitalization ($328.54, $243.74, $255.30, and $277.02 billion, respectively).
6. The author's definition of "insider" is different from that of the SEC. The authors define "insider" as any member who is likely to be positively biased in determining CEO compensation. This includes, for example, former employees, individuals with ties to or representing businesses with significant business dealings with the firm, and interlocking directors.
7. Outliers were removed using the least restrictive measure of either (1) ± 2 standard deviations, or (2) Grubbs' Test for Outliers (Grubbs, 1969; Stefansky, 1972).
8. A *t*-test for Equality of Means was computed for the individual and group scores. Individual means for attempt 1 and 2 were 72.4375 and 73.7410, respectively, for a *p*-value of 0.670. Group means for attempt 1 and 2 were 75.4884 and 71.4320, respectively, for a *p*-value of 0.303.

REFERENCES

Adut, D., Cready, W. H., & Lopez, T. J. (2003). Restructuring charges and CEO cash compensation: A reexamination. *Accounting Review*, *78*(1), 169–192.

Akhigbe, A., Madura, J., & Ryan, H. (1997). CEO compensation and performance of commercial banks. *Managerial Finance*, *23*(11), 40–55.

Anonymous. (2004). Leaders: Running out of options; Pay for performance. *The Economist*, December 11, *373*, p. 10.

Aronson, E., & Carlsmith, J. M. (1968). Performance expectancy as a determinant of actual performance. *Journal of Abnormal Social Psychology*, *65*, 178–182.

Bainbridge, S. M. (2002). Why a board? Group decisionmaking in corporate governance. *Vanderbilt Law Review*, *55*(1), 1–55.

Baron, R. S., Hoppe, S. I., Kao, C. F., Brunsman, B., Linneweh, B., & Rogers, D. (1996). Social corroboration and opinion extremity. *Journal of Experimental Social Psychology*, *32*(6), 537–560.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529.

Belliveau, M. A., O'Reilly, C. A., & Wade, J. B. (1996). Social capital at the top: Effects of social similarity and status on CEO compensation. *Academy of Management Journal*, *39*(6), 1568–1593.

Burns, J. (2003). Everything you wanted to know about corporate governance... but didn't know to ask. *Wall Street Journal* (October 27), R6–R7.

Daily, C. M., Johnson, J. L., Ellstrand, A. E., & Dalton, D. R. (1998). Compensation committee composition as a determinant of CEO compensation. *Academy of Management Journal*, *41*(2), 209–220.

Darley, J. M., & Latane, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*, 377–383.

Dechow, P. M., Huson, M. R., & Sloan, R. G. (1994). The effect of restructuring charges on executives' cash compensation. *The Accounting Review*, *69*(1), 138–156.

Duru, A., & Iyengar, R. J. (1999). Linking CEO pay to firm performance: Empirical evidence from the electric utility industry. *Managerial Finance*, *25*(9), 21–33.

Duru, A., Iyengar, R. J., & Thevaranjan, A. (2002). The shielding of CEO compensation from the effects of strategic expenditures. *Contemporary Accounting Research*, *19*(2), 175–193.

Festinger, L. (1950). Laboratory experiments: The role of group belongingness. In: J. G. Miller (Ed.), *Symposium on social psychology*. New York: McGraw-Hill.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Fosberg, R. H. (1999). Leadership structure and CEO compensation. *American Business Review*, *17*(1), 50–56.

Foti, R. J., Fraser, S. L., & Lord, R. G. (1982). Effects of leadership labels and prototypes on perceptions of political leaders. *Journal of Applied Psychology*, *67*, 326–333.

Friedkin, N. E. (1999). Choice shift and group polarization. *American Sociological Review*, *64*(6), 856–875.

Gaver, J. J., & Gaver, K. M. (1998). The relation between nonrecurring accounting transactions and CEO cash compensation. *Accounting Review*, *73*(2), 235–253.

Gomez-Mejia, L. R., Tosi, H., & Hinkin, T. (1987). Managerial control, performance, and executive compensation. *Academy of Management Journal*, *30*(1), 51–70.

Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. *Advances in Experimental Social Psychology*, 61–139.

Grubbs, F. (1969). Procedures for detecting outlying observations in samples. *Technometrics*, *11*(1), 1–21.

Hains, S. C., Hogg, M. A., & Duck, J. M. (1997). Self-categorization and leadership: Effects of group prototypicality and leader stereotypicality. *Personality and Social Psychology Bulletin*, *23*, 1087–1099.

Hogg, M. A. (1996). Social identity and group cohesiveness. In: E. H. Witte & J. A. Davis (Eds), *Understanding group behavior, Vol. 2: Small group processes and interpersonal relations* (Vol. 2, pp. 227–253). Mahwah, NJ: Erlbaum.

Iyengar, R. J. (2003). Executive compensation in perennially loss-making firms. *Finance India*, *17*(1), 199–214.

Janis, I. (1972). *Victims of groupthink*. Boston: Houghton-Mifflin.

Lambert, R. A., & Larcker, D. F. (1987). An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research*, *25*(Suppl.), 85–125.

Lorsch, J. W., & MacIver, E. (1989). *Pawns or potentates: The reality of America's corporate boards*. Boston: Harvard Business School Press.

Lublin, J. S. (2005). CEO bonuses rose 46.4% at 100 big firms in 2004. *Wall Street Journal,* February 25, p. A1.

Moscovici, S., & Zavalloni, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology*, *12*, 125–135.

Myers, D. G., & Arenson, S. J. (1972). Enhancement of dominant risk tendencies in group discussion. *Psychological Science*, *6*, 10–19.

Myers, D. G., & Bishop, G. D. (1970). Discussion effects on racial attitudes. *Science*, *169*, 778–779.

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*, 602–627.

Mynatt, C., & Sherman, S. J. (1975). Responsibility attribution in groups and individuals: A direct test of the diffusion of responsibility hypothesis. *Journal of Personality and Social Psychology*, *32*, 1111–1118.

Natarajan, R. (1996). Stewardship value of earnings components: Additional evidence on the determinants of executive compensation. *Accounting Review*, *71*(1), 1–22.

Newman, H. A., & Mozes, H. A. (1999). Does the composition of the compensation committee influence CEO compensation practices? *Financial Management*, *28*(3), 41–53.

O'Reilly, C. A., Main, B. G., & Crystal, G. S. (1988). CEO compensation as tournament and social comparison: A tale of two theories. *Administrative Science Quarterly*, *33*(2), 257–274.

Perel, M. (2003). An ethical perspective on CEO compensation. *Journal of Business Ethics*, *48*(4), 381–391.

Pruitt, D. G. (1971). Choice shifts in group discussion: An introductory review. *Journal of Personality and Social Psychology*, *20*(3), 339–360.

Sandelands, L. E., Brockner, J., & Glynn, M. A. (1988). If at first you don't succeed, try, try again: Effects of persistence–performance contingencies, ego involvement, and self-esteem on task persistence. *Journal of Applied Psychology*, *73*, 208–216.

Sheikholeslami, M. (2001). EVA, MVA, and CEO compensation. *American Business Review*, *19*(1), 13–17.

Smith, E. R., & Mackie, D. M. (2000). *Social Psychology* (2nd ed). Philadelphia: Psychology Press, Taylor & Francis.

Sniezek, J. A., & Henry, R. A. (1989). Accuracy and confidence in group judgment. *Organizational Behavior and Human Decision Processes*, *43*, 1–28.

Staw, B. (1980). Rationality and justification in organizational life. In: B. Staw & L. Cummings (Eds), *Research in Organizational Behavior*. Greenwich, CT: JAI Press.

Staw, B. M. (1976). Knee-deep in the big muddy – A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, *16*(1), 27.

Stefansky, W. (1972). Rejecting outliers in factorial designs. *Technometrics*, *14*, 469–479.

TIAA-CREF. (2002). TIAA-CREF policy statement on corporate governance. Retrieved on December 3, 2002.

Tosi, H. L., Werner, S., Katz, J. P., & Gomez-Mejia, L. R. (2000). How much does performance matter? A meta-analysis of CEO pay studies. *Journal of Management*, *26*(2), 301–339.

Whyte, G. (1993). Escalating commitment in individual and group decision making: A prospect theory approach. *Organizational Behavior and Human Decision Processes*, *54*(3), 430–455.

Young, M. N., & Buchholtz, A. K. (2002). Firm performance and CEO pay: Relational demography as a moderator. *Journal of Managerial Issues*, *14*(3), 296–313.

Zajac, E. J., & Westphal, J. D. (1996). Who shall succeed? How CEO/Board preferences and power affect the choice of new CEOs. *Academy of Management Journal*, *39*(1), 64–90.

Zimbardo, P. G. (1970). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In: W. J. Arnold & D. Levine (Eds), *Nebraska Symposium on Motivation* (Vol. 17). Lincoln, Nebraska: University of Nebraska Press.

APPENDIX A. CORPORATE GOVERNANCE – A VERY SHORT OVERVIEW

What is Corporate Governance?

> Corporate governance is a hefty-sounding phrase that really just means oversight of a company's management – making sure the business is run well and investors are treated fairly. (Burns, *Wall Street Journal*, October 27, 2003)

Publicity traded companies are those whose stock is traded in a public forum, usually over the New York Stock Exchange (NYSE), the American Exchange (AMEX), National Association of Securities Dealers Automated Quotation System (NASDAQ), or other regional exchanges such as Philadelphia or San Francisco. As such, any company can literally have thousands of "owners".

It is difficult for a company to be managed simultaneously by potentially thousands of different owners; therefore, the owners or stockholders elect a board of directors as their representatives. Board sizes vary with an average of 9–11 members.

The board of directors hires management, such as the chief executive officer (CEO), chief financial officer (CFO), and other vice-presidents, to run the company, but the board *oversees* their activities. This oversight is often conducted within a sub-committee of the board, such as the audit committee, the compensation committee, or the nominating committee. As an example, selected members of the board may be on the compensation committee. Their job is to determine the compensation of the CEO, which is a significant fiduciary duty as a board member.

> The board's most important job is hiring, firing, and setting compensation for a company's chief executive, who runs the company day-to-day. (Burns, *Wall Street Journal*, October 27, 2003)

The membership of the board is often comprised of the CEO or other firm insiders, CEO's from other industries, bankers, retired politicians, academicians, and professional directors (often representing mutual or retirement funds). *Note: Although often on the board of directors, the inside CEO cannot be a member of his own compensation committee.* Board members provide not only oversight but also expertise and advice, often meeting three to five times a year in addition to the (usual) legally required once a year meeting.

Source: Information and statistics taken from the *Wall Street Journal* article by Judith Burns, October 27, 2003.

APPENDIX B

Instructions

Please read the following and the attached, then answer all the questions on the following pages to the best of your ability.

Thank you for your time and assistance.

The Board of Directors is the governing body for a publicly held corporation. The board represents the shareholders, decides the major investment and social policies for a company, and hires and determines the compensation of the executive management.

In this case, you serve on the board of directors of Putt Company. This is not your full-time employment. Please read the details of the attached case for a description of your occupation. One of your duties while serving on the board of directors is to serve on the compensation committee.

TIAA-CREF, a major retirement pension fund in the United States, describes the importance of this function in their 2002 policy statement as such

> The governance of the executive compensation process is a critically important and highly visible responsibility of the board of directors of a corporation. In a real sense, it represents a window through which the effectiveness of the board may be viewed. TIAA-CREF (2002)

Subject #____________

Demographic Questionnaire

This section captures basic information related to you, the survey participant.

1. Age in years ____________
2. Gender (please circle) male female
3. Number of years of full-time employment ____________
4. Please list any degrees obtained (e.g., B.S. in B.A.)

____

____

____

5. Please list any certifications obtained (e.g., C.P.A.)

6. Would you describe your current position as

a. staff
b. front-line supervisor
c. mid-level supervisor
d. junior executive
e. senior executive

APPENDIX C

Case 1

You are the Chief Executive Officer (CEO) of ABC Company. You are on the board of directors of Putt Company, an industrial company that manufactures golf equipment. Within the board of directors, one committee on which you serve is the compensation committee.

Your company, ABC Company, does not perform any services for Putt, nor does it anticipate doing so. You serve on the compensation committee of the board of directors for Putt Company as an independent director. Serving with you on the compensation committee are three other members: Two are also CEOs of other companies and one is an academic at a large eastern business school.

Your personal information

- You are the CEO of ABC Company.
- Your personal compensation as CEO of ABC is in Qwerts, a non-denominational monetary unit.
- You currently make 70 Qwert as the CEO of ABC.

Putt Company information

- The golf equipment industry grew 10% this past year.
- Putt Company grew at a 6% pace.
- Putt Company's closet competitor grew at a rate of 10%.
- Last year's earnings per share for Putt Company was 0.100 Qwert. This year's earnings per share for Putt Company is 0.106 Qwert.

- The size of Putt Company is comparable to the industry average, as is the total sales volume.
- Putt Company's operating margins and net income levels are below industry averages.

The compensation committee of the board of Putt Company performs an annual compensation review of the CEO. Your task as a member of this committee is to set the compensation level of the CEO in Qwerts.

The compensation level you decide will be kept private and confidential.

Based on the information provided, what compensation in Qwerts will you award the CEO of Putt Company?

____________*Qwerts*

Case 2

You are the Chief Executive Officer (CEO) of DEF Company. You are on the board of directors of Putt Company, an industrial company that manufactures golf equipment. Within the board of directors, one committee on which you serve is the compensation committee.

Your company, DEF Company, does not perform any services for Putt, nor does it anticipate doing so. You serve on the compensation committee of the board of directors for Putt Company as an independent director. Serving with you on the compensation committee are three other members: Two are also CEOs of other companies and one is an academic at a large eastern business school.

Your personal information

- You are the CEO of DEF Company.
- Your personal compensation as CEO of DEF is in Qwerts, a non-denominational monetary unit.
- You currently make 90 Qwert as the CEO of DEF.

Putt Company information

- The golf equipment industry grew 10% this past year.
- Putt Company grew at a 6% pace.
- Putt Company's closet competitor grew at a rate of 10%.
- Last year's earnings per share for Putt Company was 0.100 Qwert. This year's earnings per share for Putt Company is 0.106 Qwert.
- The size of Putt Company is comparable to the industry average, as is the total sales volume.

- Putt Company's operating margins and net income levels are below industry averages.

The compensation committee of the board of Putt Company performs an annual compensation review of the CEO. Your task as a member of this committee is to set the compensation level of the CEO in Qwerts.

The compensation level you decide will be kept private and confidential.

Based on the information provided, what compensation in Qwerts will you award the CEO of Putt Company?

____________*Qwerts*

Case 3

You are the Chief Executive Officer (CEO) of GHI Company. You are on the board of directors of Putt Company, an industrial company that manufactures golf equipment. Within the board of directors, one committee on which you serve is the compensation committee.

Your company, GHI Company, does not perform any services for Putt, nor does it anticipate doing so. You serve on the compensation committee of the board of directors for Putt Company as an independent director. Serving with you on the compensation committee are three other members: Two are also CEOs of other companies and one is an academic at a large eastern business school.

Your personal information

- You are the CEO of GHI Company.
- Your personal compensation as CEO of GHI is in Qwerts, a non-denominational monetary unit.
- You currently make 110 Qwert as the CEO of GHI.

Putt Company information

- The golf equipment industry grew 10% this past year.
- Putt Company grew at a 6% pace.
- Putt Company's closet competitor grew at a rate of 10%.
- Last year's earnings per share for Putt Company was 0.100 Qwert. This year's earnings per share for Putt Company is 0.106 Qwert.
- The size of Putt Company is comparable to the industry average, as is the total sales volume.

- Putt Company's operating margins and net income levels are below industry averages.

The compensation committee of the board of Putt Company performs an annual compensation review of the CEO. Your task as a member of this committee is to set the compensation level of the CEO in Qwerts.

The compensation level you decide will be kept private and confidential.

Based on the information provided, what compensation in Qwerts will you award the CEO of Putt Company?

___________*Qwerts*

Case 4

You are a business professor at a large eastern U.S. college. You are on the board of directors of Putt Company, an industrial company that manufactures golf equipment. Within the board of directors, one committee on which you serve is the compensation committee.

Your university does not perform any services for Putt, nor does it anticipate doing so. You personally do not perform any services, such as consulting, for Putt. You serve on the compensation committee of the board of directors for Putt Company as an independent director. Serving with you on the compensation committee are three other members who are chief executive officers (CEOs) of other companies.

Your personal information

- You are a business professor at a large eastern U.S. college.
- Your personal compensation as a professor is in Qwerts, a non-denominational monetary unit.
- You currently make 30 Qwert as a professor.

Putt Company information

- The golf equipment industry grew 10% this past year.
- Putt Company grew at a 6% pace.
- Putt Company's closet competitor grew at a rate of 10%.
- Last year's earnings per share for Putt Company was 0.100 Qwert. This year's earnings per share for Putt Company is 0.106 Qwert.
- The size of Putt Company is comparable to the industry average, as is the total sales volume.
- Putt Company's operating margins and net income levels are below industry averages.

The compensation committee of the board of Putt Company performs an annual compensation review of the CEO. Your task as a member of this committee is to set the compensation level of the CEO in Qwerts.

The compensation level you decide will be kept private and confidential.

Based on the information provided, what compensation in Qwerts will you award the CEO of Putt Company?

__________*Qwerts*

APPENDIX D

Subject #__________

Post-Case Questionnaire – Individual

1. Describe your role in this case
   a. A chief executive officer (CEO) serving on the board of directors of Putt Company.
   b. The CEO of Putt Company.
   c. A business school professor serving on the board of directors of Putt Company.
2. On a scale of 1–7, rate Putt Company's performance.

1	2	3	4	5	6	7
below			average			above

3. Please rate the difficulty in determining the compensation level.

1	2	3	4	5	6	7
difficult			average			easy

4. Please rate your confidence in your answer.

1	2	3	4	5	6	7
not very			somewhat			very

5. Please rate the importance of the compensation decision.

1	2	3	4	5	6	7
not very			somewhat			very

APPENDIX E

Subject #__________

Post-Case Questionnaire – Individual in Team as Chair

1. Describe your role in this case
   a. A chief executive officer (CEO) serving on the board of directors of Putt Company.
   b. The CEO of Putt Company.
   c. A business school professor serving on the board of directors of Putt Company.
2. On a scale of 1–7, rate Putt Company's performance.

1	2	3	4	5	6	7
below			average			above

3. Please rate the difficulty in determining the compensation level.

1	2	3	4	5	6	7
difficult			average			easy

4. Please rate your confidence in your answer.

1	2	3	4	5	6	7
not very			somewhat			very

5. Please rate your agreement with the committee's decision.

1	2	3	4	5	6	7
did not agree						strongly agreed

6. Please rate the importance of the compensation decision.

1	2	3	4	5	6	7
not very			somewhat			very

7. Please rate the effectiveness of the committee chair.

1	2	3	4	5	6	7
not very			somewhat			very

8. Rate your agreement with the following statement: The committee possessed sufficient information to make a quality decision.

   1 2 3 4 5 6 7
   do not agree — strongly agree

9. Rate your agreement with the following statement: I was concerned with decision agreement and consensus with my team members.

   1 2 3 4 5 6 7
   do not agree — strongly agree

10. Please rate the degree to which you relied upon the arguments or beliefs of others.

    1 2 3 4 5 6 7
    very little — very much

11. In your opinion, was the group discussion.

    1 2 3 4 5 6 7
    collegial? — argumentative?

12. In your opinion, to what degree was the compensation decision based upon the accounting performance indicated?

    1 2 3 4 5 6 7
    very little — very much

13. Please rate the degree to which you feel your opinion or belief was heard.

    1 2 3 4 5 6 7
    very little — very much

14. As the committee chair, please rate the support of the committee in regards to the compensation decision reached.

    1 2 3 4 5 6 7
    very little — very much

15. Please rate the support of the committee in regard to your position as committee chair.

    1 2 3 4 5 6 7
    very little — very much

16. Did you feel the committee chair position was an effective position in regard to determining a group decision?

1	2	3	4	5	6	7
no						yes

17. As committee chair, please rate your management style.

1	2	3	4	5	6	7
autocratic						consensus driven

18. Please comment, if applicable, on the group decision process.

____
____
____
____

19. Additional comments:

____
____
____

Subject #__________

Post-Case Questionnaire – Individual in Team as Member

1. Describe your role in this case
   a. A chief executive officer (CEO) serving on the board of directors of Putt Company.
   b. The CEO of Putt Company.
   c. A business school professor serving on the board of directors of Putt Company.
2. On a scale of 1–7, rate Putt Company's performance.

1	2	3	4	5	6	7
below			average			above

3. Please rate the difficulty in determining the compensation level.

1	2	3	4	5	6	7
difficult			average			easy

4. Please rate your confidence in your answer.

1	2	3	4	5	6	7
not very			somewhat			very

5. Please rate your agreement with the committee's decision.

1	2	3	4	5	6	7
did not agree						strongly agreed

6. Please rate the importance of the compensation decision.

1	2	3	4	5	6	7
not very			somewhat			very

7. Please rate the effectiveness of the committee chair.

1	2	3	4	5	6	7
not very			somewhat			very

8. Rate your agreement with the following statement: The committee possessed sufficient information to make a quality decision.

1	2	3	4	5	6	7
do not agree						strongly agree

9. Rate your agreement with the following statement: I was concerned with decision agreement and consensus with my team members.

1	2	3	4	5	6	7
do not agree						strongly agree

10. Please rate the degree to which you relied upon the arguments or beliefs of others.

1	2	3	4	5	6	7
very little						very much

11. In your opinion, was the group discussion.

1	2	3	4	5	6	7
collegial?						argumentative?

12. In your opinion, to what degree was the compensation decision based upon the accounting performance indicated?

1	2	3	4	5	6	7
very little						very much

13. Please rate the degree to which you feel your opinion or belief was heard.

    1 2 3 4 5 6 7
    very little very much

14. As a member of the committee, please rate the support of the committee in regards to the compensation decision reached.

    1 2 3 4 5 6 7
    very little very much

15. Please rate the support of the committee in regard to the position of committee chair.

    1 2 3 4 5 6 7
    very little very much

16. Did you feel the committee chair position was an effective position in regard to determining a group decision?

    1 2 3 4 5 6 7
    no yes

17. Please rate the management style of the committee chair.

    1 2 3 4 5 6 7
    autocratic consensus driven

18. Please comment, if applicable, on the group decision process.

    ____
    ____
    ____
    ____

19. Additional comments:

    ____
    ____
    ____

STRATEGY, STRUCTURE, PERFORMANCE MANAGEMENT, AND ORGANIZATIONAL OUTCOME: APPLICATION OF BALANCED SCORECARD IN CANADIAN HEALTH CARE ORGANIZATIONS

Yee-Ching Lilian Chan and Alfred Seaman

ABSTRACT

This article looks at the alignment of performance management system with the strategy, structure, and organizational outcome in Canadian health care organizations. In this study, balanced scorecard is the framework adopted for assessing the health care organization's performance management system (PMS) and outcome. CEO and clinical unit managers were surveyed for their perceptions on their organization's strategy, autonomy structure, PMS, and organizational performance. Path analysis was the methodology used in examining the relationship about the above organizational variables. The results indicate that patient satisfaction is the primary and most significant perspective of the depicted balanced scorecard in organizational performance. Patient satisfaction

Advances in Management Accounting, Volume 17, 151–180
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017007

and research criteria, on the other hand, are the significant perspectives of a balanced scorecard in an organization's PMS, which are linked to strategy, autonomy structure, and organizational performance. Moreover, the results show that the strategy/structure links operated as suggested. Surprisingly, strategy on service innovation has a negative impact on the organizational outcome of patient satisfaction. Uncertainty from continuous development and organizational change in pursuing service innovation and cost-cutting measures in response to fiscal constraints are plausible explanations of the adverse impact reported.

INTRODUCTION

The significance of the health care sector in the Canadian economy is evident. As reported by the Canadian Institute of Health Information, Canada's expenditure on health care has increased from $106.7 billion or $3,440 per capita in 2001 to $148.0 billion or $4,548 per capita in 2006, an increase from 9.6 to 10.3 percent of Canada's Gross Domestic Product in six years.[1] The increase in total and per capita health care spending averages 6.5 and 5.4 percent per annum, respectively. Likewise, public sector expenditure on hospitals has increased from $29.6 billion to $40.1 billion over the same time period,[2] an average increase of 5.9 percent per annum. Expenditure on hospitals remains the largest component of health care spending, about 29.8 percent, in Canada.

Given the substantial amount of resources spent on hospitals, there is an increasing demand on the accountability of hospital performance. Ontario Hospital Association and the Ministry of Health and Long-Term Care in Ontario have adapted the balanced scorecard as an external accountability system in their work on *Hospital Report*, the objective of which is "to help people in Ontario better understand and assess the performance of their local hospitals and of the province's hospitals as a whole. It also supports efforts by hospitals to improve the quality of their services."[3] The first *Hospital Report '98– A System-Wide Review of Ontario Hospitals* (1998) was published in November 1998. It provided information on the performance of Ontario acute care hospitals at both the regional and the provincial levels. The second *Hospital Report '99– A Balanced Scorecard for Ontario Acute Care Hospitals* (1999) was released in December 1999. In addition to scorecards at the system and group levels, *Hospital Report '99* included individual balanced scorecard for each of the 89 Ontario hospitals.

The *Hospital Report* series for acute care hospitals is currently in its ninth publication and the scorecard has evolved over time. Besides acute care hospitals, the scorecard framework has been extended to cover complex continuing care, emergency department care, rehabilitation, and mental health.

Apart from achieving a greater level of public accountability, another objective of publishing individual hospital's balanced scorecard is to assess performance and stimulate improvement efforts (Pink et al., 2001). The scorecard of the *Hospital Report* for acute care hospitals currently includes four performance quadrants: clinical utilization and outcome, financial performance and condition, patient satisfaction,[4] and system integration and change, which examines various aspects of a hospital's operation and performance. The performance indicators included in the *Hospital Report*, however, may not be in alignment with individual hospital's strategy. Yap, Siu, Baker, Brown, and Lowi-Young (2005) report that the 22 acute-care hospitals with organizational balanced scorecards all have a fairly high average number of indicators, which are different from those in the *Hospital Report*. This suggests that hospitals would develop balanced scorecards with different indicators in order to reflect their unique organizational strategies. In view of that balanced scorecard is not only useful in assessing performance and enhancing public accountability of Ontario hospitals, but also valuable in instigating a hospital's strategic management system.

Since there is consensus in the management and accounting literature that the match, or "fit", between an organization's strategy, structure, and management processes impacts organizational survival (Miles & Snow, 2003), this study examines the role of the balanced scorecard in establishing the strategy, structure, and performance linkage. Instead of adopting the common organizational performance measures of effectiveness and efficiency, "fit" is assessed in terms of a modified version of the balanced scorecard in the *Hospital Report*. The results of this study provide empirical evidence of the nature of the linkages between strategy, structure, performance management, and organizational outcomes of Canadian health care organizations. The evidence also sheds some light on the use of performance indicators, financial and non-financial indicators, in balanced scorecards of Canadian health care organizations.

The organization of the article is as follows. We will first provide an overview of the application of balanced scorecard in the health care sector, which is followed by a review of the literature relevant to the variables of interest in the model. We will then describe the construct of the variables and sample selection. The final sections will include a discussion of the results with conclusion and limitations of the study.

APPLICATION OF BALANCED SCORECARD IN THE HEALTH CARE SECTOR

In their inaugural paper, Kaplan and Norton (1992) advocate the balanced scorecard as an improvement to the then existing performance measurement system, which has a focus on financial performance measures. With their continuous research, Kaplan and Norton (1996a, 1996b, 1996c, 2001a, 2001b, 2001c) promote the adoption of the balanced scorecard as a strategic management system. They argue that the balanced scorecard framework can be effective in translating an organization's vision into specific strategies with business plans set, resources allocated, and performance monitored to facilitate ongoing strategic review and learning. Kaplan (2001) continues to assert that the balanced scorecard is a useful tool in managing not-for-profit organizations (NFPs) because it bridges the gap from the organization's vague mission and strategy statements to day-to-day operational measures; it facilitates a process by which NFPs can achieve strategic focus; it shifts an NFP's focus from programs and initiatives to the outcomes that the programs and initiatives are supposed to accomplish; and it helps NFPs to avoid the illusion that they have a strategy because they are managing a diverse and non-cumulative set of programs and initiatives. Moreover, balanced scorecard enables NFPs to align initiatives, departments, and individuals to work in ways that reinforce each other so that dramatic performance improvements can be achieved.

Apart from the *Hospital Report*, balanced scorecard has been implemented in the Henry Ford Health System (Sahney, 1998) since the mid-1990s in measuring and improving performance of the managed care organizations, which were subject to extensive competition as well as pressure to cut costs and improve financial performance while maintaining quality of health services. In both cases, the balanced scorecard was used to assess performance of hospitals/organizations in a system with an objective to engage in performance improvement.

More recently, with the establishment of regional health authorities in a number of provinces in Canada as well as the mandate of a Performance Agreement and/or Accountability Agreement between regional health authorities and their provincial ministry of health, there has been a growing interest in implementing balanced scorecard for assessing system performance. For instance, British Columbia's Vancouver Coastal Health[5] as well as Alberta's Chinook Health Region[6] and David Thompson Health Region[7] have implemented the balanced scorecard as part of their performance measurement and accountability framework.

The use of balanced scorecard in hospitals as part of their performance management and strategic management system has increased substantially over the past 10 years as well. The Women's College Hospital in Toronto and the University of Alberta Hospitals (Baker & Pink, 1995) are among the pioneers which have adopted report cards and dash boards for performance measurement. These scorecards incorporated the concern of the hospitals' stakeholders, focused on the hospitals' processes, and included both financial and non-financial indicators for performance measurement. Peel Memorial Hospital in Brampton also adopted the balanced scorecard in 1995 (Harber, 1998) because the hospital had a tired Mission Statement and the employees were unclear about the organization's strategic direction and the linkage of various programs and initiatives undertaken. The balanced scorecard at Peel Memorial Hospital included six categories of business with 23 data elements where patient and community focus was the center of the framework. The other five categories of business were management leadership, human resource management, patient care process management, quality tools and information utilization, and performance results, and their interrelationship was identified in the framework. Although the development of the balanced scorecard was a major undertaking and the development of performance measures a challenge, the implementation of balanced scorecard at Peel Memorial Hospital was quite successful as there were increases in both patient and staff satisfaction levels. Furthermore, the balanced scorecard provided Peel Memorial Hospital the ability to translate the hospital's strategic objectives into a coherent set of performance measures as well as to align the seemingly disparate elements with organizational objectives.

As described in Zelman, Pink, and Matthias's (2003) study, the balanced scorecard is well into its growth phase in the health care sector. The use of balanced scorecard in the health care sector diverges from evaluating organizational performance to developing clinical pathway. Applications of balanced scorecard can be found in hospital systems, hospitals, long-term care facilities, psychiatric centers, university departments, and government units. Among the increasing number of hospitals and health care units which have implemented balanced scorecard are Duke's Children's Hospital (Meliones, 2000), Duke's Women's Services Clinical Business Unit (Jones & Filip, 2000), Mayo Clinic (Curtright, Stolp-Smith, & Edell, 2000), St. Elsewhere Hospital (Kershaw & Kershaw, 2001), and Yale New Haven System Hospital (Gumbus, Bellhouse, & Lyons, 2003). Moreover, SMDC Health System in Minnesota[8] has applied a strategy mapping process and developed a strategy map and balanced scorecard that helps the entire organization understand which strategic drivers are critical to achieving service excellence, clinical excellence, and management excellence while satisfying various stakeholders' needs. By and large, administrators of hospitals, which have implemented balanced scorecard, commented that the framework was essential to connecting clinical and organizational practices, outcomes, quality, value and cost; aligning performance measurement with meeting organization's vision, primary value, core principles, and operational strategies; and achieving a balance between productivity and quality. The anecdotal reports cited thus far are consistent with Chan and Ho's (2000) findings that the majority of the hospital administrators perceived their hospitals' undertaking of balanced scorecard as moderately successful, and they expected their hospitals' use of balanced scorecard would have more significant changes over the next five years. In fact, more and more Canadian hospitals[9] have adopted balanced scorecard as their strategic management system. Moreover, the Ministry of Health and Long-Term Care in Ontario has recently published a report on *Aligning Performance Measurement with Corporate Strategy: A Toolkit for Ontario's Hospitals* (Paul et al., 2006) which provides step-by-step instructions on how to design a hospital's strategy map (Kaplan & Norton, 2001a, 2001b), how to select and evaluate performance indicators, how to use the balanced scorecard for quality improvement, how to set performance targets, and how to manage the change. Evidently, the growing use of balanced scorecard in the health care sector and the anecdotes reported thus far, coupled with endorsement from government agency, support Kaplan's (2001) advocacy that balanced scorecard is a useful management tool for NFPs and health care organizations.

As discussed earlier, an objective of this study is to examine the role of balanced scorecard in establishing the linkage among strategy, structure, performance management, and organizational outcome of Canadian health care organizations. What follows is a review of the literature relevant to the variables of interest in our modeling of the strategy, structure, and performance linkage.

LITERATURE REVIEW

Strategic Choice

Porter and Teisberg (2004) argue that competition in the health care sector in the United States should change. As in other manufacturing and service sectors, the locus of competition in the health care sector should be switched from "Who pays?" to "Who provides the best value?" Since "cost leadership" strategy in the health care sector may be perceived as equivalent to inferior health services, health care providers should continue to work on cost containment while developing clear strategies around their organization's unique expertise and to redirect investment to facilities in areas where they can become distinctive. Although it is implausible for hospital executives to decide on what not to do, yet according to Porter and Teisberg (2004), they have to develop their organization's uniqueness and expertise in certain conditions and treatments rather than try to be all things to all people. The strategy of differentiation and focus, as recommended by Porter and Teisberg (2004), may not gain complete acceptance in Canada's health care system which is built upon the principles of public administration, comprehensiveness, universality, portability, and accessibility (Canada Health Act, 1984, Chapter 6, Section 7). Nonetheless, Canadian hospitals have somehow decided what types of health services are to be provided to their community. For instance, rural community hospitals in Canada, in general, still have to provide basic acute care and many have elected not to provide tertiary care because they do not have the resources needed to support specialized medical services. A number of community hospitals and teaching hospitals in suburban and urban centers, on the other hand, have elected to provide specialized medical programs to differentiate their services from other hospitals as well as to receive incentive funding from the government for these new and expanded programs.

In the past, Canadian hospitals have reduced the level of service as a temporary measure to cope with insufficient funding. It is inconceivable that specific types of health services will be discontinued permanently in response to escalating costs. Canadian hospitals can still work to develop expertise in specific areas of health services with approval and financial support from the government. Thus, many Canadian hospitals will continue to be "all things to all people" while some will pursue the "differentiation" strategy by developing and growing specialized medical programs. This is consistent with the proposition of Shortell, Anderson, Erickson, and Mitchell (1996) that the focus of service innovation is an important dimension of competitive strategy in the hospital sector. In fact, Abernethy and Lillis (2001) suggest that a hospital's commitment to service innovation falls on a continuum. At one end of the continuum, there are hospitals which respond quickly to changes in technology and market demand. These hospitals gain pride from having the latest technology and continually developing new service offerings. Conversely, there are hospitals at the other end of the continuum which provide a relatively narrow and stable set of services. These hospitals do not actively seek opportunities for adopting new technology or expanding service offerings. In addition to patient and service innovation, hospitals may broaden the scope of their strategy to include innovation in research and teaching. Thus, Canadian hospitals can either pursue a strategy of innovation or a strategy of the *status quo*, that is, conduct business as usual. This is analogous to Miles and Snow's (2003) typology of organizations as *Prospectors* and *Defenders* which represent the opposite ends of the innovation continuum based on an organization's strategic choice.

By and large, it is unlikely that Canadian hospitals can sustain their operations as a *Reactor* (Miles & Snow, 2003, pp. 81–93), which is an unstable organization type that fails to implement a consistent response mechanism in face of changing environments, because the Public Hospital Act (R.S.O. 1990, Chapter P. 40) explicitly specifies the governance and management of Canadian hospitals as well as their power, rights, and responsibilities. Furthermore, given the fiscal constraints imposed by both the provincial and the federal governments, Canadian hospitals which elect to be *Analyzers* (Miles & Snow, 2003, pp. 68–80) may not be able to compete for incentive funding for new programs against *Prospectors*, which are always looking out for market opportunities. Consequently, an increase in funding to one *Prospector* hospital implies that other hospitals' share decreases even though the total annual expenditure on health care in Canada continues to increase over time. Thus, Canadian hospitals either have to pursue a differentiation/innovation strategy (*Prospector*) or to focus on operating at status quo in a cost-effective way (*Defender*).

INNOVATION STRATEGY AND ORGANIZATIONAL STRUCTURE

As suggested earlier, there is consensus in the management literature that organizational structure affects the effectiveness of strategy implementation. For organizations which pursue innovation as a strategic priority, they have to respond quickly to market forces by assessing the viability of current product/service offerings and developing new products/services to satisfy changing demands on a continuous basis (Slater & Olson, 2000). Their organizational structure must also facilitate the flow of information, both horizontally and vertically, in an efficient manner such that individual work units are encouraged to collaborate in their assigned tasks (Bouwens & Abernethy, 2000). The creation of autonomous work units is one effective way of ensuring an efficient flow of information within the organization (Scott & Tiessen, 1999). Furthermore, managers of autonomous work units can respond rapidly to changing market conditions as they acquire information (Kaplan & Atkinson, 1998) and collaborate with each other by sharing resources (Lei, Hitt, & Goldhar, 1996).

In the health care system, information flow poses additional challenges because most health care organizations adopt a bilateral structure with the chief executive/operating officer responsible for administration while the chief of staff is accountable for medical services. In most health care organizations, the senior management team lacks the clinical expertise to make decisions related to clinical matters while the medical professionals, with support from the chief of staff, have considerable autonomy over clinical processes and patient care outcomes. Thus, for health care organizations pursuing service innovation as a strategic priority, they need an organizational structure that facilitates functional coordination and effectively harnesses the knowledge and expertise of clinical unit managers when responding to changing market forces. In their study, Abernethy and Lillis (2001) found a positive relationship between the strategic focus on service innovation and the extent of autonomy granted to clinical units over output and resource management decisions.

Strategy, Autonomy, Performance Management, and Organizational Performance

Since the senior management team in most health care organizations lacks the clinical expertise to develop new services in response to changing market conditions, the management of clinical resources has to be delegated to the medical professionals. The creation of autonomous work units limits the ability of senior management to closely monitor the actions of clinical units (Merchant, 1998). Moreover, clinicians will strongly resist any types of bureaucratic controls that threaten their autonomy (Freidson, 1975) because they prefer to work under the norms and values imposed by the medical profession itself (Abernethy & Stoelwinder, 1991). In addition, increased autonomy to clinical units could be dysfunctional for the health care organizations due to goal incongruence (Kaplan & Atkinson, 1998). Hence, mechanisms need to be in place to hold clinical unit managers accountable for their actions. A performance management system (PMS), which incorporates measures for both resource management and outcome, is a tool that facilitates managing accountability of clinical unit managers.

Abernethy and Lillis (2001) found significant relationships between the level of autonomy and the extent of the use of resources as well as clinical management performance criteria. They argue that while strategy is a significant determinant of structural autonomy, it is more likely that the importance attached to performance measures is directly affected by delegation of autonomy to work unit management and not as much by strategic choice. Results of their study support the proposition that the relationship between the focus on service innovation as a strategic priority for health care organizations and the extent of the use of performance measures on resource and clinical management performance is an indirect one operating via structural autonomy. Finally, as discussed earlier, there is consensus that organizational structure affects the effectiveness of strategy implementation and the best fit between strategy and structure is likely to yield a positive impact on organizational performance. Thus, for health care organizations pursuing a strategic focus on service innovation, their performance will be enhanced if a high level of autonomy is granted to their clinical units. In fact, the use of clinical performance measures is found to have a positive effect on the achievement of organizational effectiveness, whereas the use of resource performance measures has a positive effect on the achievement of organizational efficiency (Abernethy & Lillis, 2001). Accordingly, the relationship between a health care organization's strategic focus on service innovation and performance is not a direct one but rather an indirect one operating via structural autonomy and the use of performance measures in monitoring accountability.

A critical premise of performance management is that it allows the organization to develop a set of reinforcing signals that direct a manager's attention to criteria that are of importance to the chosen strategy (Dixon, Nanni, & Vollman, 1990). Consequently, strategic choice will be a driver of the PMS. Noticeably, the PMS in the form of a balanced scorecard not only provides an opportunity to assess management performance but also to determine if management activities support and promote the strategic objectives and organizational outcomes of the health care organizations. Thus, an objective of this article is to investigate the use and usefulness of balanced scorecard in performance management of Canadian health care organizations, as depicted in the hypothetical model of the strategy, structure, performance management, and organizational outcome linkage included in Fig. 1.

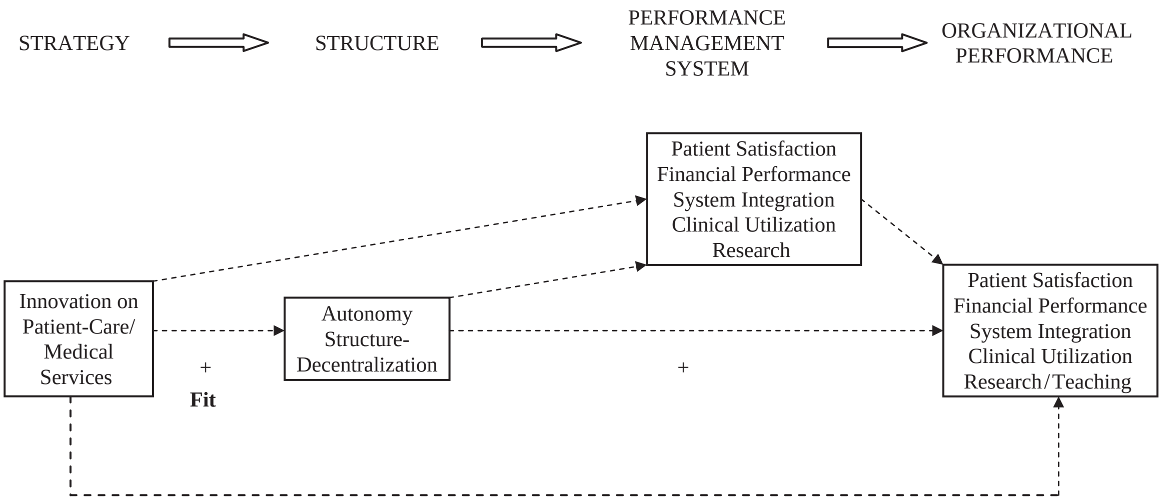


Fig. 1. Hypothetical Model.

METHODOLOGY

Measurement of Variables

Measures for this study were developed from established instruments and from pilot testing. Items in the survey instrument developed to measure the variables of interest are given in appendix.

(1) Strategy (Service Innovation)

To capture the extent of innovation in patient-care and medical service offerings, we modified the item in Abernethy and Lillis' (2001) study and applied the Miles and Snow (2003) typology, i.e., the *prospector* typology describes innovators who are early adaptors constantly seeking and finding new market opportunities while the *defender* typology provides the description for non-innovators who conduct business as usual at *status quo*. Innovation of patient-care and medical services was measured on a continuum with the extremes being the *prospector*, with many changes in services, and *defender*, with few or no changes in services. CEOs, as chief administrators of the hospitals, lead the development of their organization's vision and mission statements. They are also knowledgeable about their hospital's strategies. Accordingly, the survey instrument includes a question which asks CEOs to compare their hospital to the two descriptors included in appendix (1) and provide a rating of the degree of change in their organization's patient-care and medical service offerings on a scale of 1–7 on the continuum.

(2) Structural Autonomy

An abbreviated form of the Aston concentration of authority scale (Inkson, Pugh, & Hickson, 1970; Pugh, Hickson, Hinings, & Turner, 1968) is used to measure the delegation of decision-making in an organization. The survey instrument includes a question which asks clinical unit managers to identify the most junior job level which has the authority to make 23 policy decisions, varying from the appointment and promotion of supervisors to the allocation of resources and determination of new services (appendix (2)). The position level of the policy decision-makers identified in the question ranges from 1 (supervisor) to 5 (individuals and agencies outside the organization). The sum of the scores of the 23 items is used to assess the degree of structural autonomy where higher score represents higher degree of centralization or more concentration of authority. The scores of the 23 items on structural autonomy are then reverse coded and added to develop the decentralization construct, i.e., higher score represents higher degree of decentralization.

(3) Performance Management System

This study adapted a modified version of the performance quadrants of the *Hospital Report* series to capture the PMS. The survey instrument includes a question which asks clinical unit managers to indicate the extent to which the 16 items in a list of performance measures are used to measure their performance on a scale of 1 (to a little extent) to 7 (to a great extent) (see appendix (3)). The items would likely fall into the four performance quadrants of the scorecard depicted in the *Hospital Report* series. The sum of the scores of the items included in each performance quadrant is the construct for the hospital's PMS.

(4) Organizational Performance

Similar to the construct for PMS, this study adapted a modified version of the *Hospital Report* series in measuring organizational performance. Since the CEO is more knowledgeable about the hospital's organizational performance, the survey instrument includes a question which asks CEOs to assess their hospital's organizational performance on a scale of 1(below average) to 7 (above average) (see appendix (4)). The sum of the scores of the items included in each performance quadrant is the construct for the hospital's organizational performance.

Accordingly, this study requires two stages of data collection. First, we have to send the questionnaire to CEOs of health care organizations asking about their organization's strategy on service innovation and performance. In the first questionnaire, we also request CEOs to provide us with names of clinical unit managers and/or medical program directors for a follow-up survey because in the bilateral management system of Canadian health care organizations, medical program directors are responsible for delivering medical services while clinical unit managers provide the administrative functions. In the second stage, we have to send another questionnaire to the clinical unit managers[10] of health care organizations in which the CEOs have responded to the first questionnaire. To the extent that clinical unit managers are close to the day-to-day operations of their units, they are considered an appropriate source to provide information on the decentralization of decision-making and the use of PMS in assessing their individual performance. Hence, we survey the clinical unit managers about their assessment on their organization's autonomy structure and PMS. This matched sample approach reduces common-rater bias in survey data used for path analysis (Kren, 1997).

Sample Selection and Characteristics

Since the size of an organization affects its degree of decentralization and autonomy structure, this study includes only larger health care organizations when examining the linkage among strategy, structure, performance management, and organizational outcome in Canadian health care organizations. Accordingly, from the *Guide to Canadian Health Care Facilities* (Canadian Healthcare Association, 2007), we identified a total of 127 health care organizations with at least 100 beds in nine provinces. Hospitals in Quebec are excluded from the sample because of the language barrier while hospitals from the three territories are excluded because of their relatively small size. Some of health care organizations included in the sample are facilities of regional health authorities while others are health care organizations with multiple hospital sites, e.g., University Health Network in Toronto, Ontario.

We sent the first questionnaire to the CEOs of these 127 health care organizations and 55 questionnaires were returned, yielding a response rate of 43.3 percent. With the names of clinical unit managers given by the CEOs and a search on the organizations' websites, we mailed the second questionnaire to 459 clinical unit managers, an average of eight questionnaires per organization. A total of 124 questionnaires were returned by the clinical unit managers, yielding a response rate of 27.0 percent. The number of questionnaires returned by the clinical unit managers per organization ranges from 1 to 5 with a median and a mode of two questionnaires per organization. These resulted in a usable paired final sample size of 33 health care organizations, yielding a response rate of 26.0 percent, which is comparable to other studies in the health care sector.

Comparisons of early and late respondents on their organizations' size, as measured by the number of employees and beds, were conducted to test for non-response bias. The results indicate that there is no significant difference between early and late respondents in the size of their organizations. Non-response bias is non-existent.

Further analysis of the data indicates that CEOs had been in their current position for an average of 4.9 years, whereas clinical unit managers held their position for an average of 5.2 years.

ANALYSIS AND RESULTS

Descriptive Statistics

The mean, standard deviation, theoretical, and actual range of the variables are presented in Table 1. By and large, the sample of 33 health care organizations included in this study made frequent changes in their sets of patient-care and medical services. They consistently attempted to pioneer in new areas of patient-care and medical services and responded quite rapidly to early signals of opportunities or needs for specialized patient-care and medical services. On the basis of the responses on the concentration of authority, this group of health care organizations has a decentralized structure where clinical unit managers have relative autonomy in making various policy decisions. A Cronbach (1951) α statistic of 0.85 for this measure of structural autonomy confirmed its reliability.

As shown in Table 2, the PMS of these health care organizations includes a fifth performance perspective on research in addition to the four

Table 1. Descriptive Statistics on Strategy, Structure, Performance Management System, and Organizational Performance (sample size = 33).

Variable	Mean	Standard Deviation	Actual Range	Theoretical Range
Strategy on service innovation				
Patient care	4.845	1.372	2–7	1–7
Autonomy structure				
Decentralization	88.090	8.611	63–102	23–115
Performance management system				
Patient satisfaction	13.606	5.744	3–21	3–21
Financial performance	21.818	8.229	5–35	5–35
Systems integration	18.212	6.224	4–28	4–28
Clinical utilization	8.242	2.926	2–14	2–14
Research	7.091	3.422	2–14	2–14
Organizational performance				
Patient satisfaction	24.878	6.259	9–34	5–35
Financial performance	18.394	3.929	10–25	4–28
Systems integration	8.909	2.097	4–12	2–14
Clinical utilization	13.242	3.072	7–21	3–21
Research/teaching	13.333	3.739	3–20	3–21

Table 2. Results of Factor Analyses for Performance Management System and Organizational Performance (sample size = 33).

Factor	Survey Items	Eigenvalue	Percentage of Variance Explained	Cronbach α
Patient satisfaction				
Performance management system	1, 2, 3	3.098	19.364	0.945
Organizational performance	1, 2, 3, 4, 9	4.546	26.739	0.914
Financial performance				
Performance management system	5, 6, 7, 8, 9	3.274	20.465	0.864
Organizational performance	5, 6, 7, 8	1.738	10.226	0.616
Systems integration				
Performance management system	12, 13, 14, 15	2.240	11.501	0.845
Organizational performance	13, 14	1.579	9.286	0.522
Clinical utilization				
Performance management system	10, 11	1.202	10.011	0.766
Organizational performance	10, 11, 12	2.406	14.155	0.796
Research/teaching				
Performance management system	4, 16	1.804	11.273	0.610
Organizational performance	15, 16, 17	2.619	15.407	0.834

performance quadrants depicted in the *Hospital Report*. Similarly, organizational performance for this sample of health care organizations is evaluated in terms of these five areas of patient satisfaction, financial performance, system integration, clinical utilization, and research/teaching. The fifth performance perspective on research/teaching reflects the characteristics of this sample of hospitals, larger teaching hospitals which focus not only on providing patient-care and medical services but also on teaching students in their medical programs and conducting research in different fields of medical science. Apparently, the *Hospital Report* series does not include this fifth performance perspective on research/teaching because of the diversity of hospitals participated in the initiative.

The percentage of variance explained by the constructs varies from 10.0 to 20.5 percent and 9.3 to 26.7 percent for the five performance perspectives of the PMS and organizational performance, respectively. The Cronbach α for all but one measure on research of the PMS is greater than 0.75, which confirms the reliability of the constructs of the four performance perspectives on patient satisfaction, financial performance, system integration, and clinical utilization. On the other hand, the Cronbach α for three of the five measures on organizational performance is about 0.800 or greater. This implies that measures for financial performance and system integration in assessing organizational performance may not be as reliable as the other three measures on patient satisfaction, clinical utilization, and research/teaching.

By and large, as shown in Table 1, clinical unit managers perceived all performance perspectives except system integration were used moderately in assessing their individual performance via their organization's PMS while CEOs, in general, indicated that their organization's performance on all aspects is about average.

Path Analysis and Model

The hypothetical model, Fig. 1, was tested using Lisrel 8 (Joreskog & Sorbom, 1993) and structural equation modeling techniques. A series of nested models (Anderson & Gerbing, 1988) was used to estimate the parameters of the model. Starting with the unconstrained model, Fig. 1, the path of least significance was constrained and the model re-estimated. Using the χ^2 test of difference to compare the two models, this process was continued until no further improvements could be attained. The result of this analysis is shown in Fig. 2. The fit statistics indicate a good fit of the data to the model: $\chi^2 = 0.4$, $p = 0.88$, $\mathrm{df} = 1$; $\mathrm{GFI} = 0.99$; $\mathrm{NFI} = 0.98$. All paths shown are standardized coefficients significant at $p < 0.05$.

Among the five performance perspectives on organizational performance, patient satisfaction is the only one with significant path coefficients from service innovation, autonomy structure, and decentralization as well as PMS (see Fig. 2). This affirms health care organization's mission which is to provide quality health services in patient care while other aspects of performance, including financial performance, clinical utilization, system integration, and research/teaching are of secondary importance. Thus, while balanced scorecard provides performance assessment on various aspects of an organization's performance, patient satisfaction remains the primary perspective of performance outcome to this sample of health care organizations. Likewise, clinical unit managers indicated that patient satisfaction and research are the two significant performance perspectives used in assessing their individual performance (see Fig. 2), whereas the other three: financial performance, system integration, and clinical utilization, are not as important. This again shows that not all performance perspectives of a balanced scorecard are equally important in assessing performance of

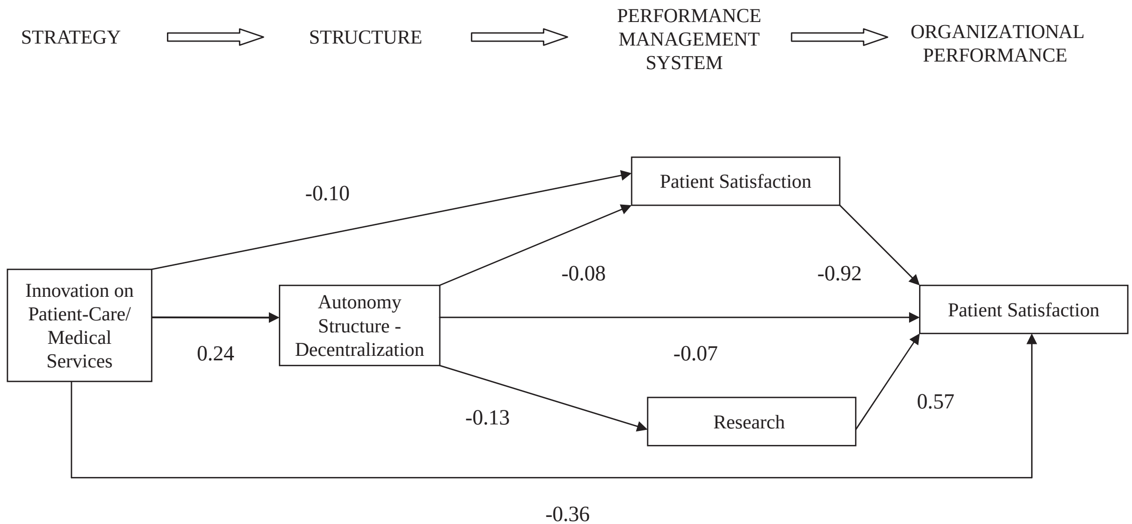


Fig. 2. Path Diagram.

clinical unit managers in health care organizations. For this sample of clinical unit managers, patient satisfaction and research are the focus in their organization's PMS because these performance perspectives are in alignment with their organization's mission of providing quality health services in patient care and pursuing a strategy on service innovation in which research is imperative.

As illustrated in Fig. 2, there are significant paths from the strategy of service innovation to decentralization (autonomy structure), to performance perspective of patient satisfaction in an organization's PMS (patient satisfaction) as well as to performance outcome of patient satisfaction in organizational performance evaluation (OPE patient satisfaction).[11] As a health care organization adopts the strategy of service innovation, more decision-making authority is delegated to clinical unit managers (path coefficient $= 0.24$). This is as expected and is generally supported in the management literature (Hambrick, 1983). However, since the correlation coefficient between the strategy of service innovation and PMS patient satisfaction is -0.135 and 77 percent of the variance is explained by the direct effect of a strategy of service innovation on PMS patient satisfaction (i.e., the path coefficient after controlling for autonomy structure is -0.10), autonomy structure is not an important intervening variable between a strategy of service innovation and PMS patient satisfaction. This suggests that it is the organization's strategy that drives the use of PMS patient satisfaction items in evaluating performance of clinical unit managers. The significant path coefficient (-0.10) between the strategy of service innovation and PMS patient satisfaction means that the use of PMS patient satisfaction in evaluating clinical unit managers decreases with a health care organization's strategy on enhancing service innovation. On the contrary, the use of research items in an organization's PMS (research) for evaluating clinical unit managers is solely a result of decentralization (path coefficient $= -0.13$), as the path coefficient from the strategy of service innovation to PMS research is insignificant. This suggests that accountability within clinical units is examined through the management of complication and infection rates as well as research output of the clinical unit, evaluation criteria which clinical unit managers perceived to have been used to a greater extent (path coefficient $= -0.13$) as compared to the patient satisfaction criteria (path coefficient $= -0.08$). This can be attributed to the fact that these research items consist of hard, objective data directly related to the operations of the clinical units while patient satisfaction items are based on patients' subjective assessment of the quality, outcome, and coordination of patient-care. Moreover, the clinical units' activities may not be directly affecting patient care, whereas research output is definitely an outcome of the clinical units' efforts.

The path coefficient of the strategy of service innovation to OPE patient satisfaction is -0.36, which suggests that organizational performance on patient satisfaction decreases when the strategic focus of the health care organization is on service innovation. However, unlike the path between the strategy of service innovation and PMS patient satisfaction, it is informative to consider the clinical unit, as depicted by its autonomy structure and PMS, as an intervening variable between the strategy of service innovation and OPE patient satisfaction as the correlation coefficient between the strategy of service innovation and OPE patient satisfaction is -0.279 and clinical unit intervention accounts for a 20 percent increase in OPE patient satisfaction.

Thus, one can conclude that for health care organizations with a strategic focus on service innovation, their use of patient satisfaction to evaluate clinical unit manager's performance decreases and their organizational performance on patient satisfaction deteriorates. A plausible explanation for the findings is that health care administrators and medical professionals will pursue service innovation in areas that they believe are beneficial to their patients. They are also more confident about their professional assessment on the quality and effectiveness of health care services provided than patients' subjective judgment on the quality and effectiveness of health services received, thereby resulting in decreased use of patient satisfaction in performance evaluation and management. Furthermore, as health care organizations pursue service innovation, there are continuous developments and changes in the organization. Some of these organizational changes could have created an atmosphere of uncertainty and caused inconvenience to the patients, who would be more likely to assess the health services and care they received as unsatisfactory. This partially explains why a health care organization's strategic choice on service innovation could have an adverse impact on its performance in patient satisfaction.

Decentralization and autonomy structure, on the other hand, have three paths to consider, all of which have negative path coefficients. The negative paths to PMS patient satisfaction and research indicate that the use of these items in PMS decreases as more decision-making authority is delegated to the clinical unit managers. In addition, autonomy structure has a negative impact upon OPE patient satisfaction. This may be attributed to the fact that as clinical unit managers are given greater decision-making authority, they are held accountable for multiple aspects of their unit's performance. Thus, the use of patient satisfaction and research items in the PMS of clinical units decreases with higher degree of structural autonomy.

With increasing responsibility, clinical unit managers may not have been able to commit the efforts and time to provide the quality of health services that patients demanded. Consequently, the organization's performance on patient satisfaction suffers as well.

DISCUSSION AND CONCLUSIONS

This study has developed a model which allows us to consider the use and usefulness of the balanced scorecard in Canadian health care organizations. Of particular interest a priori are the potential roles of the balanced scorecard as a strategic learning tool (Kaplan & Norton, 1996c), as a communication and goal alignment tool (Kaplan & Norton, 2001a, 2001b), and as a tool for designing and implementing the PMS.

The results of this study suggest that balanced scorecard is capable of serving all of these roles, but not all performance perspectives in the balanced scorecard are of equal importance in these roles to our sample of health care organizations. Evidently, financial performance, system integration, and clinical utilization are not of significant value. This is contrary to the common belief that the government is keen on holding health care organizations accountable for the financial support provided, reducing the administrative burden, integrating the provision of health services, and implementing standard protocol for patient care in Canada's health system. Unsurprisingly, patient satisfaction is the primary measure in the PMS and organizational performance for this sample of health care organizations.

A direct link from a strategy of service innovation to the PMS shows the informational/strategic monitoring role of the balanced scorecard. In this study, increasing innovation in patient services signals an emphasis on the use of quality and outcome of care as well as coordination among caregivers for assessing patient satisfaction, one of the performance perspectives identified in the balanced scorecard. Thus, the balanced scorecard, as a health care organization's PMS, can be considered as a source of information capable of monitoring strategic impact (Simons, 1995). The role of the balanced scorecard as a goal alignment and communication tool as well as a designing tool for PMS is established through the linkage from autonomy structure to the PMS (Abernethy & Lillis, 2001). As service innovations are adopted, the use of the measures in the PMS as established is de-emphasized signaling that the design of the PMS must change to include different measures that are more appropriate to measuring and monitoring the strategic innovations adopted. As a result of changing the PMS, a more appropriate measuring and monitoring tool would emerge complementing the design/redesign cycle. Nonetheless, for this sample of health care organizations, there is a greater emphasis on the use of research items (path coefficient $= -0.13$) than patient satisfaction items (path coefficient $= -0.08$) in the PMS. This is because research is imperative to service innovation on patient care (driver and leading indicator), which can affect patient care and satisfaction in the future (outcome and lagging indicator).

Our findings also support the argument that a strategy of service innovation is a determinant of autonomy structure. As a result of delegation of authority to the clinical unit managers, the requirement for accountability increases. Clinical unit managers in this sample marginally de-emphasize measures in the PMS as established by the balanced scorecard incorporated in the public accountability framework. By de-emphasizing coordination and unit-based care (path coefficient between PMS patient satisfaction and OPE patient satisfaction $= -0.92$) and emphasizing complication and infection rates as well as research output (path coefficient between PMS research/teaching and OPE patient satisfaction $= 0.57$), clinical unit managers steward their units toward research and increased patient satisfaction in the long term while overlooking the immediate impact of their units' action on patient satisfaction. The delegation of authority and decision making to clinical unit managers to pursue the organization's strategy with a long-term perspective, i.e., preference of PMS research (driver and leading indicator) over PMS patient satisfaction (outcome and lagging indicator), has positive impact upon organizational performance in the long run. Nonetheless, the misalignment between the PMS and organizational performance criteria in the short term should not be considered as detrimental to the mission of health care organizations.

Strategy of service innovation, however, has a significant adverse impact on organizational outcome as measured by patient satisfaction. One could speculate that patients have expectations upon entering a hospital for care and that change or innovation introduces uncertainty, which could have a negative impact on their satisfaction with their stay at or visit to the hospital. Another contributing factor is likely the turbulence encountered in the health care sector. For a number of years, the rising costs of health care have been greeted with funding reductions. Consequently, health care organizations have had to undergo severe cost-cutting measures, which usually means a reduction of or change in patient services that would lead to reduced patient satisfaction. It would seem plausible to suggest that, in addition to pursuing a strategy of service innovation, there are a myriad of events or circumstances contributing to patient satisfaction as captured in this data set.

Our results also allow an assessment of the "fit" of the structural arrangements within the model. Organizational performance is enhanced when structure and the PMS align or complement each other (Gresov, 1989). As noted earlier, the structural and PMS choices made by clinical unit managers do enhance performance in this model, thereby indicating that there is some measure of "fit" in the system. However, it is revealing that the results suggest a de-emphasis by clinical unit managers of the PMS, which reflects the balanced scorecard incorporated in the public accountability framework. Despite the usefulness of the balanced scorecard measures in the public accountability role, clinical unit managers revert to other activities associated with their medical tasks, such as research, which are not captured by the balanced scorecard measures of the *Hospital Report*. This would suggest that balanced scorecard measures in the *Hospital Report* are not particularly useful for accountability of clinical processes and medical services within a health care organization. Interviews with two hospital CEOs supported this conjecture indicating that medical staff are accountable through other medical-centric accountability systems. Clearly, it is questionable whether the PMS could be designed to capture clinical activities that reflect the strategic choices of top management.

The usual limitations for path analysis with cross-sectional data apply to this study. That is, this study does not provide evidence of causal relationships. However, we can say that the result is consistent with the explanations offered. In addition, this study was conducted in the Canadian health care sector. Consequently, the results may not be applicable to other industry sectors or health care organizations in a different political climate.

Despite limitations on the use of cross-sectional data from questionnaire survey for path analysis, this study does offer some general insights into the strategy, structure, and performance relationships in the Canadian health care setting. Service innovation choices of top management do impact the delegation of authority, the structure of the PMS, and organizational outcomes. It seems reasonably clear that clinical unit activities are not being captured by the balanced scorecard items used to provide public accountability and insight into hospital performance. This study highlights the importance of PMS design on organizational outcome. Providing a PMS that captures the activities of a clinical unit and links these activities to performance would likely enhance both the informational and accountability roles of the PMS. Further research into these linkages and associations could provide substantial insight into the design and implementation of PMS in the Canadian health care sector.

NOTES

1. Canadian Institute of Health Information. *National Health Expenditure Trends, 1975–2006*, available at http://secure.cihi.ca/cihiweb/dispPage.jsp?cw_page = PG_592_E&cw_topic = 592&cw_rel = AR_31_E, accessed on 7 January 2008.
2. See note 1.
3. Ontario Hospital Association http://www.oha.com/Client/OHA/OHA_LP4W_LND_WebStation.nsf/page/Hospital+Report, accessed on 9 January 2008.
4. The scorecard in the 2007 *Hospital Report Acute Care* includes an additional perspective on patient satisfaction pediatric care with eight indicators.
5. Vancouver Coastal Health, available at http://www.vch.ca/accountability/scorecard.htm, accessed on 9 January 2008.
6. Chinook Health Region, available at http://www.chr.ab.ca/bins/doc.asp?rdc_id = 5346, accessed on 9 January 2008.
7. David Thompson Health Region, available at http://www.dthr.ab.ca/resources/documents/reports/ThreeYearHealthPlan.pdf, accessed on 9 January 2008.
8. Poisson, B. (2007). *Executing Strategy through the Balanced Scorecard.* Available at http://www1.umn.edu/osci/download/Possin_070110.ppt, accessed on January 9, 2008.
9. University Health Network, Toronto East General Hospital, St. Michael's Hospital, and The Hospital for Sick Children in Toronto are among the 22 acute care hospitals (Yap et al., 2005) in Ontario which have implemented balanced scorecards as part of their accountability framework and strategic management system.
10. For this study, clinical unit managers and medical program directors are simply referenced as clinical unit managers.
11. For the five performance perspectives and measures in an organization's performance management system and performance outcome, the path coefficients from decentralization and/or structural autonomy to financial performance, system integration, and clinical utilization are insignificant. The path coefficients from decentralization and/or structural autonomy to the use of patient satisfaction in an organization's performance management system and performance evaluation, respectively, are significant, whereas only the path coefficient from decentralization and/or structural autonomy to the use of research/teaching in an organization's performance management is significant.

ACKNOWLEDGMENT

The authors would like to acknowledge the funding support from Arts Research Board, McMaster University.

REFERENCES

Abernethy, M. A., & Lillis, A. M. (2001). Interdependencies in organization design: A test in hospitals. *Journal of Management Accounting Research*, *13*, 107–129.

Abernethy, M. A., & Stoelwinder, J. U. (1991). Budget use, task uncertainty, system goal orientation and subunit performance: A test of the "fit" hypothesis in not-for-profit hospitals. *Accounting, Organizations and Society*, *16*(2), 105–120.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *15*(May), 411–423.

Baker, G. R., & Pink, G. H. (1995). A balanced scorecard for Canadian hospitals. *Healthcare Management FORUM*, *8*(4), 7–13.

Bouwens, J., & Abernethy, M. A. (2000). The consequences of customization on management accounting system design. *Accounting, Organizations and Society*, *25*(3), 221–241.

Canada Health Act (1984). *Government of Canada*. Ottawa, ON: Government of Canada.

Canadian Healthcare Association (2003). *Guide to Canadian healthcare facilities*. Ottawa, ON: Canadian Healthcare Association.

Chan, Y. L., & Ho, K. S. (2000). Performance measurement and the use of balanced scorecard in Canadian hospitals. *Advances in Management Accounting*, *9*, 145–169.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

Curtright, J. W., Stolp-Smith, S. C., & Edell, E. S. (2000). Strategic performance management: Development of a performance measurement system at the Mayo Clinic. *Journal of Healthcare Management*, *45*(1), 58–68.

Dixon, J. R., Nanni, A. J., & Vollman, T. E. (1990). *The new performance challenge: Measuring operations for world-class competition*. Homewood, IL: Dow Jones-Irwin.

Freidson, E. (1975). *Doctoring together: A study of professional social control*. Chicago, IL: University of Chicago Press.

Gresov, C. (1989). Exploring fit and misfit with multiple contingencies. *Administrative Science Quarterly*, *34*(3), 431–435.

Gumbus, A., Bellhouse, D., & Lyons, B. (2003). A three year journal to organizational and financial health using balanced scorecard: A case study at a Yale New Haven System hospital. *Journal of Business and Economic Studies*, *9*(2), 54–64.

Hambrick, D. C. (1983). Some tests of the effectiveness and functional attributes of Miles and Snow's strategic types. *Academy of Management Journal*, *26*(1), 5–26.

Harber, B. W. (1998). The balanced scorecard – Solution at Peel Memorial Hospital. *Hospital Quarterly*, *1*(3), 59–62.

Hospital Report Acute Care (2007). *Hospital report research collaborative*. Ottawa, ON: Canadian Institute of Health Information.

Hospital Report '98 – A System-Wide Review of Ontario Hospitals (1998). *Ontario Hospital Association*. Toronto, ON: Ontario Hospital Association.

Hospital Report '99 – A Balanced Scorecard for Ontario Acute Care Hospitals (1999). *Ontario Hospital Association*. Toronto, ON: Ontario Hospital Association.

Inkson, J. H. K., Pugh, D. S., & Hickson, D. J. (1970). Organizational context and structures: An abbreviated replication. *Administrative Science Quarterly*, *15*, 318–329.

Jones, M. L. H., & Filip, S. J. (2000). Implementation and outcomes of a balanced scorecard model in women's services in an academic health care institution. *Quality Management in Health Care*, *8*(4), 40–51.

Joreskog, K. G., & Sorbom, D. (1993). *Lisrel 8: Structural equation modeling with the SIMPLIS command language*. Hillside, NJ: Lawrence Earlbaum.

Kaplan, R. S. (2001). Strategic performance measurement and management in non-profit organizations. *Nonprofit Management and Leadership*, *11*(3), 353–370.

Kaplan, R. S., & Atkinson, A. A. (1998). *Advanced management accounting*. Upper Saddle Rive, NJ: Prentice Hall.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard-measures that drive performance. *Harvard Business Review*, *70*(1), 71–79.

Kaplan, R. S., & Norton, D. P. (1996a). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, *74*(1), 75–85.

Kaplan, R. S., & Norton, D. P. (1996b). *The balanced scorecard – Translating strategy into action*. Boston, MA: Harvard Business School Press.

Kaplan, R. S., & Norton, D. P. (1996c). Strategic learning and the balanced scorecard. *Strategy and Leadership*, *24*(5), 19–24.

Kaplan, R. S., & Norton, D. P. (2001a). Transforming the balanced scorecard from performance measurement to strategic management: Part I. *Accounting Horizon*, *15*(1), 87–104.

Kaplan, R. S., & Norton, D. P. (2001b). Transforming the balanced scorecard from performance measurement to strategic management: Part II. *Accounting Horizon*, *15*(2), 147–160.

Kaplan, R. S., & Norton, D. P. (2001c). *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*. Boston, MA: Harvard Business School Press.

Kershaw, R., & Kershaw, S. (2001). Developing a balanced scorecard to implement strategy at St. Elsewhere Hospital. *Management Accounting Quarterly*, *2*(1), 28–35.

Kren, L. (1997). *The role of accounting information in organizational control*. New York, NY: Wiley.

Lei, D., Hitt, M. A., & Goldhar, J. D. (1996). Advanced manufacturing technology: Organizational design and strategic flexibility. *Organizational Studies*, *17*(3), 501–524.

Meliones, J. (2000). Saving money, saving lives. *Harvard Business Review*, *78*(6), 57–65.

Merchant, K. A. (1998). *Modern management control systems: Text and cases*. Upper Saddle River, NJ: Prentice Hall.

Miles, R. E., & Snow, C. C. (2003). *Organizational strategy, structure and processes*. Palo Alto, CA: Stanford University Press.

Paul, J., Blackstein-Hirsch, P., Brown, A., Magistretti, A., Gilbart, E., Huynh, T., & Veillard, J. (2006). *Aligning performance measurement with corporate strategy: A toolkit for Ontario's hospitals*. Toronto, ON: Government of Ontario, The Ministry of Health and Long-Term Care.

Pink, G. H., McKillop, I., Schraa, E. G., Pregra, C., Montgomery, C., & Baker, G. R. (2001). Creating a balanced scorecard for a hospital system. *Journal of Health Care Finance*, *27*(3), 1–20.

Porter, M. E., & Teisberg, E. O. (2004). Redefining competition in health care. *Harvard Business Review*, *82*(6), 64–76.

Public Hospital Act (1990). *Government of Ontario*. Toronto, ON: Government of Ontario.

Pugh, D. S., Hickson, D. J., Hinings, C. R., & Turner, C. (1968). Dimensions of organization structure. *Administrative Science Quarterly*, *13*, 33–47.

Sahney, V. K. (1998). Balanced scorecard as a framework for driving performance in managed care organizations. *Managed Care Quarterly*, *6*(2), 1–8.

Scott, T. W., & Tiessen, P. (1999). Performance measurement and managerial teams. *Accounting, Organizations and Society*, *24*(3), 263–285.

Shortell, S. M., Anderson, D. A., Erickson, K. M., & Mitchell, J. B. (1996). *Remaking health care in America: Building organized delivery systems*. San Francisco, CA: Jossey Bass.

Simons, R. (1995). *Levers of control*. Boston, MA: Harvard Business School Press.
Slater, S. F., & Olson, E. M. (2000). Strategy type and performance: The influence of sales force management. *Strategic Management Journal*, *21*(8), 813–829.
Yap, C., Siu, E., Baker, G. R., Brown, A. D., & Lowi-Young, M. P. (2005). A comparison of systemwide and hospital-specific performance measurement tools. *Journal of Healthcare Management*, *50*(4), 251–263.
Zelman, W. N., Pink, G. H., & Matthias, C. B. (2003). Use of balanced scorecard in health care. *Journal of Healthcare Finance*, *29*(4), 1–16.

APPENDIX. MEASUREMENT OF VARIABLES PER SURVEY ITEMS

(1) Strategy (Service Innovation)

Given the following descriptions of two hospitals, please circle on the scale where you would place your hospital ranging from 1 to 7, where 1 represents a Hospital A-type and 7 represents a Hospital B-type, with respect to your organization's patient-care/medical service.

Hospital A It offers a relatively stable set of services. Generally, it is not at the forefront of new services or developments in its field. It tends to ignore changes that have no direct impact on current areas of operation and it concentrates instead on doing the best possible job in its existing arena.

Hospital B It makes frequent changes in its set of services. It consistently attempts to pioneer in new areas of service activity, even if not all of these efforts ultimately prove to be highly successful. It responds rapidly to early signals of opportunities or needs.

	Hospital A						Hospital B
Patient-care/Medical services	1	2	3	4	5	6	7

(2) Structural Autonomy Survey Items

The following questions relate to the lack of autonomy of your particular clinical unit/medical program. Please indicate the level at which each of the policy decision items in the following list is determined in your organization using the score indicated as below:

1 Supervisor,
2 Clinical Unit Manager/Medical Program Director,
3 V.P., Medical Program and Professional Services/ V.P., Administration/Business/Finance/Operations,
4 C.E.O./President, Chief of Staff, and
5 Outside the organization, e.g., the Ministry of Health, the Board of Directors.

	Policy Decisions	Level
(a)	Supervisory establishment.	—
(b)	Appointment of supervisory staff from outside the organization.	—
(c)	Promotion of supervisory staff.	—
(d)	Salaries of supervisory staff.	—
(e)	To spend un-budgeted or unallocated money on capital items.	—
(f)	To spend un-budgeted or unallocated money on expense/revenue items.	—
(g)	What type, or what brand, of new equipment is to be.	—
(h)	To determine a new medical service.	—
(i)	To determine service territories covered.	—
(j)	The extent and type of service to be aimed for.	—
(k)	What shall be costed.	—
(l)	What shall be inspected.	—
(m)	What operations shall be work studied.	—
(n)	Dismissal of a medical supervisor.	—
(o)	Training methods to be used.	—
(p)	Buying procedures.	—
(q)	Which suppliers of medical supplies are to be used.	—
(r)	What and how many welfare facilities (e.g., gymnasium and day care center for employees) are to be provided.	—
(s)	The price of the output.	—
(t)	To alter responsibilities/areas of work of specialist departments.	—
(u)	To alter responsibilities/areas of work of departments.	—
(v)	To create new clinical units/medical programs.	—
(w)	To create new job.	—

(3) Performance Management System (Balanced Scorecard) Survey Items

Please indicate the extent to which the following items are used to measure your performance as a clinical unit manager/medical program director on a scale ranging from 1 (to a little extent) to 7 (to a great extent).

	Performance Measures							
(1)	Patient satisfaction with quality of unit-based care, physician care and nursing care.	1	2	3	4	5	6	7
(2)	Patient satisfaction with outcome.	1	2	3	4	5	6	7
(3)	Patient satisfaction with coordination of care among caregivers in my organization (the hospital).	1	2	3	4	5	6	7
(4)	Complications and infection rate in my unit.	1	2	3	4	5	6	7
(5)	Length of stay of patients as compared to clinical units at hospitals of similar size and functions.	1	2	3	4	5	6	7
(6)	Throughput targets.	1	2	3	4	5	6	7
(7)	Budget performance.	1	2	3	4	5	6	7
(8)	Patient case cost as compared to clinical units at hospitals of similar size and functions.	1	2	3	4	5	6	7
(9)	Patient care hours to total staff hours in my unit.	1	2	3	4	5	6	7
(10)	Intensity of use of clinical information in patient care in my unit.	1	2	3	4	5	6	7
(11)	Adherence to standard procedures and clinical pathways.	1	2	3	4	5	6	7
(12)	Coordination of care and cooperation with other units in hospitals.	1	2	3	4	5	6	7
(13)	Coordination and continuity of care of patients in community.	1	2	3	4	5	6	7
(14)	Innovation in training programs and employee practices in my unit.	1	2	3	4	5	6	7
(15)	Harmony of my unit.	1	2	3	4	5	6	7
(16)	Research output of my unit.	1	2	3	4	5	6	7

(4) Organizational Performance Survey Items

Please assess your hospital's organizational performance on a scale ranging from 1 *(below average)* to 7 *(above average)* on the following dimensions:

	Organizational Performance and Outcomes							
(1)	Patient satisfaction with overall quality of care.	1	2	3	4	5	6	7
(2)	Patient satisfaction with outcome.	1	2	3	4	5	6	7
(3)	Patient satisfaction with coordination of care among caregivers in your organization (the hospital).	1	2	3	4	5	6	7
(4)	Complications and infection rate in your organization (the hospital).	1	2	3	4	5	6	7
(5)	Length of stay of patients as compared to hospitals of similar size and functions.	1	2	3	4	5	6	7
(6)	Ability to win resources.	1	2	3	4	5	6	7
(7)	Total margin (grants less costs).	1	2	3	4	5	6	7
(8)	Patient case cost as compared to hospitals of similar size and functions.	1	2	3	4	5	6	7
(9)	Patient care hours to total staff hours.	1	2	3	4	5	6	7
(10)	Intensity of use of clinical information in patient care.	1	2	3	4	5	6	7
(11)	Adherence to standard procedures and clinical pathways.	1	2	3	4	5	6	7
(12)	Coordination of care of clinical units in your organization (the hospital).	1	2	3	4	5	6	7
(13)	Coordination and continuity of care of patients in the community.	1	2	3	4	5	6	7
(14)	Innovation in training programs and employee practices.	1	2	3	4	5	6	7
(15)	Reputation of medical programs.	1	2	3	4	5	6	7
(16)	Undergraduate and graduate medical/health professional training.	1	2	3	4	5	6	7
(17)	Research.	1	2	3	4	5	6	7

SYSTEM INTEGRATION AND THE BALANCED SCORECARD: AN EMPIRICAL STUDY OF SYSTEM INTEGRATION TO FACILITATE THE BALANCED SCORECARD IN THE HEALTH CARE ORGANIZATIONS

Bea Chiang

ABSTRACT

The Balanced Scorecard (BSC) emphasizes on the of information system to track a limited number of balanced metrics (measures and indicators) that are closely aligned with organization's goals. This study investigates how system integration in different forms is related to the success of using the BSC for performance measurement. The use of a BSC in performance evaluation is considered in five contexts: determining cost, measuring efficiency, ensuring quality and customer satisfaction measure, promoting continuous innovation and monitoring contract negotiation. The findings indicate that system integration defined in the study positively relates to the success of using the BSC in all five decision perspectives. The findings conclude that hospitals need a streamlined, information integration across

Advances in Management Accounting, Volume 17, 181–201
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017008

the continuum of care to better assess the operation results, in both organizational and technical perspectives.

INTRODUCTION

Increased payer power, growing health care competition and constraining regulations creates tremendous pressure for health care to balance complicated trade-offs among cost, quality, access and consumer choices (Shortell, Gillies, Anderson, Erickson, & Mitchell, 2000; Inamdar, Kaplan, & Bower, 2002). Health care providers have begun to use the Balanced Scorecard (BSC) as a new business tool to help them take a more strategic approach to measure business performance. The BSC is a multidimensional framework for implementing and managing strategy at all levels of business by linking objectives and measures to an organization's strategy (Kaplan & Norton, 1993, 1996). It is essentially a customized performance measurement system that goes beyond traditional accounting and is based upon organizational strategy.

Much literature relates to the BSC focus primarily in two areas: (1) introduction of the BSC concepts and its application in health care organizations (Baker & Pink, 1995; Beauchamp, 1999; Castaneda-Mendez, Mangan, & Lavery, 1998; Chow, Ganulin, Teknika, Haddad, & Williamson, 1998; Gordon, Chapman, Kunov, Dolan, & Carter, 1998; Shortell et al., 2000; Oliveira, 2001; Malina & Selto, 2001; Sugarman & Watkins, 2004; Wicks & St. Clair, 2007) and (2) case study or field experience of implementing the BSC in a particular health care setting (Harber, 1998; MacDonald, 1998; Peters & Ryan, 1999; Curtright, Stolp-Smith, & Edell, 2000; Mathias, 2001; Inamdar et al., 2002; Chong, Verma, Mythily, Poon, & McGorry, 2008). No prior literature empirically tests the organizational information system as one of factors that affects the use of BSC in the health care organization. The objective of this study attempts to address this research gap. Specifically, this study focuses on examining how the system integration in different forms facilitates the use of a BSC in the health care organization.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Balanced Scorecard (BSC)

The BSC is both a performance measurement framework and a management methodology. It was developed by Robert Kaplan and David Norton after an extensive research project in 1990. The scorecard typically measures an organization's performance across four dimensions: financial perspective, customer perspective, internal perspective and learning and growth perspective. The BSC enables organizations to track short-term financial and operating results while monitoring progress in building the capabilities and acquiring the intangible assets needed for future growth and success (Kaplan & Norton, 1993). In translating the vision, specific objectives and measures are developed which are in alignment with the organization's vision, mission and strategy. In communication, management strategy is communicated throughout the organization via a concise easily interpreted, graphic scorecard that enables communicating and educating, goal setting and linking organizational rewards to performance measures. Kaplan and Norton (1996) argue that the BSC is a strategic planning device and requires communication to (1) provide strategic guidance to divisional managers and (2) describe links among lagging and leading measures of financial and nonfinancial performance.

In health care, scorecards appear even more diverse than in the business sector (Voelker, Rakich, & French, 2001; Wicks & St. Clair, 2007; Chong et al., 2008). Health care organizations face different stakeholder groups, which include patients and their families, employers, health plans, physicians, employee administrators, shareholders, communities and regulators. Applying the BSC to health care presents some unique challenges because health care faces a group of stakeholders that often has a higher order BSC perspective than "financial." Another challenge is that health care is an information-intensive business, generating huge volumes of data from hospital, clinics, laboratories and other related parties. However, because of the nature of their business, health care organizations have consisted of independent and autonomous units with little clinical incentive to share information.

It is important to realize that the business processes of operating units in the health care are highly interrelated and their information systems should be able to capture the interactions among organizations and operating units to allow the organizations to measure and evaluate the performance. Traditionally, health care organizations are using a "report card" to measure and communicate performance. The report card focuses on two dimensions: clinical outcomes and cost measures. They do not provide enough information for assessing all of the critical success factors of the organization (Voelker et al., 2001). In addition, the single doctor–patient relationship is being replaced by one in which the patient is managed by a team of health care professionals each specializing in one aspect of care. Such shared care depends critically on the ability to share information. Specifically, the health care industry generates a large quantity of clinical information which needs to be efficiently integrated in order to maintain quality of care (Bojan & Koncar, 2007). Grimson, Grimson, and Hasselbring (2000) argue that "the inability to share information across systems and between care organizations automatically is one of the major impediments to progress toward shared care and cost containment." Likewise, the integrated information systems play a critical role to facilitate an effective implementation of BSC to capture the information and measure the performance of the organization. This study intends to address the following related research question:

Q1. Does system integration relate to the success of implementing BSC?

Clinical Autonomy, System Integration and BSC

Much literature has documented the concept of system integration, and there is no one accepted definition of integration. It is a concept that has many different meanings to different people (Zachman, 1987). A comprehensive review of the literature in the areas of information systems, operations and production management, technology, organizations and systems was undertaken to identify a range of possible definitions and models of integration. Waring and Wainwright (2000) have grouped the definitions and models into four distinct areas comprising technical, systems, organizational and strategic domains. Boaden (1991) defines integration as "Integrated information supports the integrated organization by providing a shared source of accurate, timely data as the basis for communication and decision." To narrow down system integration to the information system aspect, Grimson et al. (2000) defines information integration into three layers of integration: the business architecture, the application architecture and the technology architecture layers. The business architecture layer defines the integration in the form of organization structure and business processes. The application architecture layer defines the actual integration of the business concepts in terms of enterprise applications. The technology architecture layer refers to the information and communication infrastructure.

Expanding the concept of system integration, O'Sullivan (1992) proposes a specific model regarding system integration. The model describes system integration at two levels: social and technical. Social integration involves the integration of people and decision-making process. Technical integration is more concerned with the integration of technical subsystems including equipment techniques and procedures.

Overall, we may expect that organization structure and technical integration may provide a context to facilitate the implementation of BSC. In this study, the organization structure is viewed in terms of level of clinical unit autonomy (described in the following section). Technical integration is viewed in terms of integration of system applications (software, operation systems, etc.) and data accessibility among operating units and affiliated organizations.

BSC and Autonomy of Clinical Units

The literature in performance measure and management control has focused on how to design a performance measurement system to provide a set of measures that directs subordinates' attention to the organization's strategy (Dixon, Nanni, & Vollman, 1990; Kaplan & Norton, 2001; Lipe & Salterio, 2000, 2002) and at the same time considers the demand of autonomy and accountability of subunits of the organization (Lawrence & Lorsch, 1967; Vancil & Buddrus, 1979; Flamholtz, Das, & Tsui, 1985; Abernethy & Lillis, 2001). Autonomy is defined in many ways in the literature to mean possessing self-direction (Lewis, 1975), being left on one's own to work (Johnson & Tingey, 1976), demonstrating behavior not controlled by an external agent (Katz, 1968) and showing independence (Mundinger, 1980). Dworkin (1988) defines autonomy as the capacity to reflect upon one's motivational structure and to make changes in that structure. Some others define autonomy in the medical setting as the exercise of independent judgment and freedom to make discretionary decisions, actions and plans according to one's scope of practice, which requires practitioners be self-directed, intellectually flexible, responsible and accountable for their own actions (Batey & Lewis, 1982; McKay, 1983; Dempster, 1994; Cullen, 2000; Cajulis & Fitzpatrick, 2007).

Managers of clinical units in hospitals are generally physicians who have traditionally had considerable autonomy over clinical processes and patient care outcomes. This autonomy gives physicians complete control over operating tasks without responsibility for the financial consequence of clinical decisions (Young & Saltman, 1985; Weiner, Maxwell, Sapolsky, Dunn, & Hsiao, 1987; Burns, Anderson, & Shortell, 1993). Clinicians will strongly resist imposed bureaucratic rules and procedures that threaten their autonomy in medical treatment decisions. (Freidson, 1975; Cajulis & Fitzpatrick, 2007). Much literature discusses autonomy and accountability in health care. Batey and Lewis (1982) suggest that level of autonomy held by nursing is a function of the degree of responsibility assumed by the nursing practitioners and assigned by supervisors. Cajulis and Fitzpatrick (2007) founds that the level of autonomy of nurse practitioners has positive relation to their responsibility. The exercise of autonomy and assignment of responsibility requires that we clarify both the boundaries and the scope of practice which needs a performance measurement system to insure that freedom to act be consistent with both responsibility and authority (Batey & Lewis, 1982).

Maas and Jacox (1977) argue that "accountability for behavior is a corollary of autonomy. Accountability implies responsibility and answerability to authority for one's actions." It can be seen that, if individuals are prepared to act autonomously, they must accept that they are answerable for their actions. It is expected that increased autonomy will be accompanied by increased accountability and reflected in greater importance attached to performance measurement systems. Abernethy and Lillis (2001) indicate that the performance measurement system creates accountability for outcomes while at the same time enables professionals to maintain their desired autonomy. Moreover, they find that there is a significant positive relationship between the level of autonomy and the extent of use of performance measurement criteria.

As clinical units operate more independently, the resulting increased accountability helps the organization to define and develop specific criteria for performance measurement to hold the operating units accountable. Performance measure systems such as BSC provide a comprehensive accountability system designed to capture the activities performed in the operating units. If social context for implementing BSC is viewed in terms of the level of autonomy, it is argued that the use of the BSC is more likely to be positively affected by the level of autonomy. This expectation is expressed as follows:

H1. The degree of success of using a BSC is positively related to the level of autonomy of operating units.

BSC and System Integration – Technical Perspective

Health care organizations need to reduce fragmented care delivery by bringing together the information system of provider organizations across the continuum of care into one cooperative structure. It is an essential step to improve quality and reduce the cost of care delivery through tighter integration among providers. A typical example of technical integration could be the integrated subsystems that support a wide range of activities of a hospital such as the patient logistics process, scheduling processes, ordering management, medications, discharge and appointment scheduling. Data integration is another perspective of technical integration. It is an architecture with a central databank containing patients' administrative data, medical data and available resources and their utilization. The system can be divided in a number of subsystems that each support a more or less logical set of functionality and consist of a cluster of co-operating programs. Information on performance in each co-operation program can then be captured and measured.

Several studies address many aspects of technical integration. Krol, Reich, and Dupont (2005) specifically address system integration issues in terms of integrating different application software packages built on different computer platforms. Paré and Sicotte (2001) and Jaana, Ward, Paré, and Wakefield (2005), both studies argue that clinical information technology can be assessed within three clinical domains: (1) patient management, (2) patient care activities and (3) clinical support activities. The patient management domain includes admission/discharge/transfer applications and covers issues related to medical records. The patient care domain includes computer-based applications and technologies supporting physicians, nurses and emergency department and the operating suite. And the clinical support activities domain includes clinical information technology applications and technologies present in laboratories, radiology and pharmacy. Hasselbring (2000) focuses on information system integration in terms of business architecture, applications architecture and technology architecture among organizations. Thomas, Robinson, Waring, Wainwright, and Maguire (1995) provide a "hospital information support system" to describe an integration vision for a large acute hospital. The support involves integrating departmental and support systems such as pathology, radiology and pharmacy to be capable of transferring requests (e.g. for diagnostic tests) and associated results. At the same time, the daily transactions recorded on these systems should be integrated with a case mix management system. The case mix management system would be capable of relating information on the activity of the hospital not only to managers responsible for formulating contracts in the health care market, but also possibly to medical staff requiring information for medical audit. Consequently, systems could not only be integrated, but could also capture the information required for managerial and clinical decision-making.

Integration among subsystems will assist in collecting necessary information for performance evaluation in the key areas. Grimson et al. (2000) commented that the inability to share information across systems and between care units is just one of the major impediments in implementing the BSC. Many health care organizations do not get good results from the BSC implementation, not because they lack balanced perspectives, but simply because they do not develop appropriate information systems to support the scorecard (Jensen & Sage, 2000). Wicks and St. Clair (2007) make similar comments that an integrated information system plays a critical role in the communication of performance evaluation. The findings of Hospital Peer Review Report (HPRA, 2000a) also suggest that the process of design and implementation in health care scorecards requires exquisite attention to communications with multiple stakeholders and integration among information systems. This view of prior research leads to the following hypotheses:

H2a. The degree of success of using a BSC in performance measurement is positively related to technical integration among software or operating systems within the operating unit.

H2b. The degree of success of using a BSC in performance measurement is positively related to technical integration in terms of information accessibility and sharing among the operation units.

H2c. The degree of success using a BSC in performance measurement is positively related to technical integration in terms of information accessibility and sharing among the affiliated health organizations.

Specifically, the use of the BSC is measured in the following five areas: determining cost and measuring efficiency (internal business perspective), assessing quality and customer satisfaction (customer perspective), promoting continuous innovation (innovation and learning perspective) and monitoring contract negotiation (financial perspective).

The relationship of system integration and the success of using the BSC are summarized in Fig. 1.

RESEARCH METHODOLOGY

Data Description

Six hundred hospitals were selected from the HCIA Directory of Health Care Professionals (2000) for the mailing of surveys on the basis of inpatient

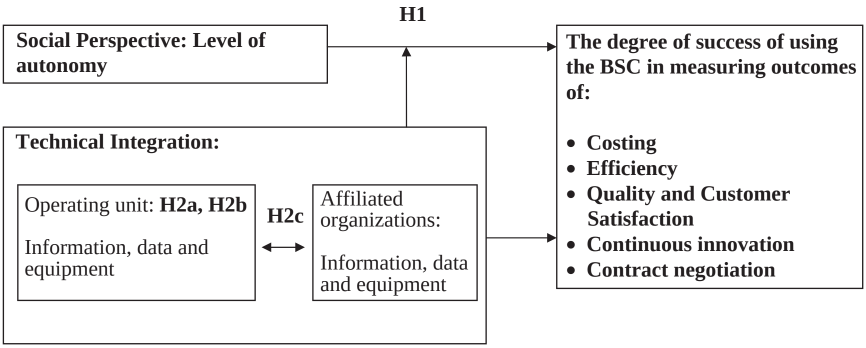


Fig. 1. System Integration and the Success of Using BSC.

bed size (above 300 licensed beds). The questionnaire allows respondents to indicate their feedback based on a scale ranging from 1 (strongly agree) to 7 (strongly disagree) (see appendix). One of the major concerns over the self-rating of prior literature has been the problem of common-rater bias (Abernethy & Lillis, 2001). To reduce the concerns of common-rater bias, questionnaires were sent to two subjects of each hospital – accounting director and CFO.

The questionnaire consisted of five groups of questions. The first group of questions covered background information such as the number of inpatient beds, annual operating revenue location (state) and the respondent's position. Information about system integration in the social dimension was obtained by asking respondents to rate the level of autonomy of clinical units of the hospital. The measure of the level of autonomy follows the study of Abernethy and Lillis (2001) in that autonomy is measured by asking respondents if clinical units are responsible for costs (COST), throughput (T), both cost and revenue (COST&R), both cost and throughput (COST&T) and if the clinical units are treated as a business unit (BUSUNIT).

This article follows the study of Chow et al. (1998) to construct an instrument to measure four perspectives of performance based on BSC: customer, internal business, innovation and learning and financial. Customer perspective includes quality of service and customer satisfaction. The survey conducted by Chow et al. (1998) shows that the hospital administrators consider patient satisfaction to be of primary importance. This emphasis on patient satisfaction encompasses quality of medical care, prompt service, satisfaction with the medical staff and food served.

The study also indicates that one of the administrators' goals is to control cost and gauge the effectiveness and efficiency of using resources. Following the same classification, this article uses identifying costs and measuring efficiency to assess internal business performance. Promoting continuous innovations is used to measure the innovation and learning perspective. As Chow et al. comment, financial goals of hospitals indicate a need to increase contracts with affiliated organizations, state, federal, insurance, philanthropy and fund-raising agencies. These goals are tied to the measures as dollars generated from new contracts and percentage of contracts relative to competitors. The increase in financial capital from contracts suggests enhanced financial perspective of performance.

For this research, survey instrument items related to technology integration were developed based on the research of O'Sullivan (1992), Hasselbring (2000) and Jaana et al. (2005), who suggest that integration can be assessed in three dimensions: integration among data and software applications, operating units and affiliated health organizations. Data integration refers to medical records, general ledger accounts and billing documents exchanged and shared among authorized users. Two types of information technology integration are vertical integration and horizontal integration. Vertical integration can be viewed as integration of data, software application and operation systems (O'Sullivan, 1992; Hasselbring, 2000; Jaana et al., 2005). Horizontal integration can be measured as inter-business units and inter-organizational system integration which allows information flow among operation units/departments and across affiliated organizations (Hasselbring, 2000). The information system integration focuses on the information sharing and accessibility among operation units/departments' specific systems such as payroll, material management, nursing acuity and operating room information systems. To summarize, the relationship between BSC and integration of data, software and operation systems is tested in H2a; the relationship between BSC and integration of information technology among operation units/departments and affiliated health organizations is tested in H2b and H2c respectively.

Information was obtained by asking respondents to rate the integration of software or operating systems within the operating unit (general ledger, medical records and billing process), denoted as OPR SYS and data accessibility among operating units (OPR DATA) and affiliated organizations (AFFIAD DATA). The last group of questions covers the information regarding the success of using the BSC in determining cost, measuring efficiency, assessing quality and customer satisfaction, promoting continuous innovation and monitoring contract negotiation.

Ninety-five usable questionnaires with all the responses completed were returned for a response rate of 16%. A follow-up letter was sent out to the nonrespondents about two weeks after the initial questionnaire had been sent out. The nonresponse bias was examined by first testing the differences between respondents and nonrespondents based on published data such as inpatient beds and geographical locations. A second test compared the early respondents (returned the questionnaire within two weeks of initial questionnaire being sent out) and late respondents (who had returned the questionnaire after receiving the follow-up letter about one month after the initial questionnaire had been sent out). The comparison of responses based on χ^2 analysis revealed no significant differences ($p<0.05$), suggesting the absence of any obvious nonresponse bias. To address the concerns of common-rater bias (self-rating bias), the responses of accounting director and CFO were matched to test for the differences between these two groups. The comparison indicates that there is no significant difference (χ^2, $p<0.05$) between these two sets of respondents. As a result, the self-rating bias should not be a concern.

RESULTS

A descriptive summary of the survey responses is provided in Table 1. Variable correlations are summarized in Table 2 and indicate that multicollinearity is not a problem in the analysis. Linear regression was run to estimate the relationships among variables. Annual operating revenue, inpatient beds and location were entered as controlled variables to control and adjust the effect of hospital size and geographical locations.

Table 3 provides the results of the regressions using the success of using the various categories of BSC as a dependent variable. The success of using the BSC is viewed in costing, measuring efficiency, assessing quality and customer satisfaction, promoting continuous innovation and monitoring contract negotiation. Regression was run based on these five perspectives of the BSC as dependent variables. Only the level of autonomy measured by clinical unit held responsible for managing throughput, cost and throughput and is treated as business unit shows significant relationships to the success of using the BSC in evaluating the results of contract negotiation ($p<0.1$). The results also show significant effect on the success of using the BSC on promoting continuous innovation when the operating unit is held responsible for cost and revenue ($p<0.1$). Overall, the level of autonomy (measured by COST, T, COST&T, COST&R and BUSUNIT) does not

Table 1. Descriptive Statistics ($n = 95$).

Variables	Mean	Standard Deviation
BSC	3.7566	1.58157
COST	2.8604	1.68550
T	3.4604	1.57215
COST&T	3.2501	1.35511
COST&R	3.0765	1.28999
BUSUNIT	3.2101	1.22662
OPR SYS	4.5170	1.49740
OPR DATA	3.8868	1.40850
AFFIAD DATA	4.5415	1.58229

show as a significant factor that relates to the use of the BSC in measuring different perspectives of performance. These findings do not support the first hypothesis (H1) and conflicts with the findings of Abernethy and Lillis (2001) that the level of clinical units' autonomy is positively related to the extent of using a specific performance measurement system.

When technology system integration was tied to the level of autonomy of clinical units in order to see how the technology integration in different forms may facilitate the information communication and result in the success of using BSC in measuring performance, the results show that it is significant when the software and operating systems are integrated among clinical units. The significant statistics are shown in all interaction terms of OPR SYS and the autonomy terms. When operating systems are integrated among operating units, the clinical units that are only responsible for cost (OPR * COST) are found significantly related to the success of using BSC in determining costs ($p < 0.01$), measuring efficiency ($p < 0.1$), assessing quality and customer satisfaction ($p < 0.1$) and monitoring contract negotiation ($p < 0.1$), but do not show significant results for promoting continuous innovation. Similar results of relationships are also found for the rest of the autonomy levels at the significance levels of $p < 0.05$ and $p < 0.1$. When technology systems are integrated, the relationship of autonomy levels and the success of using BSC in measuring performance is particularly significant for clinical units that are treated as business units. These results are shown in significant interaction terms of system integration and autonomy ($p < 0.05$ and $p < 0.01$).

From the technical perspective, the results indicate that integration among software and only operating systems (OPR SYS) is positively and significantly associated with the use of the BSC in determining costs

Table 2. Correlation of the Independent Variables.

	INTERCEPT	COST	T	COST&T	COST&R	BUSUNIT	OPR SYS	OPR DATA	AFFIADDATA
INTERCEPT	1.0000	0.1915	0.0562	0.0021	0.0151	0.1155	0.0254	0.0078	0.0121
COST		1.0000	0.0462	0.0137	0.0142	0.0015	0.0264	0.0321	0.0128
T			1.0000	0.0125	0.0254	0.0256	0.0111	0.0012	0.0142
COST&T				1.0000	0.0189	0.0122	0.1122	0.0157	0.0364
COST&R					1.0000	0.2051	0.0195	0.0037	0.0257
BUSUNIT						1.0000	0.0459	0.0047	0.0382
OPR SYS							1.0000	0.0235	0.0332
OPR DATA								1.0000	0.1219
AFFIADDATA									1.0000

Table 3. Regression Analysis Results.

Dependent Variables[a]	Costing	Efficiency	Quality and Customer Satisfaction	Continuous Innovation	Contract Negotiation
Independent variables					
INTERCEPT	4.23	5.02	3.89	3.66	4.25
COST	3.01	2.55	1.99	2.87	3.01
T	2.89	2.99	2.01	2.56	3.01*
COST&T	2.35	1.88	2.56	2.33	3.21*
COST&R	1.35	2.55	1.59	3.01*	1.89
BUSUNIT	3.55	3.01	2.01	2.88	3.28*
OPR SYS * COST	4.75***	3.25*	3.23*	2.97	3.24*
OPR SYS * T	3.26*	4.28**	3.28*	3.55*	2.58
OPR SYS * COST&T	3.95*	3.15*	3.33*	3.64*	2.56*
OPR SYS * COST&R	4.05**	3.28*	4.05**	2.57	3.18*
OPR SYS * BUSUNIT	4.55**	4.38**	5.08***	4.25**	5.21***
OPR DATA * COST	4.33**	3.26*	3.03*	4.10**	3.13*
OPR DATA * T	3.75*	3.68*	4.38**	3.55*	3.14*
OPR DATA * COST&T	3.89*	4.32**	3.81*	4.68**	3.26*
OPR DATA * COST&R	4.05**	3.65*	4.27**	4.57**	3.18*
OPR DATA * BUSUNIT	4.15**	4.68**	5.08***	4.75**	4.98**
AFFIAD DATA * COST	2.85	3.25*	3.23*	3.37*	2.88
AFFIAD DATA * T	4.04**	3.58*	3.38*	3.55*	3.18*
AFFIAD DATA * COST&T	3.25*	3.25*	4.23**	3.53*	4.23**
AFFIAD DATA * COST&R	3.95*	3.71*	4.27**	4.50**	3.48*
AFFIAD DATA * BUSUNIT	4.13**	4.68**	5.08***	4.75**	4.98***
OPR SYS	3.33*	2.01	1.05	2.89	2.01
OPR DATA	4.67**	3.45*	5.25***	4.56**	3.89**
AFFIAD DATA	3.66*	3.33*	5.05***	5.55***	5.01***
F value	14.3	17.33	24.04	17.35	11.66
R^2	0.32	0.30	0.37	0.41	0.42
Adjusted R^2	0.30	0.28	0.32	0.37	0.39

***Significant at the 0.01 level; **Significant at the 0.05 level; *Significant at the 0.10 level.
[a]Dependent variables: the success of using the BSC in measuring costing decisions, efficiency measures, quality and customer satisfaction measures, continuous innovation measures and contract negotiation practices.

($p<0.1$). Data integration among operating units (OPR DATA) is positively and significantly related to determining costs ($p<0.05$), measuring efficiency ($p<0.1$), assessing quality and customer satisfaction ($p<0.01$), promoting continuous innovation ($p<0.05$) and monitoring contract negotiation ($p<0.05$). Data integration among affiliated hospitals (AFFILIAD DATA) is found to be positively and significantly associated with the extent of using the BSC in all decision perspectives ($p < 0.1$ for costing and efficiency, $p < 0.01$ for quality, continuous innovation and contract negotiation). The statistics indicate that the null hypotheses of H2b (OPR DATA) and H2c (AFFILIAD DATA) cannot be rejected. However, the results do not show significant support for H2a (OPR SYS).

DISCUSSION AND CONCLUSION

The BSC strategic approach emphasizes the use of the information systems to track a limited number of balanced metrics (measures and indicators) for performance measurement. This study adds to our understanding of the way different levels of autonomy and different forms of system integration relate to the success of using the BSC for performance measurement. The success of using the BSC in performance evaluation is considered in five contexts: cost determination, efficiency measurement (internal business perspective), quality and customer satisfaction measurement (customer perspective), continuous innovation (innovation and learning perspective) and contract negotiation performance (financial). The findings suggest that hospitals need streamlined, information integration across the continuum of care to better assess their operating results, in both organizational and technical perspectives.

From a social perspective, the results indicate that the level of autonomy of clinical units is not significantly associated with the use of the BSC. However, if the performance measure of clinical units ties with the technology integration, then the level of autonomy is found to be significantly related to the use of the BSC. The results are particularly significant when the clinical units are treated as business units. Considering performance evaluation of a hospital's subunit, different operating units within a hospital may have deployed different performance measures. Performance measure for cost centers such as security, housekeeping and medical records departments will be based on cost, efficiency and quality of service measures. On the other hand, performance measure for profit centers like radiology and pharmacy departments will not only use cost and quality measures but also profitability measure. However, in some departments or organizations, it is difficult to measure individual performance. In these instances, individual performance measures may not be appropriate because they do not provide a good indication of team work and cooperation within a department that may be essential to its success. For example, for measuring the performance of performing an open heart surgery, one cannot rely solely on the cardiac department's profit or efficiency measures because the process involves many operating units across the hospital to complete the case. In this case, information about multiple financial and nonfinancial performance measures based on the BSC needs to be gathered not only from a single unit but also from other clinical units at both individual and group levels in order to capture the results of the care process. Integration among departments, including agreement on group-level performance measures and exchange of performance-related information, is critical to facilitate the implementation of the BSC. When the information systems are integrated, this study demonstrates that a significant relationship exists between level of autonomy and the use of BSC when the clinical units are treated as business unit. Bouwens and Abernethy (2000) comment that the organizational structure must facilitate the efficient flow of information, both horizontally and vertically, to develop collaborative delivery of core health care series within the organization. This article, however, argues that the effectiveness of the organization structure to successfully facilitate BSC depends on technical system integration that supports the work flow and communicates information for performance measurement.

Structurally, access to information is characterized by an organization's information-flow system, including receipt of reports, content of reports, distribution of memos and related information and participation in discussions that lead to formal decisions. The results indicate that connected and integrated computer operating systems software will assist in reporting and providing real-time information to support performance measures based on the BSC. The BSC could be designed specifically to be a computer-based "quick glance" performance-tracking tool. It could provide detailed reports that linked to the organization's key performance indicators. For instance, a manager in the nursing department would be able to access trended data on staff shift assignments or the turnover rate for all of the clinical cost centers. At a detailed level, the manager could compare the figures with related cost data, service hours by nurse type including overtime hours and daily and weekly census to arrange different types of nursing staff for different tasks at different care units. Registered nurses can be substituted for licensed nurses, or vice versa. Managers can reassign registered nurses and licensed nurses from one shift to another within a 24-h period to optimize resource utilization within their budget constraints.

In addition, the findings suggest that information integration across affiliated hospitals is positively related to the extent of using the

BSC in all decision contexts. Integrated systems allow hospital management to identify profit margin by payer, by physician, by service or by product line across affiliated health systems and at the same time support the appropriate decisions about resource allocation and contract negotiation.

Overall, this study suggests that hospitals should develop an appropriate information system to support the implementation of a BSC. Incomplete knowledge of operations, data validity and consistency issues due to lack of integration among affiliated hospitals, operating units and operating systems may limit the quality of the performance measures facilitated by the BSC. It is especially important for the affiliated health care organizations to facilitate patient case management and shared care while provide both financial and nonfinancial information that is useful to their associated partners in economic decision making, quality assessment and resource management.

LIMITATIONS AND FUTURE RESEARCH DIRECTION

Because the instrument used to measure system integration was intentionally developed for a hospital setting, it requires further testing of its reliability and validity. The findings of this study were limited to the hospital context and further research needs to be undertaken to examine the research questions in other industries. The length of the questionnaire in this study limited the ability to explore completed concepts of system integration; future research should take a more business process view of health care delivery and identify the appropriate organizational and information infrastructure measures for system integration.

The success of using the BSC is also influenced by other management innovations that are concurrently adopted by the hospitals. For instance, a hospital may be employing a total quality management program as it begins to alter processes as a result of implementing a BSC. The management control literature identifies characteristics of control systems that may be critical to the successful implementation of a BSC. Future research may focus on whether management innovation efforts such as benchmarking and total quality management promote a desired environment that allows an integrated system to provide appropriate information for a BSC to evaluate performance in various perspectives.

REFERENCES

Abernethy, M. A., & Lillis, A. M. (2001). Interdependencies in organization design: A test in hospitals. *Journal of Management Accounting Research*, *13*, 107–129.

Baker, G. R., & Pink, G. H. (1995). A balanced scorecard for Canadian hospitals. *Healthcare Management Forum*, *8*(4), 7–21.

Batey, M. V., & Lewis, F. M. (1982). Clarifying autonomy and accountability in nursing service: Part 1. *Journal of Nursing Administration*, *12*(9), 13–18.

Beauchamp, G. R. (1999). The five "S" levels of enterprise health. *Physician Executive*, *25*(3), 25–29.

Boaden, R. J. (1991). Organizing for CIM: Project management, technology and integration. *Journal of Computer-Integrated Manufacturing systems*, *4*(2), 60–70.

Bojan, B., & Koncar, M. (2007). HL7 and DICOM based integration of radiology departments with healthcare enterprise information systems. *Journal of Medical Informatics*, *76*, S425–S432.

Bouwens, J., & Abernethy, M. (2000). The consequence of customization on management accounting system design. *Accounting, Organization and Society*, *25*(3), 221–242.

Burns, L. R., Anderson, R. M., & Shortell, S. M. (1993). Trends in hospital/physician relationships. *Health Affairs*, *12*(3), 213–223.

Cajulis, C. B., & Fitzpatrick, J. J. (2007). Levels of autonomy of nurse practitioners in an acute care setting. *Journal of the American Academy of Nursing Practitioners*, *19*, 500–507.

Castaneda-Mendez, K., Mangan, K., & Lavery, A. M. (1998). The role and application of the balanced scorecard in the healthcare quality management. *Journal of Healthcare Quality*, *20*(1), 13–19.

Chong, S. A., Verma, S., Mythily, S., Poon, L. Y., & McGorry, P. (2008). The early psychosis intervention programme in Singapore: A balanced scorecard approach to quality care. *Journal of Mental Health*, *17*(1), 79–91.

Chow, C. W., Ganulin, D., Teknika, O., Haddad, K., & Williamson, J. (1998). The balanced scorecard: A potent tool for energizing the focusing healthcare organization management. *Journal of Healthcare Management*, *43*(3), 263–280.

Cullen, C. (2000). Autonomy and the nurse practitioner. *Nursing Standard*, *14*(21), 53–56.

Curtright, J. W., Stolp-Smith, S. C., & Edell, E. S. (2000). Strategic performance management: Development of a performance measurement system at the Mayo clinic. *Journal of Healthcare Management*, *45*(1), 58–68.

Dempster, J. S. (1994). Autonomy in practice: Conceptualization, construction and psychometric evaluation of an empirical instrument. *Dissertation Abstract International*, *50*(07), 3320A(UMI No. 9030752).

Dixon, J. R., Nanni, A. J., & Vollman, T. E. (1990). *The new performance challenge: Measuring operations for world-class competition*. Homewood, IL: Dow Jones-Irwin.

Dworkin, G. (1988). *The theory and practice of autonomy*. Cambridge: Cambridge University Press.

Flamholtz, E. G., Das, T. K., & Tsui, A. S. (1985). Toward and integrative framework of organizational control. *Accounting, Organizations and Society*, *10*(1), 35–50.

Freidson, E. (1975). *Doctoring together: A study of professional social control*. Chicago, IL: University of Chicago Press.

Gordon, E., Chapman, R., Kunov, H., Dolan, A., & Carter, M. (1998). Hospital management decision support: A balanced scorecard approach. *Medinfo*, *9*(Part 1), 453–456.

Grimson, J., Grimson, W., & Hasselbring, W. (2000). The SI challenge in the health care. *Communications of the ACM*, *43*(6), 49–55.

Harber, B. W. (1998). The balanced scorecard at Peel Memorial Hospital. *Hospital Quarterly*, *1*(4), 59–63.

Hasselbring, W. (2000). Information system integration. *Communications of ACM*, *43*(6), 32–38.

HCIA Directory of Health Care Professionals. (2000). Baltimore, MD: HCIA Inc.

Hospital Peer Review Archives (HPRA). (2000a). Achieve balance from your scorecard strategy. *Hospital Peer Review Archives*, Document ID: PB20000914010000248, pp. 1–4.

Inamdar, N., Kaplan, R. S., & Bower, M. (2002). Applying the balanced scorecard in healthcare provider organizations. *Journal of Healthcare Management*, *47*(3), 179–195.

Jaana, M., Ward, M. M., Paré, G., & Wakefield, D. S. (2005). Clinical information technology in hospitals: A comparison between the State of Iowa and two provinces in Canada. *International Journal of Medical Informatics*, *74*(9), 719–731.

Jensen, A., & Sage, A. (2000). A systems management approach for improvement of organizational performance measurement systems. *Information Knowledge Systems Management*, *1*(1), 33–61.

Johnson, G. V., & Tingey, S. (1976). Matrix organization: Blue print of nursing care organization for the 80's. *Hospital and Health Services Administration*, *21*, 27–39.

Kaplan, R. S., & Norton, D. P. (1993). Putting the balanced scorecard to work. *Harvard Business Review*, *71*(5), 134–139.

Kaplan, R. S., & Norton, D. P. (1996). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, *74*(1), 75–86.

Kaplan, R. S., & Norton, D. P. (2001). *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*. Boston, MA: Harvard Business School Publishing Corporation.

Katz, F. E. (1968). *Autonomy and organization*. New York: Random House.

Krol, M., Reich, D., & Dupont, J. (2005). Multi-platforms medical computer systems integration. *Journal of Medical Systems*, *9*(3), 259–270.

Lawrence, P. R., & Lorsch, J. W. (1967). *Organization and environment*. Boston, MA: Harvard Business School.

Lewis, E. P. (1975). Everybody's patient is nobody's patient. *Nursing Outlook*, *23*, 551.

Lipe, M. G., & Salterio, S. E. (2000). The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review*, *75*(3), 283–299.

Lipe, M. G., & Salterio, S. E. (2002). A note on the judgmental effects of the balanced scorecard's information organizations. *Accounting, Organizations and Society*, *27*, 531–540.

Maas, M., & Jacox, A. K. (1977). *Guidelines for nurse autonomy/patient welfare*. New York: Appleton Century Crofts.

MacDonald, M. (1998). Using the balance scorecard to align strategy and performance in long-term care. *Healthcare Management Forum*, *11*(3), 33–38.

Malina, M., & Selto, F. (2001). Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard. *Journal of Management Accounting Research*, *13*, 47–90.

Mathias, J. M. (2001). Balanced scorecard helps ORs to weigh clinical. *Fiscal Issues*, *17*(9), 1, 18–20, 22.

McKay, P. S. (1983). Interdependent decision-making: Redefining professional autonomy. *Nursing Practitioners Administrative Quarterly*, *7*(4), 21–30.

Mundinger, M. (1980). *Autonomy in nursing*. Germantown, MD: Aspen Systems.

Oliveira, J. (2001). The balanced scorecard: An integrative approach to performance evaluation. *Healthcare Financial Management*, *55*(5), 42–46.

O'Sullivan, D. (1992). Development of integrated manufacturing systems. *Computer Integrated Manufacturing Systems*, *5*(1), 39–53.

Paré, G., & Sicotte, C. (2001). Information technology sophistication in health care: An instrument validation study among Canadian hospitals. *International Journal of Medical Informatics*, *63*(3), 205–223.

Peters, K., & Ryan, H. (1999). An integrated dialysis delivery network in Ontario. *Journal of the Canadian Association of Nephrology Nurses and Technicians*, *9*(1), 20–23.

Shortell, S., Gillies, R., Anderson, D., Erickson, K., & Mitchell, J. (2000). *Remarking healthcare in America*. San Francisco: Jossey-Bass.

Sugarman, P. A., & Watkins, J. (2004). Balancing the scorecard: Key performance indicators in a complex healthcare setting. *Clinician in Management*, *12*(3), 129–132.

Thomas, R., Robinson, J., Waring, T., Wainwright, D., & Maguire, S. (1995). Information management and technology in England's large acute NHS Hospitals: National strategy versus local reality. *Journal of Management in Medicine*, *9*(1), 40–49.

Vancil, R. F., & Buddrus, L. E. (1979). *Decentralization: Managerial ambiguity by design*. Homewood, IL: Dow Jones-Irwin.

Voelker, K., Rakich, J., & French, R. (2001). The balanced scorecard in healthcare organizations: A performance measurement and strategic planning methodology. *Hospital Topics*, *79*(3), 13–24.

Waring, T., & Wainwright, D. (2000). Interpreting integration with respect to information systems in organization – Image, theory and reality. *International Journal of Information Management*, *15*(2), 131–147.

Weiner, S. L., Maxwell, J. H., Sapolsky, H. M., Dunn, D. L., & Hsiao, W. C. (1987). Economic incentives and organizational realities: Managing hospitals under DRGs. *Mibank Memorial Fund Quarterly*, *65*(4), 463–487.

Wicks, A. M., & Clair, L., St. (2007). Competing values in healthcare: Balancing the (UN) balanced scorecard. *Journal of Healthcare Management*, *52*(5), 309–324.

Young, D., & Saltman, R. (1985). *The hospital power equilibrium – Physician, behavior and cost control*. Baltimore, MD: John Hopkins University Press.

Zachman, J. A. (1987). A framework for information systems architecture. *IBM Systems Journal*, *26*(3), 276–292.

APPENDIX. QUESTIONNAIRE INSTRUMENT

The Extent of Using Balanced Scorecard in Health Care Organizations

---

Organization name: ______________________________ (optional)
Location of your hospital:______________ (state)
Position in organization:_____________________

---

1. Number of inpatient beds: ______________________________
Average annual operating revenue: ______________________

2. To what extent would you agree with the following items? **(1-strongly agree, 7-strongly disagree)**

• Clinical units are responsible for costs incurred in their units?	1	2	3	4	5	6	7
• Clinical units are responsible for managing throughput in their units?	1	2	3	4	5	6	7
• Clinical units are now being treated as business units	1	2	3	4	5	6	7
• Clinical units are responsible for both costs and revenues	1	2	3	4	5	6	7
• We have developed "contracts" with our clinical unit managers that hold them accountable for both costs and throughput targets.	1	2	3	4	5	6	7

3. To what extent would you agree with the following items: **(1-strongly agree, 7-strongly disagree)**

• General ledger, medical records and billing processes are highly Integrated. ("Processes are highly integrated" means information is exchanged and shared among authorized users.)	1	2	3	4	5	6	7
• Payroll, material management and department specific systems (e.g. nursing acuity, or operating room information systems) are highly Integrated.	1	2	3	4	5	6	7
• Software or operation systems are highly integrated among departments.	1	2	3	4	5	6	7
• Information systems are highly integrated across affiliated health organizations.	1	2	3	4	5	6	7

4. How would you rate the success of your hospital's balanced scorecard in evaluating the following performance

(1-very successful, 7-very unsuccessful)

• Identifying costs of intermediate and final products	1	2	3	4	5	6	7
• Measuring efficiency	1	2	3	4	5	6	7
• Evaluating quality and customer satisfaction	1	2	3	4	5	6	7
• Promoting continuous innovations	1	2	3	4	5	6	7
• Monitoring contract negotiation	1	2	3	4	5	6	7

*** Balanced Scorecard is an approach that extends performance evaluation from merely looking at financial results to formally incorporating measures that look at customer satisfaction, internal business processes, and the learning and growth potential of the organization.

Please put any comments here:

Your participation in this research study will be greatly appreciated.

FINDING AN INTERNAL OPTIMUM IN THE CLASSIFICATION OF MANAGEMENT ACCOUNTING INFORMATION: THE ROLE OF FUZZY SETS

Harry Zvi Davis, Roger Meszník and John Y. Lee

ABSTRACT

This article contributes to the fuzzy logic application literature in accounting by examining a key issue in the use of fuzzy logic: how to find an optimum number of classes to minimize the decision maker's cost. Two costs are assumed: (1) we assume fuzziness is costly and thus should be minimized and (2) we assume that adding categories is costly. In order to address the issue of finding the optimal number of classes, we define the objective function as being cost minimization. We seek to determine the costs and benefits of increasing the number of classifications and ask whether an internal optimum is identifiable and achievable. We assume, ceteris paribus, less fuzziness is preferable to more fuzziness, but fuzziness can only be reduced through the use of more categories whose creation is costly. More fuzziness is costly, but so is the creation of additional categories to alleviate the fuzziness. When we arrive at the optimal number of clusters that corresponds to a minimal total cost, that number may not be the same as the "natural" number of categories. It is, nonetheless,

Advances in Management Accounting, Volume 17, 203–216
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017009

a useful and practical way of deciding on the number of classifications. The approach we employ in this study is not confined to a management accounting information environment. It can be applied to any information environment where measurable classifications exist.

INTRODUCTION

Fuzzy sets, originally introduced by Zadeh (1965) as an extension of the classical and binary sets, allow the gradual assessment of the membership of elements in a set. Since assigning observations to classes is fundamental to the process of scientific discovery, fuzzy set theory can play a very significant role in various disciplines.

In decision sciences, for example, the members of a set exhibit the presence or absence of an attribute along a continuum, and not as a distinct binary condition. Strictly speaking, the answer to the question whether or not a certain attribute exists is usually not a simple yes or no. Between the "clear yes" and the "unambiguous no" there is a whole range of "a bit", "both", "rather more of the one", etc. And yet, in the classical modes, the decision maker pretends that the answer is a clear "yes" or "no" when asked to create classes. The classes, by definition, are mutually exclusive. The boundary is unique. Any member of the set belongs to one class, and only one class.

Having bisected the continuum,[1] the scientist creates the illusion of a clear break, a mutual exclusion, and unambiguous class memberships. In this Aristotelian view, there is no *excluded middle*. All set members are uniquely classified. The benefit of adding more classes is neither conveniently measurable, nor even necessarily evident.

Fuzzy logic provides an opportunity to deal differently with the problem of classifications.[2] Instead of forcing mutually exclusive classifications on the members of the set, which are part of a continuum, it allows for partial, or fuzzy, memberships in more than one class. The levels of inventory, for example, are not seen as either correctly or incorrectly valued. Rather, they are seen as being valued more or less correctly. A member being classified is not assigned a binary 0 or 1, but a fractional number, which designates partial membership in the class.

Fuzzy logic has its own formal structure, assumptions, and restrictions. The theory has recently been applied in various business disciplines, including accounting. In this article, we address one issue that has not been studied thus far: finding an internal optimum in the classification of management accounting information. We base our research on the relationship between class design efficiency and cost minimization in the context of fuzzy logic. The findings can be applied to a broad spectrum of scientific and managerial applications.

The article is organized as follows: in the next section, we provide a literature review in the application of fuzzy logic and describe our research objective. Then we develop our model using the main propositions. The following section discusses the implications of the assumptions made in the study and extends the model applicability to other areas. The applicability of our model is further interpreted and illustrated using a numeric analysis. Lastly, we provide a brief summary and conclusion.

LITERATURE REVIEW

Many studies have explored the possibility of applying fuzzy logic to various areas of business. In the 1990s, there were a number of studies on the feasibility of using fuzzy logic in financial decision-making (Brewer, Gation, & Reeve, 1993; De Korvin, Strawser, & Siegel, 1995; Gulledge, 1993; Chiu & Park, 1994; Lai, Liu, & Hwang, 1994; Masui, Terano, Yumino, & Mimori, 1994; Turtle, Bector, & Gill, 1994; Wong, Wang, Goh, & Quek, 1992).

More recently, fuzzy logic has been applied to option pricing (Lee, Tzeng, & Wang, 2005), financial risk forecasting (Thiagarajah & Thavaneswaran, 2006), adaptive hierarchical system development using generic algorithms (Mohammadian & Kingham, 2004), product-mix decisions (Bayou & Reinstein, 2005), project management (Lam et al., 2001), modeling constraint satisfaction (Yager, 2004), stock allocation in a distribution supply chain (Xie & Petrovic, 2006), sampling methods in quality research (de Korvin & Shipley, 2005), finding system solutions in linear programming (Muzzioli & Reynaerts, 2007), and financial modeling under uncertainty (Zmeskal, 2005).

Studies of fuzzy logic have also dealt with various accounting issues. The theory has been applied to evaluating investments in advanced manufacturing technology (Abdel-Kader & Dugdale, 2001), valuing financial instruments (Simonelli, 2001), human resource allocation in a CPA firm (Kwak, Shi, & Jung, 2003), evaluating the quality function deployment and value engineering for target costing (Gandhinathan, Raviswaran, & Suthakar, 2004), assessing internal control risks in a computer-based accounting information system (De Korvin, Shipley, & Omer, 2004), exposition of the antecedents of audit fees (Beynon, Peel, & Tang, 2004), selecting the optimum mechanism for developing accounting standards (Bayou, de Korvin, & Reinstein, 2007), and testing the relationship between organization configurations and management accounting system changes (Cassia, Paleari, & Redondi, 2005).

This eclectic selection of recent contributions shows the breadth and impact of fuzzy logic on a wide range of applications and uses. Each contribution accentuates the benefits of using fuzzy logic. While the studies on the application of fuzzy logic in business venues have been varied, vast, and rapidly expanding, the common denominator of all the studies is clear and evident: it is based on the fuzziness in object classifications. Despite the appeal of employing fuzzy logic as a new theory in varied applications, studies have not yet advanced to a stage where the use of fuzzy logic is proven to be cost effective compared to the classical non-fuzzy classification system.

RESEARCH OBJECTIVE

We attempt to contribute to the fuzzy logic application literature in accounting by examining a key issue in the use of fuzzy logic: how to find an optimum, which reduces a decision maker's costs of reducing fuzziness. In particular, we posit the following question:

If a set of items is to be divided into classes (categories) using fuzzy logic, what is the optimal number of classes?

The question about the optimum number of classes is not new. For example, Thorndike (1953) raised the issue long time ago. However, when raised, the issue has always been an attempt to identify the best "natural" number of classes. When correctly analyzed, the nature of the items to be classified and the objective of the classification are supposed to reveal to the researcher what the "natural" classes ought to be, and what optimal number of classes will fulfill the requirements of the classification. The broad context of the classification is supposed to determine the classification structure and the number of categories.[3]

In this article, we try to answer the same question, but we approach it differently. We define the objective function as being cost minimization. We seek to determine the costs and benefits of increasing the number of classifications and ask whether an internal optimum is identifiable and achievable. We assume, ceteris paribus, less fuzziness is preferable to more fuzziness, but fuzziness can only be reduced through the use of more categories, whose creation is costly. More fuzziness is costly, but so is the creation of additional categories to alleviate the fuzziness. When we arrive at the optimal number of clusters that corresponds to a minimal total cost, that number may not be the same as the "natural" number of categories. It is, nonetheless, a useful and practical way of deciding on the number of classifications.

In an Aristotelian world, the addition of new classes creates new discontinuities in the memberships which are, by definition, mutually exclusive. Given these discontinuities, the costs and benefits of adding more classes are neither conveniently measurable nor necessarily evident. In a fuzzy logic world, the addition of classes enhances the resolution of the classifications in a continuous manner, thus enabling the researcher to exercise clear choices of costs and benefits. These are the trade-offs we model in this article. We explore the opportunity to optimize the number of classes in a fuzzy logic context. Accepting the use of fuzzy logic, we investigate the trade-offs attendant to the choice of the number of fuzzy classifications.

THE MODEL

The Basic Design

Assume a universe of m elements, defined over a domain X such that

$$X_i = \text{Element of universe } X, \quad \text{where } i = 1, 2, \ldots, m \tag{1}$$

Assume that this universe of m elements can be classified based on some arbitrary class definition D_j and that n such class definitions are possible. Then we have a population

$$D_j, \quad \text{where } j = 1, 2, \ldots, n \tag{2}$$

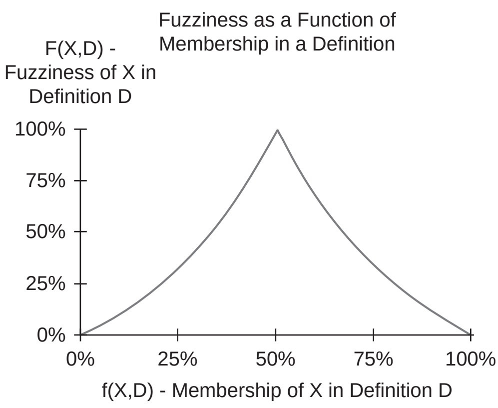
Using the theory of fuzzy logic, we stipulate a membership function $f(X_i, D_j)$, which defines for every element X_i its degree of membership in the definition set D_j. This function gives a continuous rather than a binary (yes/no) type value. This function is bounded by the closed interval $f(X_i, D_j) = [0, 1]$, and is defined at its extreme values as follows:

$$\begin{gathered} f = 1 \text{ defines full membership of } X_i \text{ in } D_j \\ \text{and} \\ f = 0 \text{ defines full non-membership of } X_i \text{ in } D_j \end{gathered} \tag{3}$$

We can now define the fuzziness of the membership in relation to a specific definition as follows:

$$F(X_i, D_j) = \left\{ \begin{array}{ll} \dfrac{f(X_i, D_j)}{1 - f(X_i, D_j)} & \text{if } f(X_i, D_j) \leq 0.5 \\ \dfrac{1 - f(X_i, D_j)}{f(X_i, D_j)} & \text{if } f(X_i, D_j) \geq 0.5 \end{array} \right\} \qquad (4)$$

For "full membership" of X_i in D_j, when $f = 1$, then $F = 0$. For "full non-membership", defined by $f = 0$, then $F = 0$. For the case that X_i, is, in equal parts, a member and a non-member, $f = 0.5$, and the function F peaks at $F = 1$. In this sense, the function F defines the *degree of fuzziness* of membership of X_i in D_i. It equals zero when X_i is clearly and totally a member of D_j, or clearly and totally a member of the complementary set "non-D_j", denoted as $\bar{D}_j$.[4] The function F peaks at 1 when X_j is at one-half a member of D_j and at the other half a non-member of D_j. A value of $F = 0$ denotes no fuzziness for X_i in an unambiguous classification. A value of $F = 1$ provides no information for the classification (see graph below).



Obviously, the fuzziness of each element X_i depends upon the membership set D_j. A membership definition, which is particularly suitable for a given X and reduces its fuzziness, may raise the fuzziness of another X element.

From the assumption that n definitions of classifications exist, we can define the fuzziness of an element X_i in our system of n definitions as follows:

$$F(X_i, n) = \min[F(X_i, D_1), F(X_i, D_2), \ldots, F(X_i, D_n)] \qquad (5)$$

The fuzziness of an element is defined as the minimum fuzziness achievable over all possible membership definitions. The intuition for this definition is that the minimum fuzziness conveys the maximum amount of information about the element.

For the total universe of elements X_i, fuzziness is defined as the arithmetic sum over the fuzziness of all individual members, viz.

$$F_n = \sum_{i=1}^{m} F(X_i, n) \tag{6}$$

Obviously, the total fuzziness of a system depends upon the number of membership definitions provided to the members, in accordance with Eq. (4).

Since classifications and memberships therein are decision variables, one can imagine another set of n membership definitions, denoted as n_2, which yields a different system fuzziness F_n. Assuming that p such groupings are possible, then we define an F_n^* such that

$$F_n^* = \min[F_{n_1}, F_{n_2}, \ldots, F_{n_p}] \tag{7}$$

F_n^* depends upon

– The number of elements in our domain m.
– The number of membership systems p.

To allow for the subsequent optimization, we offer the theorem that

$$F_{n+1}^* \leq F_n^* \tag{8}$$

This theorem is justified as follows: If more classifications increase the information value of the system, then the fuzziness for $n + 1$ classification definitions can be no more than the one encountered for n since, at worst, one of the classifications is redundant, thus not adding to, but certainly not reducing the information content.

To simplify the subsequent optimization, we suggest that expression (8) is a strict inequality. It means the function is assumed to be strictly monotonic in the number of classifications.[5] Therefore, we restate expression (8) as follows:

$$F_{n+1}^* < F_n^* \tag{9}$$

When classifying the elements, the decision maker has a goal in mind, a need to impose order on a set of elements. With this goal in mind, we can now define the *cost of fuzziness*, *CF*, as being the following function:

$$CF = CF_G(F_{n_i}) \tag{10}$$

where the subscript G is a parameter that denotes the goal of the decision maker who imposes the classification.[6] In accordance with our assumption that better classifications are more desirable, we assume that

$$\frac{\partial CF}{\partial F_{n_i}} > 0 \tag{11}$$

that is, the more the fuzziness in the classification, the higher is the cost. It follows trivially from the above that for any given number of definitions n, the optimal set is the one whose fuzziness is minimized. Thus, when $n = 2$, the set of definitions which have F_2^*-fuzziness is optimal, and, when $n = 3$, the set with F_3^* is optimal. Accordingly, the set $\{F_2^*, F_3^*, \ldots, F_n^*\}$ represents the fuzziness associated with the efficient frontier of definitions. Any definition set for which $F_i > F_i^*$ is non-optimal since the fuzziness can be reduced further while maintaining the same number of definitions.

Our second assumption is that there is a cost attached to classifying the elements and populating the classifications, and that the cost increases with the number of definitions. We therefore introduce the cost of classification of the elements and label it CC. This CC is assumed to be a function $CC = CC_G(n)$.

We can now observe the trade-off confronting a decision maker. Increasing the number of classifications n reduces the fuzziness in the system while raising the costs of the process. In terms of our notation, we have

$$CF[F_{n+1}] < CF[F_n] \quad \text{while} \quad CC[n+1] > CC[n] \tag{12}$$

An Optimum and Its Properties

Stated differently, if fuzziness is costly because of the corollary poor resolution, and if reducing fuzziness is costly because of the increased effort and cost, what are the properties of the possibly available optimum? Under what terms will the total of these two cost components be minimized?

Mathematically, we are seeking

$$\mathrm{Min}|_n\{CF_G(F_n) + CC_G(n)\} \tag{13}$$

Applying standard optimization techniques,[7] and recalling that the decision variable is n, we get the following necessary condition:

$$\frac{\partial[CF(F_n)]}{\partial n} + \frac{\partial[CC(n)]}{\partial n} = 0 \tag{14}$$

Since $CF' > 0$, and $CC' < 0$, the optimum is defined by

$$\left[\frac{\partial[CC(n)]}{\partial n}\right] = (-1)\left[\frac{\partial[CF(F_n)]}{\partial n}\right] \tag{15}$$

To elucidate the process, we now invoke a standard assumption. We assume that the benefits of a decreasing fuzziness accrue at a decreasing rate, and that the costs of reducing fuzziness increase at an increasing rate. In either case, we invoke the common assumption of diminishing marginal returns. Accepting the validity of the diminishing marginal returns, we can now show the problem and its solution in a graph (see Fig. 1).

Model Extensions

In setting up Eq. (6) that defines the total fuzziness of all elements, we presented it as the arithmetic sum of the fuzziness of the individual elements. Obviously, this presupposes that the fuzziness is linearly additive. We assumed additivity based on the fact that the m elements of our model are all

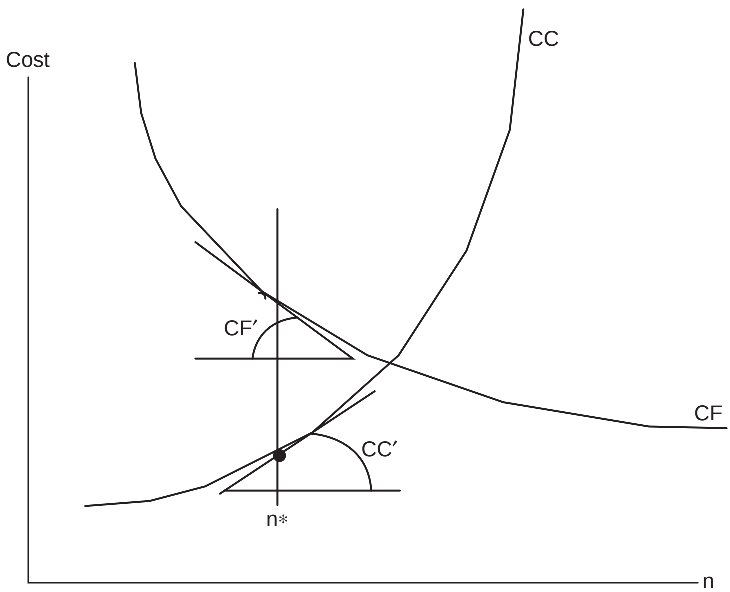


Fig. 1. Cost Minimization for n^* Optimal Categories.

considered equally important – an increase in fuzziness of one element can be offset against a loss in fuzziness of equal magnitude of another element. The linearity implies that the cost attributable to fuzziness of any one element is a linear function of the fuzziness. It is the simplest possible assumption, but certainly not the only defensible one. The model can be expanded to include a more complex cost function, such as a quadratic one. This linearity is clearly separable from the non-linearity assumed in Eq. (10), which defines the total cost of the fuzziness.

In expression (13), we assumed that both cost functions CC and *CF* to be everywhere differentiable and strictly convex. The assumed convexity is grounded in our objective of cost minimization. Local violation of the strict convexity would imply that, over some local region, fuzziness can be reduced at no added cost (function CC), or increased fuzziness can be accepted without a loss in resolution of the categorization (function *CF*).

Therefore, encountering such deviations from strict convexity would represent opportunities to gain benefits at no incremental cost. If present, such local deviations from strict convexity would represent unexercised opportunities to improve the solution at no additional cost. Their presence would violate our stated objective of cost minimization. The assumed differentiability is a matter of convenience. Its absence would require non-linear programming and be highly sensitive to the stipulated model design. While it would represent a possible extension of what we derive in this article, it would not alter the underlying process of balancing costs and benefits at the margin.

A NUMERICAL EXAMPLE

Within the environment of a management accounting information design, assume a universe of five elements X_i ($m = 5$). We want to classify these elements in accordance with a goal function G which need not be specified further.

We start by assuming one definitional set that includes three definitions with the following F-values, viz.:

(I)

	X_1	X_2	X_3	X_4	X_5
$F(X_i, D_1)$	1	0.8	0.2	0.3	0.5

$\Sigma F(X_i, D_1) = 2.8$

(II)

	X_1	X_2	X_3	X_4	X_5
$F(X_i, D_1)$	0	0.8	0.2	0.5	0.8

$\Sigma F(X_i, D_1) = 2.3$

(III)

	X_1	X_2	X_3	X_4	X_5
$F(X_i, D_1)$	0.5	0.2	0.1	0	0.1

$\Sigma F(X_i, D_1) = 0.9$

The last classification with the value of 0.9, a minimum, is the efficient frontier for this one set of classifications. It is designated below with $F(X_i, D_1)^*$.

We now introduce a second set of classifications, D_2. Assume that it has two possible classifications, and compare it to the efficient frontier achieved under D_1

(I)

	X_1	X_2	X_3	X_4	X_5
$F(X_i, D_1)^*$	0.5	0.2	0.1	0	0.1
$F(X_i, D_2)$	0.3	0.7	0.3	0.3	0.1
Minimum	0.3	0.2	0.1	0	0.1

We adopt the $F_{2,\min} = (F_{X_iD_1}, F_{X_iD_2}) = 0.7$ as the efficient frontier.

And for the second possible set of definitions for D_2, we have

(II)

	X_1	X_2	X_3	X_4	X_5
$F(X_i, D_1)^*$	0.5	0.2	0.1	0	0.1
$F(X_i, D_2)$	0.2	0.6	0.4	0.8	0.2
Minimum	0.2	0.2	0.1	0	0.1

The new efficient frontier is now given by $F_{2,\min} = 0.6$. This efficient frontier is defined over both sets of classifications.

By adding the second set of definitions, we lowered the cost of the fuzziness CF from 0.9 to 0.6. However, whether or not the second set of definitions should be used depends upon whether the cost of achieving this reduction exceeds the benefits of the reduction.

For example, assume that the costs of imposing the classifications happen to be:

$$CC^*_{(1)} = 0.5, \quad CC^*_{(2)} = 1.1 \tag{16}$$

Assume also that the costs of the fuzziness map one-to-one on the degree of fuzziness. A total fuzziness of 0.9 costs 0.9 units. Then we have the following total costs:

$$\begin{aligned} \text{Total Cost}_{(n=1)} &= CF^* + CC^* = 0.9 + 0.5 = 1.4 \\ \text{Total Cost}_{(n=2)} &= CF^* + CC^* = 0.6 + 1.1 = 1.7 \end{aligned} \tag{17}$$

In this case, the benefit of the decline in fuzziness is more than offset by the costs attendant to reducing the fuzziness with the ''better'' classification scheme.

SUMMARY AND CONCLUSIONS

A fuzzy logic application allows the decision maker to assess the ''quality'' of a classification scheme, and allows for easy comparisons among different classifications. Having dispensed with the binary, Aristotelian view, the quality of a classification is no longer a binary variable, but is rather measurable on a continuum. Most previous studies concentrated on measuring these processes.

In this article, we show how the benefits of an improved classification can be balanced against the corresponding costs. As can be clearly seen, this approach is not confined to a management accounting information environment. It is versatile for application to any information environment where measurable classifications exist.

In reality, any classification is costly. The improved classification may consist of no more than added categories to refine the process. However, even such an improvement is costly. We show that this improvement will be carried up to the point where, at the margin, the benefits from the improvement are exactly offset by the corresponding costs. To find a unique internal solution, we assume that the classification costs represent a monotonically increasing function of the classificatory quality that is being sought. We also assume a decreasing rate of benefits that accrue as refinement efforts in the classification increase.

NOTES

1. In these examples, the continuum is assumed to be divided into two classes. The arguments hold equally when the set is divided into more than two classes.
2. For a popular summary of the basics of Fuzzy Logic, see, e.g., McNeill & Freiberger (1993, passim).
3. See Bezdek (1974) and Dunn (1974).

4. Where

$$(D_j + \bar{D}_j) \equiv 1.$$

5. Accepting that additional classifications are costly, this assumption of strict monotonicity is easily justifiable on pragmatic grounds of cost minimization. Refining the classifications without reducing the fuzziness is presumably shunned by the decision maker since it increases costs without generating any incremental benefits.

6. In both functions, *CF* here and *CC* below, the term "*G*" enters only parametrically to represent the decision maker's goals and intent. Assuming the term *G* is constant, and therefore irrelevant for our optimization, does not alter the results.

7. Strictly speaking, the functions *CF* and *CC* may be made up of linked linear or non-linear segments if we assume that *n* can take on only integer values. Therefore, the functions are not everywhere differentiable, and taking derivatives may not always be the appropriate technique. We abstract from this problem by assuming a curvilinear approximation to a piecewise linear function without a loss in validity for our results.

REFERENCES

Abdel-Kader, M. G., & Dugdale, D. (2001). Evaluating investments in advanced manufacturing technology: A fuzzy set theory approach. *The British Accounting Review*, *33*(4), 455.

Bayou, M. E., de Korvin, A., & Reinstein, A. (2007). Using the fuzzy-analytic-hierarchical-process to select the optimum mechanism for developing accounting standards. *Review of Accounting and Finance*, *6*(2), 109.

Bayou, M. E., & Reinstein, A. (2005). Analyzing the product-mix decision by using a fuzzy hierarchical model. *Managerial Finance*, *31*(3), 35–48.

Beynon, M. J., Peel, M. J., & Tang, Y. C. (2004). The application of fuzzy decision tree analysis in an exposition of the antecedents of audit fees. *Omega*, *32*(3), 231.

Bezdek, J. C. (1974). Cluster validity with fuzzy sets. *Journal of Cybernetica*, *3*(3), 58–72.

Brewer, P. C., Gation, A. W., & Reeve, J. M. (1993). Managing uncertainty. *Management Accounting*, *75*(4), 39–45.

Cassia, L., Paleari, S., & Redondi, R. (2005). Management accounting systems and organisational structure. *Small Business Economics*, *25*(4), 373.

Chiu, C. Y., & Park, C. S. (1994). Fuzzy cash flow analysis using present worth criterion. *Engineering Economist*, *39*(2), 113–138.

De Korvin, A., & Shipley, M. F. (2005). Sampling rules to achieve quality, maximize benefit and minimize financial loss. *Managerial Finance*, *31*(3), 1–18.

De Korvin, A., Shipley, M. F., & Omer, K. (2004). Assessing risks due to threats to internal control in a computer-based accounting information system: A pragmatic approach based on fuzzy set theory. *Intelligent Systems in Accounting, Finance and Management*, *12*(2), 139–152.

De Korvin, A., Strawser, J., & Siegel, P. H. (1995). An application of control systems to cost variance analysis. *Managerial Finance*, *21*(3), 17–35.

Dunn, J. C. (1974). Well separated clusters and optimal fuzzy partitions. *Journal of Cybernetica*, *4*(1), 95–104.

Gandhinathan, R., Raviswaran, N., & Suthakar, M. (2004). QFD- and VE-enabled target costing: A fuzzy approach. *The International Journal of Quality and Reliability Management*, *21*(9), 1003.

Gulledge, D. E. (1993). FuziCalc. *Management Accounting*, *75*(1), 65.

Kwak, W., Shi, Y., & Jung, K. (2003). Human resource allocation in a CPA firm: A fuzzy set approach. *Review of Quantitative Finance and Accounting*, *20*(3), 277.

Lai, Y. J., Liu, T. Y., & Hwang, C. L. (1994). TOPSIS for MODM. *European Journal of Operations Research*, *76*(3), 486–500.

Lam, K. C., So, A. T. P., Hu, T., Ng, T., Yuen, R. K. K., & Lo, S. M. (2001). An integration of the fuzzy reasoning technique and the fuzzy optimization method in construction project management decision-making. *Construction Management and Economics*, *19*(1), 63.

Lee, C. F., Tzeng, G. H., & Wang, S. Y. (2005). A fuzzy set approach for generalized CRR model: An empirical analysis of S&P 500 index options. *Review of Quantitative Finance and Accounting*, *25*(3), 255.

Masui, S., Terano, T., Yumino, M., & Mimori, S. (1994). Decision making support system for house purchase by using fuzzy logic. *Computers and Industrial Engineering*, *27*(1–4), 281–284.

McNeill, D., & Freiberger, P. (1993). *Fuzzy logic: The revolutionary computer technology that is changing our world*. Old Tappan, NJ: Touchstone Books.

Mohammadian, M., & Kingham, M. (2004). An adaptive hierarchical fuzzy logic system for modelling of financial systems. *Intelligent Systems in Accounting, Finance and Management*, *12*(1), 61–82.

Muzzioli, S., & Reynaerts, H. (2007). Solving parametric fuzzy systems of linear equations by a nonlinear programming method. *Computational Economics*, *29*(2), 107–117.

Simonelli, M. R. (2001). Fuzziness in valuing financial instruments by certainty equivalents. *European Journal of Operational Research*, *135*(2), 296.

Thiagarajah, K., & Thavaneswaran, A. (2006). Fuzzy random-coefficient volatility models with financial applications. *The Journal of Risk Finance*, *7*(5), 503.

Thorndike, R. L. (1953). Who belongs in a family? *Psychometrika*, *18*, 267–276.

Turtle, H., Bector, C. R., & Gill, A. (1994). Using fuzzy logic in corporate finance: An example of a multinational cash flow netting problem. *Managerial Finance*, *20*(8), 36–53.

Wong, F. S., Wang, P. Z., Goh, T. H., & Quek, B. K. (1992). Fuzzy neural systems for stock selection. *Financial Analysts Journal*, *48*(1), 47–52.

Xie, Y., & Petrovic, D. (2006). Fuzzy-logic-based decision-making system for stock allocation in a distribution supply chain. *Intelligent Systems in Accounting, Finance and Management*, *14*(1/2), 27.

Yager, R. R. (2004). Constraint satisfaction using soft quantifiers. *Intelligent Systems in Accounting, Finance and Management*, *12*(3), 177–186.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353.

Zmeskal, Z. (2005). Value at risk methodology under soft conditions approach (fuzzy-stochastic approach). *European Journal of Operational Research*, *161*(2), 337.

AN EXPERIMENTAL STUDY OF THE EFFECT OF BUDGET FAVORABILITY ON THE FORMATION OF PSEUDO-PARTICIPATION PERCEPTIONS

Laura Francis-Gladney, Robert B. Welker and Nace Magner

ABSTRACT

Budget decision-makers are forced at times to assign budgets that deviate substantially from budget participants' requests. In these instances, budget participants likely interpret their budgetary involvement as lacking influence and perhaps as pseudo-participative. This experimental study examined two situational factors that may affect perceptions of pseudo-participation: budget favorability (receiving a much better or much worse budget than requested) and disclosure of budget intention (the decision-maker discloses or does not disclose a preliminary budget before the budget decision, with the final budget exactly matching the preliminary budget). As hypothesized, budget participants had a self-serving tendency to discount pseudo-participation as the cause of low influence when they received a favorable budget. However, contrary to a hypothesized effect, budget participants did not have a self-serving tendency to inflate

Advances in Management Accounting, Volume 17, 217–245
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017010

pseudo-participation as the cause of low influence when they received an unfavorable budget. Instead, they formed strong, unbiased pseudo-participation perceptions. Also contrary to a hypothesized effect, the budget decision-maker's disclosure of an intended budget, which should have provided clear indications of an insincere request for budget input, did not increase perceptions of pseudo-participation. Budget outcomes that indicate low influence may evoke such strong perceptions of pseudo-participation as to override other information that suggests pseudo-participation.

INTRODUCTION

Brownell (1982) describes budgetary participation as ''a process in which individuals, whose performance will be evaluated, and possibly rewarded, on the basis of their achievement of budgeted targets, are involved in, and have influence on, the setting of these targets'' (p. 124). It is a process of empowerment in which budget decision-makers, acting with sincerity, invite budget participants who have a stake in the budget outcome to provide opinions and recommendations regarding their budget. An extensive body of research has shown that participative budgeting processes can enhance organizational effectiveness by heightening budget participants' motivation (Hofstede, 1967; Searfoss & Monczka, 1973; Griffin, 1996), organizational commitment (Magner, Welker, & Campbell, 1995; Nouri & Parker, 1998), job satisfaction (e.g., Aranya, 1990; Frucot & Shearon, 1991; Dunk, 1992), and work performance (e.g., Brownell & McInnes, 1986; Govindarajan, 1986; Chalos & Haka, 1989; Mia, 1989; Aranya, 1990; Pasewark & Welker, 1990; Greenberg, Greenberg, & Houri, 1994; Chalos & Poon, 2001). An antithetical concept to budgetary participation is pseudo-participation, which involves making insincere solicitations for budget input without any intention of using the input information for the budget decision (Libby, 1999). Sham invitations for budget opinions and recommendations that successfully create the illusion of participation enable the budget decision-maker to make unilateral budget decisions without having to sacrifice budget participants' morale or risk their resentment from openly excluding them from the decision-making process. However, commentaries about the potential detriments of pseudo-participation abound (e.g., Argyris, 1952; Pateman, 1970; Nigro & Bellone, 1979; Libby, 1999; Maiga, 2005; Barsky, 1999 presents a case study of pseudo-participation effects). For instance, Foran and DeCoster (1974) state that ''Given unfavorable feedback ... about the acceptance of the proposed performance standards there would probably be a rejection of the standards and a lack of commitment to them'' (p. 753). Tiller (1983) makes the following observation from his research: ''A participative budgeting program which does not allow budget participants to perceive themselves as having exercised some choice in the budget-setting process is ... little different from a nonparticipative budgeting program'' (p. 593). Libby (1999) interprets Pasewark and Welker's (1990) results as suggesting that '' ... participation without influence is worse than no participation because it is demotivating to subordinates'' (p. 129). Conclusions such as these are consistent with early warnings by Argyris (1952) and Becker and Green (1962) that perceptions of pseudo-participation can undermine the benefits of participative budgetary processes.

In the current study, we accept the premise that it is in the organization's and the budget decision-maker's best interest to avoid perceptions of pseudo-participation, and we focus on situational factors (i.e., antecedents) that may influence pseudo-participation perceptions. Budget participants may not possess direct information as to whether pseudo-participation occurred. They may instead have to infer the decision-maker's sincerity from information they glean from events that transpire during the budget process. Budget decision-makers may behave in ways that inadvertently signal insincerity, even though their invitations for input were genuine. Our literature search did not turn up studies that look specifically at situational factors that condition budget participants' perceptions of decision-maker insincerity in soliciting budget input. That is, research has not examined the effects of budgetary events on pseudo-participation perceptions.

The focal event of our study is the release of the final budget, which is the culminating event of the budget formulation process. Its occurrence enables budget participants to determine the extent to which the budget outcome diverges from what they consider appropriate based on their budget input to decision-makers. We use the term ''budget favorability'' to refer to deviations from expectations, where a favorable budget indicates a better-than-expected budget outcome and an unfavorable budget indicates a worse-than-expected budget outcome. Participative input includes opinions and recommendations about appropriate performance targets or resource allocations to include in the budget. Conviction about the merit of one's opinions and recommendations and the presumption that the budget decision-maker made a sincere request for budget input create an expectation that the decision-maker will consider the input to be both informative and relevant and thus will make use of the input information when setting the final budget.

The final budget provides budget participants with feedback about whether the input was incorporated into the budget in accordance with expectations (Foran & DeCoster, 1974). Budgets that deviate substantially from expectations, either favorably or unfavorably, likely lead to perceptions of low influence. The budget decision-maker's disregard of the input information is contradictory to the notion that the input contained valuable opinions and recommendations. Relevant and useful information should have been incorporated into the budget. The contradiction produces incompatible cognitions. Budget participants' perception that the budget decision-maker chose to disregard the input information conflicts with their confident belief in the value of the budget input.

To reduce or eliminate the dissonance attributed to the incompatible cognitions, budget participants may seek explanations as to why the budget decision-maker did not use the input information (e.g., Festinger, 1957; Foran & DeCoster, 1974; Burris, Harmon-Jones, & Tarpley, 1997; Harmon-Jones & Mills, 1999). Causal explanations are crafted from perceived or imagined events or situational factors pertaining to the feedback. One viable causal explanation for low budget influence is pseudo-participation. However, budget participants may or may not consider it as *the* cause. For example, a possible alternative explanation for low budget influence is that the budget decision-maker was sincere in soliciting budget input but did not incorporate it into the final budget because of superior private knowledge or restrictions imposed by organizational superiors. Identification of pseudo-participation as the principal causal factor depends on whether budget participants set its causal viability at a high enough level to override the causal viability of other potential factors.

This study investigates the effect of two situational factors on budget participants' tendency to attribute pseudo-participation as the cause of low budget influence: whether the final budget deviated favorably or unfavorably from a requested budget, and whether the budget decision-maker disclosed his or her budget intention prior to the budget decision. Both situational factors should have an effect on the causal viability of pseudo-participation as an explanation for low influence. Regarding the first factor, budget participants may perceive pseudo-participation differently for favorable budgets than for unfavorable budgets because of self-serving attributions. Specifically, they may have a tendency to inflate the causal viability of pseudo-participation when they receive a budget that is less favorable than requested. They may do so to deflect the cause of unfavorable treatment away from self. Conversely, budget participants may have a tendency to discount the causal viability of pseudo-participation when they receive a budget that is more favorable than requested. They may do so in order to identify self as the cause of the favorable treatment. Thus, an important research question of the study concerns whether a budget that is substantially worse than requested has a different effect on perceptions of pseudo-participation than a budget that is substantially better than requested.

The second factor, whether the decision-maker disclosed his or her budget intention prior to the budget decision, was included in the study to ascertain whether budget participants' knowledge of the decision-maker's budget intention mitigates their tendency to produce self-serving perceptions of pseudo-participation in situations of low budget influence. A situation in which a decision-maker discloses a preliminary budget prior to the budget decision and then sets a final budget that exactly matches the preliminary budget should establish pseudo-participation as the most viable cause of low influence. The match between the preliminary budget and the final budget strongly suggests that the decision-maker intentionally disregarded the budget participant's input, and therefore pseudo-participation perceptions should be high in this situation. Budget participants cannot bias pseudo-participation upward for self-serving reasons when they receive an unfavorable budget since pseudo-participation is already perceived at a high level. Correspondingly, they cannot bias pseudo-participation downward for self-serving reasons when they receive a favorable budget since they possess information that provides a strong indication of pseudo-participation's existence. Two research questions in the current study are motivated by the theory presented above. One of these research questions addresses whether budget participants perceive more pseudo-participation when the budget decision-maker has disclosed his or her budget intention. The other research question relates to whether budget participants have less of a self-serving tendency to either reduce or inflate pseudo-participation perceptions in response to budget favorability when the budget decision-maker has disclosed his or her budget intention.

ATTRIBUTION THEORY

Attribution theory explains how people use information in their environment to form causal ascriptions of outcomes and events (Ployhart & Ryan, 1997; Schroth & Shaw, 2000). Attribution research documents people's self-serving tendency to see themselves as the cause of favorable outcomes and to lay responsibility for unfavorable outcomes on external factors such as bad luck or other people's actions (Miller & Ross, 1975; Zuckerman, 1979;

Fiske & Taylor, 1991; Brockner et al., 2002; Blader & Bobocel, 2005). Theorists believe that these tendencies stem from people's desire to either enhance or protect self-esteem (Schroth & Shaw, 2000; Blader & Bobocel, 2005). For instance, taking credit for a favorable outcome associates the outcome with personal characteristics that boost self-esteem, such as competence and diligence (Baumeister, 1982, 1989), and laying blame on others for an unfavorable outcome distances self from the outcome and, hence, from the negative connotations that an unfavorable outcome has about one's personal qualities, such as unworthiness and incompetence. Nouri, Kyj, and Dunk (1999) apply attribution theory to predict the effects of performance reports on budgetary participation perceptions. They show that experimental participants act in accordance with attribution theory, taking credit for good performance by perceiving more participation and distancing themselves from poor performance by reducing perceptions of participation.

Budget participants have self-serving reasons to judge participative budgeting processes leading to an unfavorable (worse-than-requested) budget outcome as pseudo-participative. Ascribing the unfavorable outcome to self suggests that one's budget views lack merit. To protect self-esteem, budget participants may have a self-serving tendency to place responsibility for the unfavorable outcome on an external source. Budget decision-makers represent a convenient external source to blame, and "insincere" may provide an expedient label to attach to the decision-maker, especially in the absence of information suggesting otherwise. Conversely, budget participants may have a self-serving tendency to judge participative processes leading to a favorable (better-than-requested) budget outcome as genuine. Ascription of low influence to self rather than to pseudo-participation embraces the self-favoring notion that the budget decision-maker intended to use the information contained in their budget input but decided that they deserved a more favorable budget. Pseudo-participation, on the other hand, implies disrespectful treatment of budget participants. It suggests that the budget decision-maker considered them as dupes. Disrespect can damage the budget participant's self-esteem (e.g., Lind & Tyler, 1988; Anderson & Hayes, 1996).

OVERVIEW OF EXPERIMENT AND DEVELOPMENT OF HYPOTHESES

To assess empirically whether budget participants form self-serving pseudo-participation attributions of low influence based on budget favorability, we conducted a laboratory experiment that involved budget participants, who were members of a work team assigned to perform a task, and a budget decision-maker, who set a personal budget for each member of the work team in terms of a performance target. The targets were used as a basis for computing rewards that budget participants would earn for task performance. At the invitation of the budget decision-maker, each budget participant made a formal request for a specific performance target. Budget participants received a final budget that differed substantially from the requested target. Thus, they had low budget influence. Budget favorability was manipulated by assigning a final budget that was either well above or well below the requested target. After budget participants received the final budget, they rated the level of pseudo-participation in the budget process. The experiment included three factors to assess whether the pseudo-participation perceptions were the product of a self-serving attribution bias.

Experimental Factors

One experimental factor was budget favorability, set at two levels: (a) a favorable budget, which was a much easier performance target than the one requested by the budget participant and (b) an unfavorable budget, which was a much more difficult performance target than the one requested by the budget participant. In a truly pseudo-participative process, the budget decision-maker gives no weight to the budget participant's request when setting the final budget. The budget outcome reflects the unilateral judgment of the budget decision-maker. The budget that the decision-maker has in mind can be either higher or lower than the requested budget. As such, whether the budget is favorable or unfavorable should have little bearing on assessments of pseudo-participation as long as both conditions imply the same degree of low influence. Therefore, differences in pseudo-participation perceptions between the unfavorable and favorable budget conditions, after controlling for any disparity in influence perceptions, suggest a biased causal interpretation of the budget outcome. The study included two additional factors to enhance experimental inferences about the self-serving attribution bias.

The second factor, disclosure of budget intention, concerned whether the budget decision-maker informed budget participants of his budget intention prior to making the budget decision. The factor had two levels: (a) disclosed budget intention, in which the budget decision-maker disclosed a preliminary performance target after receiving, but before reading, the participant's requested target, with the final target exactly matching the preliminary target and (b) non-disclosed budget intention, in which the budget decision-maker did not disclose a preliminary target. The disclosed budget intention condition was designed to provide budget participants with persuasive evidence that the budget decision-maker intended to use the preliminary budget without any consideration of the requested budget and thus had made an insincere solicitation for input. The disclosure of budget intention factor was added to the study to assess whether information that provided a strong suggestion of pseudo-participation had a mitigating effect on budget participants' tendency to bias their perceptions of pseudo-participation in a self-serving way.

The third factor, subject group, involved two groups of subjects: (a) budget participants and (b) a control group made up of observers who, after viewing the budget process, rated the sincerity of the budget decision-maker's input solicitation. The pseudo-participation perceptions of observers provide an indication of how people who have no stake in the budget outcome perceive low budget influence. As such, their perceptions provide a standard for measuring the strength of the budget participants' tendency to bias pseudo-participation perceptions. Our use of observers as a control group for measuring self-serving biases has support in the literature. Greenberg (1981) and Grover (1991) argue that self-serving tendencies cause individuals affected by a decision to judge the propriety of the decision process differently than do individuals who are unaffected by the decision. For instance, non-participative decision processes are rated as less fair by people with a stake in the decision than by people with no interest in the decision (Greenberg, 1981). Accordingly, we assessed budget participants' tendency to make self-serving pseudo-participation perceptions by comparing their pseudo-participation perceptions with the pseudo-participation perceptions by observers. The use of observers for obtaining a measure of unbiased perceptions is consistent with attribution theory. According to the theory, self-serving biases exist for actors but not for observers (Ross, 1977).

Hypotheses

The hypothesized effects of the three experimental factors on perceptions of pseudo-participation are graphically represented in Fig. 1 as differences between budget participants and observers. The first hypothesis of the study, which relates to the disclosure of budget intention factor, is reflected in Fig. 1 as a comparison between the top two lines (each representing the disclosed budget intention condition) and the bottom two lines (each

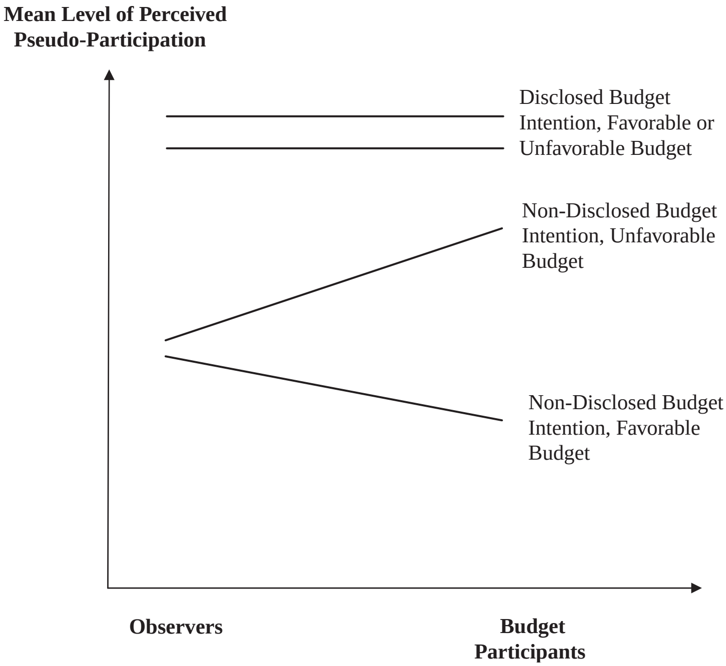


Fig. 1. Graphical Representation of Hypothesized Cell Means for the Experimental Factors (Disclosure of Budget Intention × Budget Favorability × Subject Group). The Disclosure of Budget Intention Factor has Two Levels: Disclosed Budget Intention and Non-Disclosed Budget Intention. The Budget Favorability Factor has Two Levels: Favorable Budget and Unfavorable Budget. There are Two Subject Groups: Observers and Budget Participants.

representing the non-disclosed budget intention condition). In the disclosed budget intention condition, the budget decision-maker disclosed his preliminary budget, and budget participants received a final budget that exactly matched the preliminary budget. These events should have supplied both budget participants and observers with persuasive evidence that the budget decision-maker never intended to use the information supplied by budget participants and that the participative budget process was a sham. According to Kelley's (1972) discounting principle, people discount alternative causes of an outcome when the plausibility of a given cause has been established. On this basis, we expect that budget participants and observers, in trying to explain low budget influence, will discount the viability of causes other than pseudo-participation in the presence of this strong, direct evidence of insincerity and that both groups will perceive high pseudo-participation for the disclosed budget intention condition. Thus, we advance the following hypothesis:

H1. For each budget favorability (favorable budget vs. unfavorable budget) condition, both budget participants and observers will perceive more pseudo-participation for the disclosed budget intention condition than for the non-disclosed budget intention condition

We also hypothesize that budget participants and observers will perceive similar levels of pseudo-participation for the disclosed budget intention condition, irrespective of budget favorability. The hypothesis is reflected in Fig. 1 as the two parallel lines at the top of Fig. 1. Similarity in pseudo-participation perceptions between budget participants and observers indicates that budget participants do not form self-serving perceptions of pseudo-participation for either favorable or unfavorable budget outcomes if they see that the decision-maker's preliminary budget exactly matched the final budget. While budget participants may have self-serving reasons to bias pseudo-participation perceptions upward for an unfavorable outcome and downward for a favorable budget outcome, they may be unable to do so when presented with persuasive evidence of pseudo-participation. Attribution research (e.g., Kelley, 1972; Schroth & Shaw, 2000) suggests that people are less apt to make self-serving causal attributions of outcomes when they are given persuasive evidence of the factor that caused the outcome. Thus, we advance the following hypothesis, which is presented in the null form because persuasive evidence of pseudo-participation suggests no differences in pseudo-participation perceptions between budget participants and observers:

H2. For each budget favorability (favorable budget vs. unfavorable budget) condition, budget participants will perceive pseudo-participation at the same level as observers for the disclosed budget intention condition.

The third hypothesis of the study relates to budget participants' tendency to make self-serving perceptions of pseudo-participation in situations of low budget influence. We assess their self-serving tendency by examining their pseudo-participation perceptions for favorable and unfavorable budgets for the non-disclosed budget intention condition (the two ordinal interactive lines in the lower part of Fig. 1). For an unfavorable budget, budget participants should have a tendency to perceive higher levels of pseudo-participation to deflect blame for the unfavorable outcome away from self. For a favorable budget, they should have a tendency to reduce their pseudo-participation perceptions to facilitate their attribution of the favorable outcome to self-enhancing personal attributes. These tendencies will manifest experimentally as pseudo-participation perceptions for unfavorable budgets that exceed the pseudo-participation perceptions for favorable budgets. While this difference indicates the presence of an attribution bias, the difference in itself does not indicate whether the bias is due to self-serving attributions. To identify the nature of the bias, we compared the budget participants' pseudo-participation perceptions with the pseudo-participation perceptions by observers. Since observers have no stake in the budget outcome, their pseudo-participation perceptions should provide a standard for assessing whether budget participants made self-serving pseudo-participation perceptions. For the unfavorable budget condition, pseudo-participation perceptions by budget participants should exceed pseudo-participation perceptions by observers; for the favorable budget condition, pseudo-participation perceptions by observers should exceed the pseudo-participation perceptions by budget participants. Hence, we hypothesize the following ordinal interaction between budget favorability and subject group for the non-disclosed budget intention condition:

H3. For the non-disclosed budget intention condition, budget participants will perceive more pseudo-participation than observers for the unfavorable budget condition, and will perceive less pseudo-participation than observers for the favorable budget condition.

METHOD

Hypotheses testing involved the manipulation of budget favorability (unfavorable budget vs. favorable budget) and manipulation of the disclosure of the decision-maker's budget intention (disclosed budget intention vs. non-disclosed budget intention) in a laboratory experiment. Budget participants participated directly in budget setting by submitting a requested budget in terms of a performance target to a budget decision-maker (referred to in the experiment as a supervisor). The performance target related to the amount of work that should be accomplished on a clerical task in a specified period of time. The budget decision-maker set a final budget in terms of a performance target that was either easier (favorable budget) or more difficult (unfavorable budget) than the budget participant's requested target. After receiving the final budget, budget participants rated the degree to which the budget process was pseudo-participative. Observers did not participate in the budget setting activities. They watched a reenactment of the process on videotape, after which they provided pseudo-participation perceptions.

Budget Participants and Observers

The experiment included 96 undergraduate business students who were taking or had completed a managerial accounting class. Females comprised 52 percent of the sample. The mean age of the students was 22 years (SD = 3.9 years). Students received extra course credit for their participation in the study. We randomly formed dyads containing a budget participant and an observer and randomly assigned dyads to experimental treatments. The effectiveness of randomization was checked by testing for differences in age, sex, educational background, and performance ability (performance on a practice task) between experimental treatments. None of the tests was significant (each $p > .05$), indicating that study participants on an average had similar individual-difference characteristics among treatments.

Performance Budget and Performance Task

Budgets related to performance targets for a clerical task. The clerical task was selected from Libby (1999; also see Chow, 1983). It consisted of applying a decoding key to translate symbols into alphabetic characters over a 5-min period. Performance was measured as the number of correctly decoded symbols. An advantage of the task is that it is easy to understand and there is sufficient range in people's ability to perform the task to allow flexibility in setting performance budgets (Libby, 1999).

Budget favorability was manipulated experimentally by setting either a difficult performance target (unfavorable budget) or an easy performance target (favorable budget). In order to ensure that the performance target deviated from expectations, the target was set at a level that differed substantially from the target requested by the budget participant. The unfavorable budget was set at 200 correctly decoded symbols in 5 min. Libby (1999) used this same budget level to represent an unfavorable budget. She described this budget level as sufficiently difficult for study participants to consider it as unfavorable. The favorable budget was set at 85 correctly decoded symbols in 5 min, which is approximately two standard deviations below the average level of performance as determined in a pilot test.

Dependent Variable: Perceived Pseudo-Participation

Perceived pseudo-participation was measured with two items (summed): (a) "Did the supervisor intend to use your (the student/employee's) budget input?" and (b) "Was the supervisor's request for your (the student/employee's) budget input genuine?" (response scale for each item: 1 = "not at all" and 11 = "definitely"). The items were reverse coded so that higher values indicated more perceived pseudo-participation.

Control Variable: Perceived Influence

Perceived influence was added to the study as a control variable. Perceptions of pseudo-participation depend on perceptions of low influence. While we designed the experiment to minimize disparity in the level of perceived budget influence among the experimental conditions, it is unlikely that we were able to remove such disparity completely. Therefore, we used statistical control to equalize the experimental conditions with regard to perceived influence, which was measured with two items (summed): (a) "How much influence do you feel you (the student/employee) had over the final budget?" (1 = "no influence" and 11 = "a great deal of influence") and (b) "How important do you think your (the student/employee's) contribution was to setting the budget?" (1 = "not at all important" and 11 = "very important").

Experimental Procedures for Budget Participants

Experimental sessions were held over the course of two weeks, with four to eight students attending each session. Sessions involving budget participants preceded sessions involving observers. The two manipulations related to the disclosure of budget intention factor were conducted in separate sessions. In the non-disclosed budget intention condition, 24 budget participants received an unfavorable budget and 26 received a favorable budget; in the disclosed budget intention condition, 22 budget participants received an unfavorable budget and 24 received a favorable budget. The experimental sessions were held in a room containing nine cubicles, each with a chair and a desk. Each cubicle was set up so that the budget participant could see the experimenter and the budget decision-maker (an experimental accomplice), but could not see budget participants in other cubicles. Experimental procedures were conducted in four steps. Scripts were prepared in advance to guide and standardize verbalizations by the experimenter and the budget decision-maker during experimental sessions.

In step 1 (introduction and practice), the experimenter greeted budget participants as they arrived at a session, assigned each a cubicle, and provided a general overview of their participation in the experiment. The budget decision-maker, who was a student of a similar age, arrived with the budget participants. After receiving written and oral instructions, budget participants (including the budget decision-maker) performed the clerical task in a 5-min practice session to acquaint them with the task and to provide them with insight about their ability to perform the task. As a work incentive, budget participants earned one raffle ticket for every 25 symbols they correctly decoded during the practice session. The raffle tickets provided opportunities to participate in two $100 drawings. Budget participants graded their own work.

In step 2 (compensation scheme explanation), budget participants were presented with a compensation scheme that involved earning raffle tickets for the $100 drawings. Their compensation was based on the following scheme:

(a) They earned two raffle tickets for every 25 symbols they correctly decoded.
(b) They earned a 20-ticket bonus for meeting or exceeding the performance target they received.
(c) They lost one ticket for each increment of 25 symbols that their performance fell short of their performance target.

A truth-inducing compensation scheme (Lindquist, 1995) was not used because it precluded the setting of a favorable budget.

After explaining the raffle ticket compensation scheme and its relationship to the performance target, the experimenter selected a person to serve as a work supervisor. The person selected was an experimental accomplice, who served as the budget decision-maker in all experimental sessions. As a supervisor, the budget decision-maker's responsibility was to set a budget in terms of a performance target for each budget participant. To further the deception, the experimenter explained the budget decision-maker's compensation scheme in the presence of budget participants. The budget decision-maker's pay was two raffle tickets for every 100 symbols that his work crew (participants in the session) correctly decoded and a 20-ticket bonus if the work crew correctly decoded more symbols than other work crews. After explaining the budget decision-maker's compensation scheme, the experimenter announced that she was leaving the room to allow the budget decision-maker to set budgets and told the budget decision-maker to come out and get her when he had completed his task. The experimenter then left the room.

In step 3 (budget setting), the budget decision-maker asked budget participants to assist him in setting their budgets. He distributed slips of paper to budget participants and asked them to write down a requested performance budget in terms of correctly decoded symbols. He then collected the slips of paper. It was during this time that the experimental manipulations took place. In the non-disclosed budget intention condition, the budget decision-maker looked at each request and, after a few minutes of contemplation, wrote final performance targets on slips of paper, assigning either a favorable budget (85 correctly decoded symbols) or an unfavorable budget (200 correctly decoded symbols) depending on the budget favorability manipulation. He then distributed final performance targets to budget participants. Once the budget decision-maker finished setting and distributing final performance targets, he left the room, ostensibly to get the experimenter.

In the disclosed budget intention condition, the supervisor wrote down a preliminary budget for each budget participant in terms of a performance target. These targets were pre-planned amounts that reflected the budget favorability manipulation and were written down at the same time as budget participants were writing down their budget requests. Once the budget decision-maker had written a preliminary performance target on each slip of paper, he distributed the targets to each budget participant as he collected their requested targets. This procedure ensured that budget participants did not see the budget decision-maker's preliminary budget before they decided on their requested budget and the budget decision-maker did not see the budget participants' requested budget before he decided on their preliminary budget. After this exchange of the budget decision-maker's preliminary budgets and the budget participants' requested budgets, the budget decision-maker pretended to think about the budgets to set for participants (as in the non-disclosed budget intention condition) and then distributed the final budgets. In the disclosed budget intention condition, the final budget contained the same performance target as the preliminary budget.

In step 4 (perceived pseudo-participation measurement), the experimenter returned and administered the pseudo-participation questions. We undertook the following steps to project the appearance that the pseudo-participation questioning was an unplanned event. The purpose of these steps was to counter hypothesis guessing in the measurement of perceived pseudo-participation. The experimenter distributed several yes-or-no questions pertaining to the budget formulation process. Questions included, "Did the supervisor leave the room while setting your budget?" and "Did the supervisor ask you for your budget suggestion?". The experimenter collected the answers and, after pretending to review them and feigning perplexity over their contents, announced that based on the review, she would distribute another set of questions. This new set contained the pseudo-participation questions. Budget participants next performed the decoding task and completed a post-experimental questionnaire. Perceived influence was measured in the post-experimental questionnaire. After the experimental sessions were completed, the experimenter conducted the raffle and thoroughly debriefed budget participants regarding the purpose of the experiment and the manipulation deceptions.

Experimental Procedures for Observers

The procedures for observers were the same as those for budget participants up through the practice decoding session. After the practice session, observers watched a video recording that contained a reenactment of the experimenter's and budget decision-maker's actions previously described in step 2 (compensation scheme explanation) and step 3 (budget setting). There were two versions of the video recording, one showing the budget decision-maker giving out both a preliminary budget and a final budget (the disclosed budget intention condition) and the other showing the budget decision-maker giving out only a final budget (the non-disclosed budget intention condition). After watching the video recording, each observer was shown the requested budget and final budget relating to his or her paired budget participant, and, in the disclosed intention condition, was shown the budget decision-maker's preliminary budget for that budget participant. Next, observers answered the pseudo-participation questions and completed the post-experimental questionnaire. They received 30 raffle tickets for performing the evaluation task, which entitled them to participate in two \$100 raffles. Budget participants and observers participated in separate raffles.

RESULTS

Manipulation and Validity Checks

Several tests were conducted to check the effectiveness of treatment manipulations. Two items (1–11 response scales) were included in the post-questionnaire to test the effectiveness of the budget favorability manipulation (summed; inter-item correlation of .63): (a) "How difficult was your (the student/employee's) budget to attain?" and (b) "How favorable was your (the student/employee's) budget for you (them)?". The mean budget favorability score (budget participants and observers combined) was 18.44 for the favorable budget condition and 12.86 for the unfavorable budget condition. This difference was significant ($p < .001$), which supports the effectiveness of the budget favorability manipulation. Work performances in the practice session substantiated the difficulty of the budget levels set for a favorable budget (85 correctly decoded symbols) and an unfavorable budget (200 correctly decoded symbols). In the practice session, approximately four percent of budget participants and observers combined decoded more than 200 symbols correctly and approximately 97 percent decoded more than 85 symbols correctly. The effectiveness of the manipulation pertaining to the disclosure of budget intention factor was checked by asking whether the supervisor distributed a preliminary budget. Eighty-three percent of the students provided a correct answer, which supports the effectiveness of the manipulation.

We found support for the internal and construct validities of the perceived pseudo-participation and perceived influence scales. The two items making up each scale had a high inter-item correlation (.72 for perceived pseudo-participation and .74 for perceived influence), and each scale had a high pairwise correlation (–.52 for pseudo-participation and .52 for influence) with a procedural fairness item that asked, "How fair was the procedure used to set the budget?". Perceived pseudo-participation and perceived influence had a pairwise correlation of –.50 ($p < .001$).

Testing of Perceived Influence

The first set of tests examined whether perceived influence varied among the experimental treatments. Results of maximum-likelihood ANOVA for perceived influence are reported in panel A of Table 1, and the mean perceived influence for the two (subject group) by two (disclosure of budget intention) by two (budget favorability) design are reported in panel B of Table 1. ANOVA estimates were computed with a maximum likelihood procedure because each observer viewed the budget process pertaining to a paired budget participant, thus producing correlated observations. The MIXED procedure of SAS® was used for the analysis. Main effects relating to the disclosure of budget intention factor ($F = 7.87$, $p = .005$) and the

Table 1. ANOVA Results for Perceived Influence.

Source	*F*-Statistic (df = 1.44)	Significance
Panel A: Results of maximum-likelihood ANOVA[a]		
Subject group (G)	.01	.921
Disclosure of budget intention (D)	7.87	.005
Budget favorability (F)	4.14	.048
G∗D	3.69	.061
G∗F	1.59	.214
D∗F	.02	.899
G∗D∗F	.14	.712

Disclosure of Budget Intention	Budget Favorability		Totals
	Unfavorable budget	Favorable budget	
Panel B: Mean perceived influence for budget participants and observers[b]			
Non-disclosed budget intention	Observer: 11.7 Participant: 12.1 ($n = 24$)	Observer: 12.1 Participant: 16.0 ($n = 26$)	Observer: 11.9 Participant: 14.1 ($n = 50$)
Disclosed budget intention	Observer: 10.0 Participant: 7.1 ($n = 22$)	Observer: 11.5 Participant: 10.5 ($n = 24$)	Observer: 10.8 Participant: 8.9 ($n = 46$)
Totals	Observer: 10.9 Participant: 9.7 ($n = 46$)	Observer: 11.8 Participant: 13.4 ($n = 50$)	Observer: 11.4 Participant: 11.6 ($n = 96$)
Panel C: Mean perceived influence for budget favorability and disclosure of budget intention			
Non-disclosed budget intention	11.9	14.1	13.0
Disclosed budget intention	8.6	11.0	9.8
Totals	10.3	12.6	11.4

[a]A maximum-likelihood procedure was appropriate for analysis because each observer viewed the budget process of a paired budget participant. This situation created correlated observations. The analysis was performed with the MIXED procedure of SAS®. R^2, computed with an OLS procedure, was .17 (adjusted $R^2 = .10$).

[b]"Participant" indicates budget participants; "Observer" indicates observers. The perceived influence scale ranged from 2 to 22, with higher numbers indicating greater perceived influence. In the disclosed budget intention condition, the final budget exactly matched the budget decision-maker's disclosed preliminary budget.

budget favorability factor ($F = 4.14$, $p = .048$) were statistically significant. The means for the two main effects are reported in panel C of Table 1. The mean level of perceived influence for those in the non-disclosed budget intention condition was higher than the mean for those in the disclosed intention condition (13.0 vs. 9.8). This difference indicates that the budget decision-maker's disclosure of his budget intention contained information that suggested reduced influence. The mean for budget participants who received an unfavorable budget (10.3) was lower than the mean for budget participants who received a favorable budget (12.6). This difference indicates that the budget levels used in the manipulation of budget favorability may have produced different perceptions of influence. Thus, the analysis indicates that the experimental manipulations created different perceptions of influence, confirming the need to include perceived influence as a control variable when examining the effects of the manipulations on perceived pseudo-participation.

Testing of Perceived Pseudo-Participation

Table 2 provides results for assessing the three hypotheses of the study. Panel A of Table 2 contains results of maximum-likelihood ANCOVA for perceived pseudo-participation, with perceived influence added as a covariate to isolate the effects of budget favorability and disclosure of budget intention after controlling for perceived influence. Panel B contains mean perceived pseudo-participation for each cell of the $2 \times 2 \times 2$ experimental design.

Two experimental effects were statistically significant, a subject group main effect ($F = 7.28$, $p = .010$) and the effect of the interaction between the subject group factor and the budget favorability factor ($F = 4.52$, $p = .039$). Cell means pertaining to the interaction between the subject group factor and budget favorability factor are reported in panel C of Table 2. Two sets of means are reported, one for the raw perceived pseudo-participation scores and, in parentheses, one adjusted for the effect of perceived influence. An ordinal interaction is indicated. Budget participants' mean perceived pseudo-participation for the favorable budget condition was noticeably less than the mean for the unfavorable budget condition and noticeably less than mean perceived pseudo-participation by observers.[1]

The disclosure of budget intention factor included a disclosed budget intention condition that was intended to evoke strong pseudo-participation perceptions. The purpose of including this condition in the experiment was to assess whether self-serving perceptions of pseudo-participation occurred when budget participants had convincing information about the budget decision-maker's sincerity. We found no significant main or interactive effects for the disclosure of budget intention factor. These insignificant

Table 2. ANCOVA Results for Perceived Pseudo-Participation.

Panel A: Results of maximum-likelihood ANCOVA[a]

Source	*F*-Statistic (df = 1.43)	Significance
Perceived influence (covariate)	18.79	<.001
Subject group (G)	7.28	.010
Disclosure of budget intention (D)	2.14	.150
Budget favorability (F)	2.94	.094
G∗D	.32	.575
G∗F	4.52	.039
D∗F	.92	.344
G∗D∗F	1.76	.192

Panel B: Mean perceived pseudo-participation for budget participants and observers[b]

Disclosure of Budget Intention	Budget Favorability		Totals
	Unfavorable budget	Favorable budget	
Non-disclosed budget intention	Observer: 13.8 Participant: 11.5 ($n = 24$)	Observer: 12.0 Participant: 6.9 ($n = 26$)	Observer: 12.9 Participant: 9.1 ($n = 50$)
Disclosed budget intention	Observer: 13.3 Participant: 15.5 ($n = 22$)	Observer: 15.1 Participant: 10.5 ($n = 24$)	Observer: 14.2 Participant: 12.9 ($n = 46$)
Totals	Observer: 13.6 Participant: 13.4 ($n = 46$)	Observer: 13.5 Participant: 8.6 ($n = 50$)	Observer: 13.5 Participant: 11.1 ($n = 96$)

Panel C: Mean perceived pseudo-participation for subject group and budget favorability, with (without) adjustment for influence

Budget Favorability	Observers	Budget Participants
Unfavorable budget	13.6 (1.05) ($n = 46$)	13.4 (.38) ($n = 46$)
Favorable budget	13.5 (1.39) ($n = 50$)	8.6 (–2.73) ($n = 50$)

[a]A maximum-likelihood procedure was appropriate for analysis because each observer viewed the budget process of a paired budget participant. This situation created correlated observations. The analysis was performed with the MIXED procedure of SAS®. R^2, computed with an OLS procedure, was .37 (adjusted $R^2 = .33$).

[b]"Participant" indicates budget participants; "Observer" indicates observers. The perceived pseudo-participation scale ranged from 2 to 22, with higher numbers indicating greater perceived pseudo-participation. In the disclosed budget intention condition, the final budget exactly matched the budget decision-maker's disclosed preliminary budget.

results indicate that the self-serving perceptions of pseudo-participation that occurred for the favorable budget condition were unaffected by prior knowledge of the decision-maker's budget intention. The self-serving inclination to discount pseudo-participation as the cause of a favorable budget may have been strong enough to override other information that suggests the decision-maker was insincere in soliciting budget input.

DISCUSSION OF RESULTS

H3 predicts that budget participants will form self-serving perceptions of pseudo-participation in situations of low budget influence. When they receive an unfavorable budget, budget participants will inflate pseudo-participation perceptions to discount its cause away from inadequacies of self. When they receive a favorable budget, budget participants will deflate pseudo-participation perceptions to establish positive attributes of self as the cause. The results support self-serving perceptions of pseudo-participation for a favorable budget but not for an unfavorable budget. The strength of pseudo-participation perceptions that derive from an unfavorable budget may explain why self-serving perceptions do not occur for an unfavorable budget. In situations of low influence, people may have a natural tendency to consider budgetary involvement as pseudo-participative when they receive an unfavorable budget. Low budget influence and budget decision-maker insincerity may be bound perceptions because insincerity is a causal explanation for a budgetary outcome that did not reflect the budget participant's input. For instance, research has shown that people who have knowledge of a causal relationship, say X causes Y, frequently inflate the likelihood of X existing when Y occurs, even though Y does not cause X (Mandel & Lehman, 1998). If influence and sincerity perceptions are linked in this manner, budgets that suggest low influence may evoke strong pseudo-participation perceptions even when there is no evidence to support decision-maker insincerity other than the decision outcome. Strong perceptions may severely limit the returns that budget participants can realize from biasing perceptions of pseudo-participation upward to deflect blame for the unfavorable budget away from self. The budget process is already considered to be pseudo-participative, thus rendering the perception bias superfluous.

The results did not support either H1 or H2. The argument underlying H1 and H2 was that the decision-maker's prior disclosure of a budget intention, operationalized in the experiment as the disclosure of a preliminary budget that equaled the final budget, would provide persuasive evidence of pseudo-participation. The persuasiveness of the evidence would prompt both budget participants and observers to perceive pseudo-participation as the most viable causal factor for low budget influence and would deter budget participants from biasing their pseudo-participation perceptions. On the basis of this argument, H1 posited higher perceived pseudo-participation when the budget decision-maker disclosed a budget intention than when the budget decision-maker did not disclose a budget intention, and H2 posited that the perceived pseudo-participation of budget participants would be similar to that of observers when the decision-maker disclosed a budget intention. In addition, H3, which addressed the existence of biased pseudo-participation perceptions, included a conditional clause that delimited the hypothesized bias to situations when the budget decision-maker did not disclose a budget intention. Contrary to the theorizations underlying all three hypotheses, pseudo-participation perceptions of budget participants differed little from those of observers irrespective of whether or not the decision-maker disclosed a budget intention. Since the disclosure of a budget intention was intended to evoke a strong perception of pseudo-participation, this result suggests that low budget influence may at times provide convincing evidence of pseudo-participation. An additional finding was that budget participants' biased perceptions of pseudo-participation for the favorable budget condition occurred even when they knew the decision-maker's preliminary budget. This result suggests that budget participants possess strong inclinations to bias attributions of a favorable budget outcome in order to identify self, and not a pseudo-participative process, as the agent responsible for the favorability of the outcome.

IMPLICATIONS, LIMITATIONS, AND FUTURE RESEARCH

In participative budgeting, budget decision-makers solicit input from employees who have a stake in the budget outcome. Sincerity of the budget decision-maker's request for input is considered to be an essential element for effective budget participation (e.g., Argyris, 1952; Becker & Green, 1962). Our study's results suggest that there may be times when it is difficult for budget decision-makers to avoid perceptions of pseudo-participation, even when the decision-maker acted with sincerity. Budget participants have a perceptual tendency to associate pseudo-participation with an unfavorable budget. Budget decision-makers may frequently encounter situations in which they cannot issue budgets that meet participants' budgetary requests. For particularly unfavorable budgets, budget participants may tend to perceive requests for budget input as pseudo-participation even though they possess little information other than the budget outcome for assessing the budget decision-maker's sincerity.

The preceding concerns relating to budget participants' attitudinal dispositions may extend to employees other than budget participants. Dispositions to associate pseudo-participation with low influence may generalize to all people who observe the budget process, not just to budget participants. Observers of the budget process, which can include other subordinates of the budget decision-maker or the budget decision-maker's superiors or peers, may have a tendency to regard the budget decision-maker's solicitations for budget input as insincere whenever they perceive a mismatch between budget inputs and budget outcomes. The tendency to associate pseudo-participation with low influence may contribute to observers' perceptions of unfairness (Cohen, 1985), which may impede the budget decision-maker's efforts to build or maintain intra-organizational trust relationships (Greenberg & Folger, 1985).

Our study is among the first to examine a budgetary situation where budget participants receive a final budget that is more favorable than the requested budget. The findings show that budget participants may form very different perceptions of pseudo-participation than observers when budget participants receive a favorable budget. Budget participants have a self-serving tendency to discount pseudo-participation as a viable cause of favorable (better than expected) budgets. Thus, budget decision-makers may experience puzzling situations in which observers regard the budget process leading to a favorable budget as unfair while budget participants deem the process as fair.

People's tendency to attribute their low budget influence to the budget decision-maker's insincerity has implications for budgetary participation researchers. First, the use of budget influence to manipulate participation in budget experiments (e.g., Brownell, 1981; Clinton, Hall, Hunton, & Pierce, 1996) may produce changes in perceptions of pseudo-participation in conjunction with changes in perceptions of influence. An alternative method for manipulating budgetary participation that avoids the possibility of also affecting pseudo-participation perceptions is to change perceptions of involvement, such as manipulating participation as input vs. no input (e.g., Tiller, 1983; Chalos & Haka, 1989; Kren, 1990; Lindquist, 1995; Libby, 1999). Second, budgetary participation instruments (e.g., Milani's 1975 scale and its derivatives) that include questions involving budget influence may confound the concept of involvement with the concept of pseudo-participation. Participation studies that use these instruments may be unable to resolve whether organizational attitudes and behaviors are the consequence of pseudo-participation perceptions, budget involvement perceptions, or both.

This study must be evaluated in light of its limitations. First, observers and budget participants did not experience the exact same experimental conditions. Observers watched a videotape that contained a reenactment of the participative process. It is possible that viewing a reenactment of budgetary events on videotape may yield different perceptions than do first hand experiences. Second, budgets were stated in terms of performance targets. A potential explanation for an unfavorable performance budget is that the budget decision-maker selected a challenging target to motivate budget participants to perform at a higher level (e.g., Locke, 1996, 2000). Budget participants or observers may have included this explanation as a potential cause for the unfavorable budget outcome. If they did, they may have assigned less weight to pseudo-participation as the cause of an unfavorable budget. Third, as a reviewer pointed out, the participative process in the experiment may have represented an oversimplification of pseudo-participation processes in organizations. We modeled a situation in which a team leader had the freedom to invite team members to participate in the budget setting process. In the invitation, the budget decision-maker asked budget participants to assist him in setting the budget. However, organizations often mandate budgetary participation, and the budget setting process often involves explicit or tacit bargaining and negotiations (Fisher, Frederickson, & Peffer, 2000). A process of bargaining and negotiation may evoke different expectations and thus produce different results than those found in our study. Finally, all budget participants in a given budget favorability condition were assigned the same final budget. The final budget was farther from the requested budget for some budget participants than for others. This approach might have resulted in a lower average level of perceived pseudo-participation than would have existed if the magnitude of the distance between the requested budget and the final budget had been the same for each budget participant.

Our results suggest several avenues of future research. First, research needs to identify steps that budget decision-makers can employ in instances of low budget influence to avoid, or at least minimize, perceptions of pseudo-participation. A possible counteraction is to provide convincing explanations for why budget participants were assigned either unfavorable or favorable budgets (Bies & Shapiro, 1988; Greenberg, 1991). Libby's (1999) study suggests that budget participants may be less inclined to associate pseudo-participation with low influence when the budget outcome is accompanied by an explanation, although she did not study the effect of explanations on observers. Second, research needs to assess whether familiarity with budget decision-makers affects budget participants' and observers' tendency to attribute low influence to pseudo-participation. In our study, budget participants and observers knew neither the budget decision-maker nor one another. It is possible that familiarity breeds trust, which may moderate pseudo-participation perceptions. Third, research needs to address whether budget participants' access to information that suggests alternative causes for low budget influence moderates pseudo-participation perceptions. In our study, budget participants and observers had access only to the final budget and the budget decision-maker's preliminary budget on which to judge the decision-maker's sincerity. It is possible that knowledge of situational factors, such as the company's budgetary practices or its financial health, may provide alternative causal explanations for low influence that negate tendencies to attribute low influence to pseudo-participation.

Budget researchers have not given much empirical attention to pseudo-participation, even though it is recognized as an important determinant of attitudes toward the organization and its budget decision-makers. We surmise that this scarcity is due to people's general belief that budget decision-makers can avoid pseudo-participative perceptions simply by avoiding disingenuous requests for input. However, budget participants and observers have to infer sincerity of input requests from perceptions of the budget decision-maker and the events that unfolded during the budget formulation process. As our study shows, people possess attitudinal dispositions and biases that can create mismatches between perception and reality. Thus, our study clearly demonstrates the need for more study of pseudo-participation.

NOTE

1. We performed two post-hoc tests to confirm the form of the interaction. First, we compared budget participants' mean perceived pseudo-participation for the two budget favorability conditions to confirm that the means were significantly different. The difference between the means for the two budget favorability conditions (8.6 for the favorable budget condition vs. 13.4 for the unfavorable budget condition, collapsed as to subject group and disclosure of budget intention) was significant

(Duncan and Scheffe tests, $p<.05$). Second, we excluded cells that involved a favorable budget and tested for differences in means among the remaining cells. Insignificant differences among the remaining cells provide evidence of an ordinal interaction. The difference between the highest mean (15.5 for the "budget participant, disclosed budget intention, unfavorable budget" condition) and the lowest mean (11.5 for "budget participant, non-disclosed budget intention, unfavorable budget" condition) was not significant (Duncan and Scheffe tests, $p>.05$). The results of post-hoc testing indicated an ordinal interaction between subject group and budget favorability: budget participants perceived less pseudo-participation than observers for the favorable budget condition, but about the same level of pseudo-participation as observers for the unfavorable budget condition.

REFERENCES

Anderson, D. Y., & Hayes, C. L. (1996). *Gender, identity and self-esteem*. New York, NY: Springer Publishing.

Aranya, N. (1990). Budget instrumentality, participation and organizational effectiveness. *Journal of Management Accounting Research*, *2*(Fall), 67–77.

Argyris, C. (1952). *The impact of budgets on people*. Ithaca, NY: Controllership Foundation.

Barsky, N. P. (1999). A core/periphery structure in a corporate budgeting process. *Connections*, *22*(2), 22–29.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*, 3–26.

Baumeister, R. F. (1989). Motives and costs of self-presentations in organizations. In: R. A. Giacalone & P. Rosenfeld (Eds), *Impression management in the organization* (pp. 57–72). Hillsdale, NJ: Erlbaum.

Becker, S., & Green, D. Jr. (1962). Budgeting and employee behavior. *The Journal of Business*, *35*, 392–402.

Bies, R. J., & Shapiro, D. L. (1988). Voice and justification: Their influence on procedural fairness judgments. *Academy of Management Journal*, *31*, 676–685.

Blader, S. L., & Bobocel, D. R. (2005). Wanting is believing: Understanding psychological processes in organizational justice by examining perceptions of fairness. In: S. Gillihand, D. Steiner & K. van den Bos (Eds), *Research in social issues in management: What motivates fairness in organizations?* (pp. 3–30). Greenwich, CT: Information Age Publishing.

Brockner, J., Heuer, L., Magner, N., Folger, R., Umphress, E., van den Bos, K., Vermunt, R., Magner, M., & Siegel, P. (2002). High procedural fairness heightens the effect of outcome favorability on self-evaluations: An attributional analysis. *Organizational Behavior and Human Decision Processes*, *91*, 51–68.

Brownell, P. (1981). Participation in budgeting, locus of control and organizational effectiveness. *The Accounting Review*, *56*(4), 844–860.

Brownell, P. (1982). Participation in the budgeting process: When it works and when it doesn't. *Journal of Accounting Literature*, *1*, 124–153.

Brownell, P., & McInnes, M. (1986). Budgetary participation, motivation, and managerial performance. *The Accounting Review*, *61*(4), 587–600.

Burris, C. T., Harmon-Jones, E., & Tarpley, W. R. (1997). By faith alone: Religious agitation and cognitive dissonance. *Basic and Applied Social Psychology*, *19*, 17–31.

Chalos, P., & Haka, S. (1989). Participative budgeting and managerial performance. *Decision Sciences*, *20*, 334–347.

Chalos, P., & Poon, M. (2001). Participative budgeting and performance: A state of the art review and re-analysis. *Advances in Management Accounting*, *10*, 171–201.

Chow, C. W. (1983). The effects of job standard tightness and compensation scheme on performance: An exploration of linkages. *The Accounting Review*, *158*(4), 667–685.

Clinton, D. B., Hall, T. W., Hunton, J. E., & Pierce, B. J. (1996). Perceptions of supervisor/subordinate task-specific expertise as a moderator of participative budgeting outcomes. *Advances in Accounting*, *14*, 85–106.

Cohen, R. L. (1985). Procedural justice and participation. *Human Relations*, *38*, 643–663.

Dunk, A. S. (1992). The effects of managerial level on the relationship between budgetary participation and job satisfaction. *British Accounting Review*, *24*, 207–218.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fisher, J. G., Frederickson, J. R., & Peffer, S. A. (2000). Budgeting: An experimental investigation of the effects of negotiation. *The Accounting Review*, *75*(1), 93–114.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York, NY: McGraw Hill.

Foran, F., & DeCoster, D. T. (1974). An experimental study of the effects of participation, authoritarianism, and feedback on cognitive dissonance in a standard setting situation. *The Accounting Review*, *49*(4), 264–277.

Frucot, V., & Shearon, W. T. (1991). Budgetary participation, locus of control, and Mexican managerial performance and job satisfaction. *The Accounting Review*, *66*(1), 80–99.

Govindarajan, V. (1986). Impact of participation in the budgetary process on managerial attitudes and performance: Universalistic and contingency perspectives. *Decision Sciences*, *17*, 496–516.

Greenberg, J. (1981). The justice of distributing scarce and abundant resources. In: M. J. Lerner & S. C. Lerner (Eds), *The justice motive in social behavior: Adapting to times of scarcity and change* (pp. 289–316). New York, NY: Plenum Press.

Greenberg, J. (1991). Using explanations to manage impressions of performance appraisal fairness. *Employee Responsibilities and Rights Journal*, *4*, 51–60.

Greenberg, J., & Folger, R. (1985). Procedural justice, participation, and the fair process effect in groups and organizations. In: P. B. Paulus (Ed.), *Basic group processes* (pp. 235–256). New York, NY: Springer-Verlag.

Greenberg, P. S., Greenberg, R. H., & Houri, H. (1994). Participative budgeting: A meta-analytic examination of methodological moderators. *Journal of Accounting Literature*, *13*, 117–141.

Griffin, L. (1996). The effects of participatory budgeting on both the attractiveness of budgetary compliance and motivation. *Advances in Management Accounting*, *5*, 169–187.

Grover, S. L. (1991). Predicting the perceived fairness of parental leave policies. *Journal of Applied Psychology*, *76*(2), 247–255.

Harmon-Jones, E., & Mills, J. (1999). *Cognitive dissonance: Progress on a pivotal theory in social psychology*. Washington, DC: American Psychology Association.

Hofstede, G. (1967). *The game of budget control*. Assen, The Netherlands: Van Gorcum.

Kelley, H. H. (1972). Attribution in social interaction. In: E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins & B. Weiner (Eds), *Attribution: Perceiving the causes of behavior* (pp. 1–26). Morristown, NJ: General Learning Press.

Kren, L. (1990). Performance in a budget-based control system: An extended expectancy theory model approach. *Journal of Management Accounting Research*, *2*(Fall), 100–112.

Libby, T. (1999). The influence of voice and explanation on performance in a participative budgeting setting. *Accounting, Organizations and Society*, *24*, 125–137.

Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York, NY: Plenum Press.

Lindquist, T. M. (1995). Fairness as an antecedent to participative budgeting: Examining the effects of distributive justice, procedural justice and referent cognitions on satisfaction and performance. *Journal of Management Accounting Research*, *7*(Fall), 122–147.

Locke, E. A. (1996). Motivation through conscious goal setting. *Applied and Preventive Psychology*, *5*, 117–124.

Locke, E. A. (2000). Motivation, cognition, and action: An analysis of studies of task goals and knowledge. *Applied Psychology: An International Review*, *49*(3), 408–429.

Magner, N., Welker, R., & Campbell, T. L. (1995). The interactive effect of budgetary participation and budget favorability on attitudes toward budgetary decision makers: A research note. *Accounting, Organizations and Society*, *20*(7/8), 611–618.

Maiga, A. S. (2005). Antecedents and consequences of budget participation. *Advances in Management Accounting*, *14*, 211–231.

Mandel, D., & Lehman, D. (1998). Integration of contingency information and judgment of cause covariation and probability. *Journal of Experimental Psychology: General*, *127*(3), 269–285.

Mia, L. (1989). The impact of participation in budgeting and job difficulty on managerial performance and work motivation: A research note. *Accounting, Organizations and Society*, *14*(4), 347–357.

Milani, K. (1975). The relationship of participation in budget-setting to industrial supervisor performance and attitudes: A field study. *The Accounting Review*, *50*(2), 275–284.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, *82*, 213–225.

Nigro, L. G., & Bellone, C. J. (1979). Participative management: Making it work. *The Bureaucrat*, *8*(4), 34–39.

Nouri, H., Kyj, L., & Dunk, A. (1999). The effect of performance reporting on budgetary participation: An attribution theory analysis. *Advances in Management Accounting*, *8*, 211–223.

Nouri, H., & Parker, R. J. (1998). The relationship between budget participation and job performance: The roles of budget adequacy and organizational commitment. *Accounting, Organizations and Society*, *23*(5), 467–483.

Pasewark, W., & Welker, R. (1990). A Vroom-Yetton evaluation of subordinate participation in budgetary decision making. *Journal of Management Accounting Research*, *2*, 113–126.

Pateman, C. (1970). *Participation and democratic theory*. London, England: Cambridge University Press.

Ployhart, R. E., & Ryan, A. M. (1997). Toward an explanation of applicant reactions: An examination of organizational justice and attribution frameworks. *Organizational Behavior and Human Decision Processes*, *72*, 308–335.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In: L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 174–221). New York, NY: Academic Press.

Schroth, H., & Shaw, P. (2000). Procedures, do we really want to know them? *Journal of Applied Psychology*, *85*, 462–471.

Searfoss, G., & Monczka, R. M. (1973). Perceived participation in the budget process and motivation to achieve the budget. *Academy of Management Journal*, *16*(December), 541–554.

Tiller, M. (1983). The dissonance model of participative budgeting: An empirical exploration. *Journal of Accounting Research*, *21*(2), 581–595.

Zuckerman, M. (1979). Attribution of success and failure revisited, or: The motivational bias is alive and well in attribution theory. *Journal of Personality*, *47*, 245–287.

TAXONOMY OF PERFORMANCE MEASUREMENT SYSTEMS

Jean-François Henri

ABSTRACT

This study aims to provide an integrated view of performance measurement systems (PMS) by developing a taxonomy reflecting the interdependencies among various PMS aspects. This study aims to move the study of PMS from a cartesian form of contingency fit to a configuration form. More specifically, the following research question is investigated in this study: To what extent do similar patterns across various dimensions of PMS occur with regularity? Using a survey approach to collect data from a sample of manufacturing firms, this study aims to develop a taxonomy based on three aspects of the PMS process, namely the design (i.e., the mix of financial, customer, internal processes, innovation and learning measures), the use (i.e., monitoring, strategic decision-making, attention-focusing, legitimization), and the revision (i.e., the addition, deletion, and changes in performance indicators). Three patterns of relationships reflecting the role and importance of PMS within the organization emerge: (a) PMS as an outcomes surveillance mechanism, (b) PMS as a management support tool, and (c) PMS as an institutionalized organizational process. This study contributes to the management accounting literature by providing a different understanding of the various levels of integration of PMS within organizational routines.

Advances in Management Accounting, Volume 17, 247–288
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017011

INTRODUCTION

This study aims to provide an integrated view of performance measurement systems (PMS) by developing a taxonomy reflecting the interdependencies among various PMS aspects. A considerable body of literature has studied PMS using a cartesian form of contingency fit (Gerdin & Greve, 2004, 2008). These studies have examined how one or a combination of contextual factors affects a single PMS aspect and how these PMS-context pairs affect performance. The studies in this line of research have mainly emphasized the aspect of the design of PMS, specifically measurement diversity. They have analyzed the effects of various organizational and individual factors, including strategy, structure, size, technology, and environmental uncertainty, on the mix of financial and nonfinancial measures (e.g., Hoque & James, 2000; Hoque, 2004; Gosselin, 2005; Said, Elnaby, & Wier, 2003; Baines & Langfield-Smith, 2003; Hyvönen, 2007; Van der Stede, Chow, & Lin, 2006; Widener, 2006; Hall, 2008).

However, by focusing mainly on the design of PMS, this line of research has assumed that specific performance indicators are used similarly and indefinitely without any revisions, which is not necessarily the case. The design constitutes one element of the performance measurement process along with the use and the revision (Bourne, Mills, Wilcox, Neely, & Platts, 2000). Other studies have examined the use of PMS (e.g., Bisbe & Otley, 2004; Simons, 2000; Henri, 2006a) or the revision of performance indicators (e.g., Bititci, Turner, & Begemann, 2000), but not with the same intensity. For the most part, past studies of PMS have considered the aspects of the design, the use, and the revision in isolation, and thus, have ignored the potential links among those aspects.

A recent line of research has combined more than one aspect of PMS. Those studies have mainly examined: (i) the association between pair of PMS aspects and (ii) the influence of one aspect of PMS on another aspect. More specifically, Ittner, Larcker, and Randall (2003b), Bhimani and Langfield-Smith (2007), and Naranjo-Gil and Hartmann (2006) suggest an association between the design and use of PMS, whereas Malina and Selto (2004) examine the association between the design and revision of PMS. Furthermore, while Cavalluzzo and Ittner (2004) and Naranjo-Gil and Hartmann (2007) observe the influence of the design of PMS on the use, Henri (2006b) documents the opposite relationship (i.e., the influence of the use of PMS on the design). Overall, little is yet known about how the design, use, and revision of PMS are related to each other (i.e., nature and direction of the relationships) and how sets of PMS share common profiles along those conceptually distinct dimensions.

A configuration form of contingency fit allows for the examination of PMS from a different perspective (Gerdin & Greve, 2004, 2008). Organizational configurations represent alignments of conceptually distinct characteristics that commonly occur together (Meyer, Tsui, & Hinings, 1993; Miller, 1996). Based principally on an empirically based taxonomy or a conceptually derived typology, configurations intend to explain why similar patterns occur with regularity across a number of variables. This study aims to move the investigation of PMS from a cartesian form of contingency fit to a configuration form. More specifically, the following research question is investigated in this study: To what extent do similar patterns across various dimensions of PMS occur with regularity? Using a survey approach to collect data from a sample of senior managers from manufacturing firms, this study aims to develop a taxonomy based on three aspects of the PMS process, namely the design (i.e., the mix of financial, customer, internal processes, innovation and learning measures), the use (i.e., monitoring, strategic decision-making, attention-focusing, legitimization), and the revision (i.e., the addition, deletion, and changes in performance indicators).

The identification and understanding of similar patterns across various dimensions of PMS is an important issue for management accounting research and practice. By considering simultaneously the aspects of design, use, and revision, this article mirrors the complexities surrounding PMS. At the same time, while recognizing the infinite number of PMS forms, the development of a classification scheme facilitates the understanding of PMS' dynamics by proposing a few common types. By performing such exercise, the analysis of PMS moves from a basic dichotomy "presence versus absence" to a continuum reflecting various degrees of integration of PMS within organizational routines. A classification scheme provides a common vocabulary to describe typical patterns of relationships among the range of organizational activities measured with performance indicators, the nature and intensity of use of those indicators, and their frequency of revisions to reflect the organization's external and internal changes. In sum, while the identification of common profiles provides a picture of the different roles and importance of PMS for organizations, the understanding of those profiles is important to explain the impacts of PMS within the organization.

The remainder of this article is organized as follows. The next section establishes the theoretical background. The following sections present the research method, the emergence of the taxonomy, and its validation. The final section presents a discussion and the conclusions of this study.

THEORETICAL BACKGROUND

Comparison of the Cartesian and Configuration Forms of Contingency Fit

A cartesian approach seeks to understand organizations by separately analyzing their constituent parts. This approach intends to explain how order is created in the parts of an organization and thus, invokes a reductionistic analysis as its dominant mode of inquiry (Meyer et al., 1993). Unidirectional and linear relationships are used to isolate the variable effects which are statistically analyzed using bivariate or sharply circumscribed multivariate analysis, such as correlation, regression, and structural equation models (Gerdin & Greve, 2004; Miller & Friesen, 1984). For the most part, a cartesian form of fit has been used to abstract a limited set of organizational concepts and to measure their relationships with a limited set of abstracted situational concepts (Meyer et al., 1993). One shortcoming of this approach is its inability to meaningfully capture the complexity of organizational reality (Ketchen & Shook, 1996).

On the other hand, a configurational approach considers that the parts of an organization take their meaning from the whole and cannot be understood in isolation. It intends to explain how order emerges from the interaction of those parts and thus, invokes a holistic synthesis as a dominant mode of inquiry (Meyer et al., 1993). Reciprocal and nonlinear relationships are used to understand the coherence between organizational elements. Statistically, cluster analysis and profile deviation analyses are used to capture the patterning of organizational elements and provide rich descriptions of configurations (Gerdin & Greve, 2004; Ketchen & Shook, 1996). Organizational configurations are not a theory in themselves, but rather they offer guidance towards the emergence of a theory. They are not intended to be exhaustive but merely illustrative of important relationships (Miller & Shamsie, 1996). In this study, we intend to take a step toward a more holistic view of PMS using a configurational approach. We aim to determine whether any meaningful patterns among PMS design, use, and revision can be observed.

Aspects of PMS

The aspects of PMS examined in this study are related to the PMS process, i.e., the design, the use, and the revision of performance indicators (Bourne et al., 2000). First, the design refers to the content of PMS in terms of measurement diversity (i.e., the mix of financial and nonfinancial performance indicators developed and included within PMS). Those indicators are then used by managers for different purposes (i.e., for monitoring, attention-focusing, decision-making, and legitimization). Lastly, following changing circumstances, performance indicators need to be reviewed as the content of PMS is reviewed. Those three aspects will now be discussed specifically in more detail.

Design of PMS

The design refers to the content of PMS in terms of measurement diversity. Measurement diversity is a broad concept that relates to various dimensions: driver versus outcome measures, subjective versus objective measures, internal versus external measures, aggregate versus specific, and financial versus nonfinancial measures (Kaplan & Norton, 1992, 1996; Ittner et al., 2003b). In this study, drawing on the work of Ittner et al. (2003b), Van der Stede et al. (2006), and Scott and Tiesen (1999), measurement diversity emphasizes the multiplicity and variety of performance indicators included within the PMS. It refers specifically to the extent to which managers collect and utilize information related to a broad set of indicators that can be grouped into financial and nonfinancial performances.

Several classifications have been proposed in the literature based on the mix of performance measures. For instance, Fitzgerald, Johnston, Brignall, Silvestro, and Voss (1991) propose a model built on the distinction between results-related measures and determinants-related measures, whereas de Haas and Kleingeld (1999) classify performance measures as ex post (reactive) and ex ante (proactive) measures. Atkinson, Waterhouse, and Wells (1997) present a stakeholder model, which includes measurement for the primary and secondary objectives of environmental and process stakeholders. Proposing a performance pyramid, Lynch and Cross (1991) present business units, core processes, and department measures. Neely, Adams, and Kennerley (2002) have developed a performance prism that contains measures of stakeholder satisfaction, operation and stakeholder contribution. Finally, probably one of the most influential classifications is the balanced scorecard developed by Kaplan and Norton (1992, 1996). In this model, three areas of performance have been added to the traditional financial dimension, namely customers, internal business process, as well as innovation and learning. In this study, the four dimensions of the balanced scorecard are used as a basic framework to define the measurement diversity dimension. This choice is motivated by the increased adoption of balanced scorecards in organizations and their use in recent empirical research (e.g., Banker, Chang, & Pizzini, 2004; Ittner, Larcker, & Meyer, 2003a; Libby, Salterio, & Webb, 2004; Malina & Selto, 2001; Lipe & Salterio, 2002).

Use of PMS

Unlike the diversity of measurement that refers to the mix of financial and nonfinancial measures included within PMS, the use of PMS is defined as the way those measures are used by managers. It refers to the nature and purpose of the use of performance indicators by managers. Previous research has classified the use of accounting and management control systems from different perspectives. These classifications suggest that PMS are used in several ways in an organizational setting. For example, Simon, Guetzkow, Kozmetsky, and Tyndall (1954) classify accounting information use as scorecard, problem-solving, and attention-directing. Burchell, Clubb, Hopwood, and Hughes (1980) present four roles of accounting practices, namely answering machine, learning machine, ammunition machine, and rationalization machine, whereas Boland and Pondy (1983) refer to rational and natural uses. Simons (1995) classifies control systems into four levers (diagnostic, interactive, beliefs, and boundaries), whereas Atkinson et al. (1997) present three uses of PMS (monitoring, diagnostic, and coordination). The uses of management information systems are grouped by Vandenbosch (1999) into four categories, namely score-keeping, problem-solving, attention-focusing, and legitimizing. Finally, Simons (2000) classifies the use of management control systems as decision-making, control, signaling, education and learning, and external communication. More recently, Ittner et al. (2003b) present four uses of strategic PMS, namely problem identification, capital investment, performance evaluation, and external disclosure, whereas Ahrens and Chapman (2004) classify the use of management control systems as coercive or enabling. From the overlap between those classifications, four main uses are reflected and examined in this study: (i) monitoring, (ii) strategic decision-making, (iii) attention-focusing, and (iv) legitimization.

Monitoring use refers to the formal feedback systems used to monitor and coordinate the implementation of plans and achievement of organizational goals (Simons, 1990). It relies on a cybernetic logic whereby goals are set in advance, output is measured, goals and output are compared, feedback is provided, and corrections are made when necessary (Hofstede, 1978). The information gathered is used for reporting and external disclosure. Acting as a diagnostic control (Simons, 1990) and answer machine (Burchell et al., 1980), PMS are associated with the measurement and reporting of performance in meeting stakeholders' requirements (Atkinson et al., 1997). *Strategic decision-making* use refers to the information systems used to support decision-making (Simon et al., 1954). By revealing cause-and-effect relationships between internal processes and objectives achievement (Atkinson et al., 1997), PMS are used in strategic decision-making as a learning machine (Burchell et al., 1980) and as a problem-solving tool (Vandenbosch, 1999).

Attention-focusing use refers to the signaling systems used to focus organizational attention and force dialogue throughout the organization (Simons, 1990). PMS act as an interactive control (Simons, 1995) and as an ammunition machine (Burchell et al., 1980) that promote specific positions and reflect one particular conception of the organizational mission. The signals sent indicate the primary and secondary objectives which employees should be focusing their attention on (Atkinson et al., 1997; Vandenbosch, 1999). Lastly, *legitimization* use refers to the justification and validation of past, current, and future actions and decisions (Ansari & Euske, 1987). This role is expressed by Burchell et al. (1980) in terms of a "rationalization machine" where there is often a need for a retrospective understanding of the emergence of an action (Feldman & March, 1981). By providing the impact in terms of performance, accounting and control systems are used to enhance the legitimacy of organizational activities (Markus, 1983). They have the capacity to establish authority and maintain credibility (Dermer, 1990).

Revision of PMS

The revision of PMS refers to the continual change and evolution in the measurement set over time (Vitale & Mavrinac, 1995). It is defined based on a notion of change in light of the modifications in the external and internal environment. As time passes and the firm's competitive environment and strategic direction change, performance indicators must be reevaluated to ensure their relevance and appropriateness (Kennerley & Neely, 2002, 2003). The revision refers to the periodic reevaluation of the appropriateness of the established performance measures in view of the current competitive environment (Forza & Salvador, 2000; Waggoner, Neely, & Kennerley, 1999; Wisner & Fawcett, 1991). The revision reflects dynamic capabilities which institutionalize the need for continuously changing measures and ensuring evolution in the measurement set (Dixon, Nanni, & Vollman, 1990; Kuwaiti, 2004; Vitale & Mavrinac, 1995). It represents the last step in the ongoing process for developing an effective PMS. The revision is comprised of the addition of measures, the deletion of measures, the changes in target, and the changes in the definition of measures (Bourne et al., 2000).

Relationships among Aspects of PMS

The three aspects of the PMS process are complementary and they reflect various interdependencies. Firstly, depending on the specific use of PMS, the type of performance indicators chosen may not be the same (Ittner & Larcker, 2001). For instance, when using PMS to monitor periodical divisional results, corporate managers may focus more on financial measures than nonfinancial measures. Indeed, financial information is strongly related to a traditional planning and control cycle where outcomes are compared to preset standards to identify variances and correct deviations (Nanni, Dixon, & Vollmann, 1992). On the other hand, some managers might prefer to use nonfinancial measures to focus attention throughout the organization because they are more traceable to strategic actions and actionable (Fisher, 1992). Conversely, depending on the type of performance indicators available, the use of PMS by managers may vary significantly. For instance, a limited set of performance measures may restrict the use of PMS to support strategic decision-making (Atkinson et al., 1997; Kaplan & Norton, 1992).

Empirically, Ittner and Larcker (1997) examine the use of quality-related measures for various purposes, including feedback by managers and review by the board of directors. Lingle and Schiemann (1996) and Ittner et al. (2003b) describe the importance of financial, internal processes, customer and innovation measures for monitoring, strategic decision-making, and attention-focusing. Cavalluzzo and Ittner (2004) and Naranjo-Gil and Hartmann (2007) provide evidence supporting the influence of the extent of measurement (i.e., financial and nonfinancial) on various uses of PMS. Conversely, Henri (2006b) suggests that the nature of use influences measurement diversity. Indeed, an attention-focusing and strategic decision-making use of PMS tend to lead to more diversity of measurement, whereas a monitoring and legitimization use tend to lead to less diversity. Also, Naranjo-Gil and Hartmann (2006) observe that the coercive use of management accounting systems (MAS) reflects three first-order factors, namely financial information, diagnostic use, and performance evaluation, whereas the enabling use reflects nonfinancial information, interactive use, and resource allocation. Bhimani and Langfield-Smith (2007) observe that more emphasis is placed on financial information in strategy implementation, whereas both financial and nonfinancial information are used in strategy development. Overall, these studies support the idea of reciprocal relationships between the design and use of PMS. More specifically, a larger set of performance indicators seems to be associated with a more intense use of PMS by managers.

Secondly, depending on the design and use of PMS, the revision may also vary. For instance, a system that is used occasionally and based mainly on a limited set of traditional financial measures (i.e., sales volume, return on investment, profit) might not need to be frequently revised and thus, be more static. On the other hand, a system used extensively for multiple purposes and reflecting a great diversity of performance measures may need to be revised periodically to remain relevant in light of internal and external changes (Dixon et al., 1990; Lynch & Cross, 1991). Furthermore, depending on the frequency of revision, the use of PMS can also differ. For instance, the constant addition and deletion of performance measures to reflect a spirit of continuous improvement might foster the use of PMS to support strategic decision-making and focus organizational attention (Bititci et al., 2000; Malina & Selto, 2004; Vitale & Mavrinac, 1995).

Some studies examine the links between the revision of performance indicators and the design and use. For instance, Kennerley and Neely (2002, 2003) and Waggoner et al. (1999) propose various facilitators in the evolution of PMS. Among those factors, two are related to the use of PMS, namely the integration of measurements within a business process review (monitoring) and a forum to discuss the appropriateness of performance measures (attention-focusing). Also, Bourne et al. (2000) and Malina and Selto (2004) describe how performance indicators are changed, deleted, or added due to attributes related to the design and use of PMS that are not achieved (e.g., diversity and complementarity, objectivity and accuracy, strategic communication devices, supportive of improved decision, etc.).

In sum, it is argued that PMS are comprised of several aspects that interact together following reciprocal relationships to form groups of PMS that share common profiles. Following a configurational approach, common alignments of PMS aspects occurring with regularity are expected to be identified. Despite some recent studies that have examined the links between aspects of PMS, it remains difficult to predict the exact nature of these alignments and the nature of the common profiles. Hence, no specific hypothesis or proposition is formulated.

METHODOLOGY

Data Collection

The data used in this study were collected using a cross-sectional survey. The survey implementation followed four steps: (i) pre-notification, (ii) initial mailing, (iii) first follow-up, and (iv) second follow-up. The first step consisted of a letter, phone call, or email to respondents to generate early interest. Then, a mail out package including the following three elements was sent to every contact name: cover letter, questionnaire, and business-reply envelope. In some circumstances, the questionnaire was sent by fax or email. The first follow-up was a postcard reminder three weeks after the initial mailing, while the second one was a phone call or replacement questionnaire three weeks after the first follow-up. To establish content validity, existing and validated scales used in the existing literature were employed. Moreover, the questionnaire was pre-tested in three stages. First, several academics were asked to revise the questionnaire. Three top managers were subsequently interviewed. They were asked to complete the questionnaire and to provide comments on its form and content. The questionnaire was then completed by a group of MBA students. Minor adjustments were made to the wording and presentation.

The target population consisted of 2,175 top management teams of Canadian manufacturing firms listed in the Scott's database with primary and secondary SIC codes in the range of 21–39. In this study, "firm" refers to a fully autonomous entity or a subunit of a larger firm. In all cases, firms appeared as separate entities in the database. Furthermore, the firms were large enough to ensure that organizational variables apply (Miller, 1987) and to ensure that a formal PMS is in place (Bouwens & Abernethy, 2000). Thus, the firms selected in the sample respect the following two criteria: (i) sales are at least \$20 million Canadian yearly and (ii) at least 150 people are employed. However, the lack of contact names in the database in several cases reduced the number of usable firms in the target population to 1,692. Data were collected using a structured questionnaire sent to the highest member of the "corporate" top management team (autonomous entity) or "local" top management team (subunit) for which the identity was revealed in the database.[1]

A total of 383 usable questionnaires were received, for a response rate of approximately 24%.[2] This response rate is satisfactory considering that it is similar to the 15–25% range reported in similar recent studies (e.g., Baines & Langfield-Smith, 2003; Lee, Lee, & Pennings, 2001; Spanos & Lioukas, 2001). Appendix A presents the statistics of the respondents in terms of position, experience, size (number of employees), and industry classification. To test whether respondents differed from nonrespondents, a two-step analysis was conducted. Respondents were first compared to nonrespondents in terms of sample characteristics (size, location, industry). Next, early and late respondents (used as proxies for nonrespondents) were compared to detect any difference in the mean score of each construct. Using χ^2 statistics, no significant differences were found between the size, location, and industry of respondent and nonrespondent firms. A comparison of the means of the variables found no significant difference between early and late respondents. Hence, the analysis did not reveal any systematic differences between respondents and nonrespondents.

Measurement and Validation of Construct

The design of PMS is measured with an adapted version of the instrument used by Hoque and James (2000) and Hoque, Mia, and Alam (2001). Based on the four dimensions of the balanced scorecard, this instrument asks for the extent of use of 20 financial and nonfinancial measures ranging on a seven-point Likert-type scale. To better reflect the context related to the unit of analysis (i.e., top management team) and to better balance the financial and internal processes dimensions, adjustments have been made to the instrument: (i) three items are added to better capture the financial dimension and (ii) three items of the internal processes dimension are ignored. An average score is calculated for each of the four dimensions based on their respective items. A higher score indicates the collection and utilization of a broader set of financial and nonfinancial indicators.

The use of PMS is measured using 27 items from an adapted version of the instrument of Vandenbosch (1999) and Brockmann and Simmonds (1997). The former is used to measure three of the four uses (monitoring, attention-focusing, and legitimization), whereas the latter is used to measure the strategic decision-making dimension. Two items are added to the attention-focusing dimension to better reflect the definition of the construct. To measure the strategic decision-making dimension, seven elements from the instrument of Brockmann and Simmonds (1997) are used. Those items were chosen because they are the most generic (i.e., they refer to strategic decision-making in general), while the others refer to specific strategic decisions (e.g., venturing, new regulations, etc.). All questions are asked using a seven-point Likert-type scale. An average score is calculated for each of the four uses based on their respective items. A higher score indicates a more frequent use of PMS for each purpose.

On the basis of the work of Bourne et al. (2000), a four-item instrument has been developed to measure the PMS revision. The respondents were asked to indicate on a seven-point Likert-type scale how often during the last 12 months each of the following events happened: (i) performance indicators were deleted from the measurement system, (ii) performance indicators were added to the measurement system, (iii) changes occurred in the performance indicators' targets, and (iv) changes occurred in the definition of the performance indicators. An average score is computed for the four items. Thus, a higher mean score indicates more reviews of the performance indicators.

Several tests have been conducted to assess construct validity and they reflect satisfactory results. Specifically, confirmatory factor analyses (CFA) were conducted and Cronbach α values were calculated. For every construct, all factor loadings were significant ($p < .01$), the Cronbach α coefficients exceeded the common cut-off level of 0.70 (Nunnally, 1967), and the goodness-of-fit indices mostly respect the recommended threshold values. The indices used to assess the model are among the most frequently reported, namely nonnormed fit index (NNFI), comparative fit index (CFI), and root mean square error of approximation (RMSEA). The threshold values recommended are (i) NNFI >0.90 (Tabachnick & Fidell, 2001), (ii) CFI >0.95 (Hu & Bentler, 1995), and (iii) RMSEA <0.10 (Browne & Cudeck, 1993). Also, to test discriminant validity between the four uses of PMS, and between the four types of performance indicators, a series of χ^2 difference tests have been performed between two dimensions at a time by constraining the estimated correlation parameter to 1.0 (Anderson & Gerbing, 1988). A significantly lower χ^2 value for the unconstrained model provides support for discriminant validity. For every pair of dimensions tested, the unconstrained model provides evidence of discriminant validity ($p < 0.05$). In sum, these tests suggest that the variables measured reflect strong validity and reliability. Appendix B illustrates the results of the CFA, Cronbach α, as well as the survey items for the PMS aspects. Appendix C presents the descriptive statistics of every construct (mean, standard deviation, and correlation matrix).

Data Analysis

Cluster analysis is used to identify groups reflecting common alignment of elements. This statistical technique sorts observations into similar sets or groups for which variance among elements grouped together is minimized while between-group variance is maximized (Ketchen & Shook, 1996). A two-stage procedure is used to gain benefits from both hierarchical and nonhierarchical methods (Hair, Anderson, Tatham, & Black, 1998; Punj & Stewart, 1983). First, a hierarchical algorithm is used to identify the number of clusters and cluster centroids. To increase confidence in the results, this step is conducted using two different methods, namely the Ward's and average linkage method,[3] as well as two types of distance, namely the Euclidean and Mahalanobis distance.[4] The results of this first step are used as seed points for the nonhierarchical clustering. This second step allows for the fine-tuning of the results by permitting the switching of cluster membership.

One of the most perplexing issues in cluster analysis is to determine the final number of clusters. The use of multiple methods is suggested in the literature to deal with this issue. Two techniques based on the agglomeration coefficient are used in this study (Aldenderfer & Blashfield, 1984): (i) a graph reflecting the number of clusters against the agglomeration coefficient (the appropriate number of clusters is found at the "elbow" of the graph) and (ii) an examination of the incremental changes in the agglomeration coefficient (the appropriate number of clusters is found at the step before a sudden jump occurs).

Given the absence of statistical tests to validate the results, split samples and hold out samples are used to increase reliability and validity of the complete procedure. Looking for stability and replicability, the two-stage procedure is conducted on each of the following samples built from the initial sample of 383 cases:

- Sample 1: main sample of 300 observations randomly chosen;
- Sample 1a: first half of the main sample (150 observations randomly chosen);
- Sample 1b: second half of the main sample (150 observations randomly chosen);
- Sample 2: Hold out sample (83 observations randomly chosen).

Two different sets of data are used to conduct a series of first-level and second-level analyses. The aim of the first-level analysis is to develop a taxonomy of PMS. The two-stage cluster analysis discussed above is conducted using simultaneously the three aspects of PMS as a data set. The main analyses are conducted at the category level (e.g., financial measures, customers measures, etc.) instead of the item level (e.g., ROE, market share, etc.).[5] Once the clusters are identified, multivariate analysis (MANOVA) and discriminant analysis are conducted to statistically validate the robustness of the classification. Afterwards, analysis of variance (ANOVA), Tukey's pairwise comparison, and Cohen's *d* factor are conducted to identify differences between the clusters for the various aspects of PMS.

The aim of the second-level analysis is to provide validation for the taxonomy developed. The two-stage cluster analysis is once again conducted, this time using simultaneously the three aspects of PMS and also four organizational factors as a data set. The four organizational factors considered are (i) size, (ii) decentralization, (iii) strategy, and (iv) perceived environmental uncertainty (PEU).[6] Those factors have been selected because they represent common factors used in contingency-based management accounting research (Chenhall, 2003; Chapman, 1997). Furthermore, previous research using a cartesian approach suggests a relationship between those factors and PMS. The objective of this validation phase is twofold: (i) statistical validation and (ii) content validation. Firstly, it intends to demonstrate the stability and replicability of the results by comparing the patterns of relationships observed among PMS aspects using a different data set. Secondly, by examining the potential differences between PMS groups for different organizational factors, the content of the taxonomy could be validated in light of past research. Once again, MANOVA and discriminant analysis are used to validate the clusters, whereas ANOVA, Tukey's pairwise, and Cohen's *d* factor are conducted to identify factors associated with the alignment of PMS dimensions.

RESULTS AND DISCUSSION

Development of the Taxonomy (First-Level Analysis)

As mentioned, two techniques based on the agglomeration coefficients guide the selection of the final cluster solution. These techniques are conducted on each of the four samples discussed above and suggest that a three-cluster solution is the most appropriate classification for these samples. To statistically assess the robustness of this solution, MANOVA and discriminant analysis were conducted on the main sample. The MANOVA shows that the three clusters are significantly different on each dimension of PMS ($p < .001$). A discriminant model is developed based on the dimensions of PMS and by assuming that PMS are classified into three clusters. The two discriminant functions are statistically significant based upon Wilk's λ ($p < .001$). Furthermore, group centroids for each of the three clusters differ substantially and they are graphically positioned in three different quadrants. Finally, 98.7% of the originally grouped cases are correctly classified. In sum, the discriminant analysis suggests that the pattern of PMS dimensions can correctly predict the original groups of PMS.

Table 1 presents the results of the two-stage procedure conducted on each of the four samples using the Mahalanobis and Euclidean distance, as well as the Ward's and average linkage method. Specifically, panel A presents the results obtained from the main sample (sample 1) using Ward's method and the Mahalanobis distance. It contains the mean of each PMS dimension for the three clusters, the results of Tukey's pairwise comparison between each cluster, and the Cohen's *d* factor for each pair of dimensions.[7] The results suggest that each component of the PMS dimensions differs significantly among the clusters (mostly $p < 0.001$), and that the mean differences between each pair of dimensions are generally large ($d > 0.8$). For instance, the monitoring mean scores move from 4.55 (cluster 1) to 5.73 (cluster 2) to 6.22 (cluster 3). The differences between those mean scores are significant ($p < 0.001$) and large (1.64, 2.32, 0.68).

Panel B presents the results of the two-stage procedure replicated on sample 1 but using different combinations of methods and distance measures: (i) Ward's method/Euclidean distance, (ii) average method/Euclidean, and (iii) average method/Mahalanobis. The examination of the results suggests stability and replicability of the results. In fact, the same pattern of relationships is observed among PMS dimensions and the means of the clusters are mostly similar to panel A.

Panel C presents the results of the two-stage procedure using Ward's method and the Mahalanobis distance but unlike panel A, validation samples are used: (i) sample 1a (first half of the sample), (ii) sample 1b (second half of the sample), and (iii) sample 2 (hold out sample). Once again, the same pattern of relationships is observed among PMS dimensions and the means of the clusters are mostly similar to panels A and B. Hence, this supports the stability and replicability of the results.

The three groups derived from the cluster analysis are labeled as follows: outcomes surveillance mechanism (cluster 1), management support tool (cluster 2), and institutionalized organizational process (cluster 3). In short, PMS in the first group have low mean scores on the various dimensions (< 4.0) except for the financial measures and the monitoring use that have moderate mean scores (4.84 and 4.55, respectively). PMS in the second group have mainly moderate mean scores (between 4 and 5.5), whereas PMS in the third group have mostly high mean scores (> 5.5). Table 2 summarizes the characteristics of the three PMS groups, whereas Fig. 1

Table 1. Development of a Taxonomy (First-Level Analysis).

Panel A: Main Sample (Sample 1)

(Ward's Method – Mahalanobis)	Clusters			Tukey's Pairwise Comparison			Cohen's *d* Factor		
	(1)	(2)	(3)	1–2	1–3	2–3	1–2	1–3	2–3
PMS design									
Financial	4.84	5.42	6.07	***	***	***	0.75	1.61	0.86
Internal processes	3.27	4.84	5.84	***	***	***	1.57	2.57	1.00
Innovation and learning	2.52	2.92	4.46	*	***	***	0.41	2.00	1.59
Customer	3.87	4.61	5.53	***	***	***	0.84	1.87	1.03
PMS use									
Monitoring	4.55	5.73	6.22	***	***	***	1.64	2.32	0.68
Attention focusing	3.77	5.07	5.88	***	***	***	1.77	2.88	1.11
Strategic decision-making	3.73	4.75	5.54	***	***	***	1.36	2.41	1.04
Legitimization	3.66	4.75	5.53	***	***	***	1.59	2.71	1.12
PMS revision	2.64	3.76	4.18	***	***	**	1.00	1.38	0.38
Number of cases (*N*)	49	160	91						

Panel B: Replication (Sample 1)

	Clusters (Ward's/Euclidean)			Clusters (Average/Euclidean)			Clusters (Average/Mahalanobis)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
PMS design									
Financial	4.89	5.38	6.06	4.86	5.42	6.07	4.89	5.41	6.07
Internal processes	3.06	4.85	5.84	3.09	4.89	5.84	3.06	4.88	5.84
Innovation and learning	2.37	2.94	4.44	2.42	2.98	4.45	2.37	2.99	4.45
Customer	3.76	4.61	5.51	3.79	4.61	5.57	3.76	4.63	5.57
PMS use									
Monitoring	4.58	5.69	6.21	4.55	5.71	6.24	4.58	5.69	6.24
Attention focusing	3.76	5.03	5.87	3.78	5.05	5.90	3.76	5.05	5.90
Strategic decision-making	3.76	4.71	5.54	3.77	4.72	5.59	3.76	4.72	5.59
Legitimization	3.67	4.72	5.51	3.68	4.72	5.58	3.67	4.71	5.58
PMS revision	2.63	3.74	4.17	2.65	3.72	4.24	2.63	3.72	4.24
Number of cases (*N*)	45	162	93	46	167	87	45	168	87

Table 1. (*Continued*).

Panel C: Validation Samples									
(Ward's Method – Mahalanobis)	Clusters (Sample 1a)			Clusters (Sample 1b)			Clusters (Sample 2)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
PMS design									
Financial	4.95	5.29	5.89	4.89	5.45	6.09	4.19	5.53	5.56
Internal processes	2.57	5.07	5.48	3.29	4.86	5.87	3.20	3.65	5.43
Innovation and learning	2.40	2.70	4.03	2.67	3.00	4.54	2.27	2.21	4.00
Customer	3.53	4.56	5.36	3.86	4.61	5.62	3.40	3.65	5.13
PMS use									
Monitoring	5.03	5.44	6.17	4.54	5.73	6.21	4.02	5.68	5.79
Attention focusing	4.27	4.81	5.80	3.69	5.06	5.80	3.47	5.07	5.41
Strategic decision-making	4.01	4.66	5.43	3.61	4.67	5.53	3.63	4.49	5.19
Legitimization	3.92	4.48	5.28	3.63	4.87	5.58	3.68	4.64	5.17
PMS revision	2.57	3.40	4.37	2.97	3.73	4.14	1.98	3.60	4.23
Number of cases (N)	23	64	63	25	78	47	10	31	42

Note: Samples: 1, main sample ($n = 300$); 1a, first half of the main sample ($n = 150$); 1b, second half of the main sample ($n = 150$); 2, hold out sample ($n = 83$). $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

illustrates the positioning of the three PMS groups on a three-dimension scatter plot. The large and significant distance between groups for most of the PMS dimensions and the specific area covered by each group in the scatter plot suggest that the three PMS groups reflect a high percentage of nonoverlap. They constitute different patterns of relationships among the design, use, and revision of PMS that occur with regularity within the sample. Based on Table 2 and Fig. 1, the three PMS groups are discussed specifically below.

Cluster 1: PMS as an Outcomes Surveillance Mechanism

PMS in this cluster are used occasionally, on an ad hoc basis by managers, and they are not developed to a great extent ($X_{\text{design}} = 3.63$, $Y_{\text{use}} = 3.93$, $Z_{\text{revision}} = 2.64$). Reflecting the traditional cybernetic approach, they are mostly used for the monitoring of financial measures. PMS in this group are not used extensively to focus attention, to support decision-making or to legitimate actions, and not much attention is devoted to nonfinancial measures. Narrow in scope, the information contained in this group of PMS is mostly historical and ex post. PMS in this category are generally static as reflected by the low level of revisions. In sum, being used periodically to

Table 2. Synthesis of the Three PMS Groups.

	Importance of Dimensions[a]			Distance Between Groups[b]		
	Outcomes surveillance mechanism (cluster 1)	Management support tool (cluster 2)	Institutionalized organizational process (cluster 3)	1–2	1–3	2–3
PMS design						
Financial	++	++	+++	Medium	Large	Large
Internal processes	+	++	+++	Large	Large	Large
Innovation and learning	+	+	++	Small	Large	Large
Customer	+	++	+++	Large	Large	Large
PMS use						
Monitoring	++	+++	+++	Large	Large	Medium
Attention focusing	+	++	+++	Large	Large	Large
Decision making	+	++	+++	Large	Large	Large
Legitimization	+	++	+++	Large	Large	Large
PMS revision	+	+	++	Large	Large	Small

[a]"+", mean score < 4; "++", mean score > = 4 and < = 5.5; "+++", mean score > 5.5.
[b]Based on Cohen's d factor presented in the panel A of Table 1, Cohen (1988) suggests to interpret the standardized difference between two means as small ($d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$).

review financial results, PMS in this group represent administrative mechanisms, which allow the surveillance of organizational outcomes.

Cluster 2: PMS as a Management Support Tool

PMS in this cluster are used on a regular basis by managers and they reflect a moderate degree of deployment ($X_{\text{design}} = 4.45$, $Y_{\text{use}} = 5.08$, $Z_{\text{revision}} = 3.76$). PMS in this category reflect moderate diversity in the measurement of performance indicators, as indicated by the moderate use of financial, internal processes, and customer indicators. The information contained in those PMS is broader than the previous group (outcomes surveillance mechanism), but it is more focused on specific measures while ignoring other indicators compared to the next group (institutionalized organizational process). Based on a more balanced mix of financial and nonfinancial information, PMS in this group are used extensively for

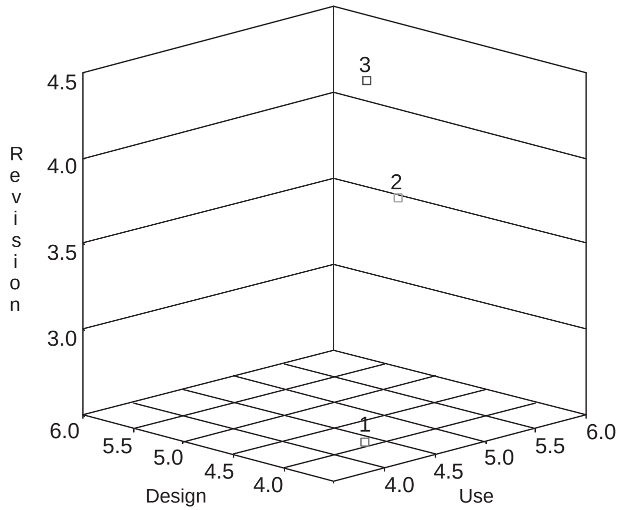


Fig. 1. Illustration of the Three PMS Groups. *Notes:* (1) The Coordinates of Each PMS Group on the 3D Scatter Plot (x,y,z) is Determined as Follows Using the Results of Table 1 (Panel A): $X_{\text{design}} = \text{Mean}$ (Financial, Internal Processes, Innovation and Learning, Customer); $Y_{\text{use}} = \text{Mean}$ (Monitoring, Attention-Focusing, Decision-Making, Legitimization); Z_{revision}. (2) The Coordinates of the PMS Group are: Cluster 1, Outcomes Surveillance Mechanism (3.63, 3.93, 2.64); Cluster 2, Management Support Tool (4.45, 5.08, 3.76); Cluster 3, Institutionalized Organizational Process (5.48, 5.79, 4.18).

monitoring purposes and they are used moderately to focus organizational attention, support strategic decision-making, and legitimate actions. Some effort has been made to revise performance indicators more frequently, but overall they remain relatively static. In sum, considered as a management tool, PMS in this group contain a broader set of performance measures which are used in multiple ways to support management activities.

Cluster 3: PMS as an Institutionalized Organizational Process

PMS in this cluster are used extensively by managers and they reflect a full degree of deployment ($X_{\text{design}} = 5.48$, $Y_{\text{use}} = 5.79$, $Z_{\text{revision}} = 4.18$). PMS are well integrated in management activities and organizational routines. Financial and nonfinancial informations are well balanced and supported by several indicators in each of the four dimensions of the balanced scorecard framework. The information available reflects a broad scope, PMS are used extensively to monitor outcomes, focus attention, support decision-making, and legitimate actions. Being used extensively, changes are made within PMS periodically to revise their components, and thus reflect a spirit of continuous improvement within the organization. In sum, PMS in this group are considered to be more than just a mechanism or a tool; instead they are a continuous and integrated process, which has multiple ramifications on organizational routines.

The content of the current taxonomy is consistent with findings of other studies related to PMS. First, past studies have reported specifically for each aspect of PMS current practices observed in various organizations (e.g., Lingle & Schiemann, 1996; Ittner et al., 2003b; Gosselin, 2005; Kennerley & Neely, 2002, 2003). Those studies suggest three main conclusions that are also supported by the results of this taxonomy: (i) despite various critics, financial measures remain among the most important performance indicators, (ii) more importance is devoted to customer and internal process measures than innovation and learning measures, (iii) among the various uses of PMS, monitoring appears to be the most widely integrated use in organizational practices, and (iv) despite the fact that the current business environment is characterized by fast changes, the majority of organizations do not appear to have systematic processes in place to manage the evolution of their PMS.

Second, some studies have examined how various types of performance indicators were used by managers for different purposes and how those indicators were revised (e.g., Cavalluzzo & Ittner, 2004; Henri, 2006b; Malina & Selto, 2004). Globally, those studies, as well as the current taxonomy, suggest interdependencies among the design, use, and revision of PMS. More specifically, the current taxonomy suggests that measurement diversity, the nature of use, and the review of performance indicators are aligned together and have evolved simultaneously. Indeed, from cluster 1 to cluster 3, the mean score of the three aspects increases simultaneously: (i) design (3.63, 4.45, 5.48), (ii) use (3.93, 5.08, 5.79), and (iii) revision (2.64, 3.76, 4.18). This suggests the presence of reciprocal relationships among the three aspects of PMS. More diversity of measurement may lead to more use of PMS because more information is available for various purposes. At the same time, depending on the needs of managers in terms of use of performance indicators, the level of measurement diversity has to be adjusted. Similarly, a more frequent use of PMS may lead to more revisions because performance indicators have to remain relevant. At the same time, more review of performance indicators may encourage the use of PMS because they contain relevant information aligned to the changing circumstances. Moreover, more diversity of measurement may lead to more revisions of PMS because they cover a large range of organizational activities that are subjected to various changes in light of internal and external change. At the same time, more revisions of performance indicators may lead to more diversity because new indicators are added to reflect important issues for the business. In sum, the examination of the three clusters suggests that the more diversity of measurement within PMS, the more they are used by managers and the more performance indicators are revised. Inversely, the less diversity of measurement within PMS, the less they are used by managers and the less performance indicators are revised. Hence, the three PMS aspects are interrelated to reflect patterns of relationships and common profiles.

Validation of the Taxonomy (Second-Level Analysis)

To provide reassurance that the observed associations are statistically robust and meaningful, cluster analysis has once again been conducted, this time using simultaneously the three aspects of PMS and four organizational factors as a data set (i.e., size, structure, strategy, and PEU). Table 3 presents the results of a two-stage cluster analysis conducted on the main sample (sample 1) using Ward's method and the Mahalanobis distance. It contains the mean of each PMS aspect and organizational factors for each cluster, the results of Tukey's pairwise comparison between each cluster, and the Cohen's *d* factor for each pair of dimensions. The results reflect the presence of three different clusters. The examination of the three aspects of PMS reveals similar patterns to those observed from the first-level analysis (Table 1). Indeed, similar group means are associated with the various components of PMS aspects. We recognize the common alignment of the "outcomes surveillance mechanism" (cluster 1), "management support tool" (cluster 2), and "institutionalized organizational process" (cluster 3). The stability and replicability of the results provide validity to the taxonomy presented previously because the pattern of relationships among PMS aspects is also captured in a holistic analysis including additional variables.[8]

The examination of organizational factors reveals significant differences among the three groups of PMS. The results suggest that firms pertaining to cluster 1 (outcomes surveillance mechanism) are significantly smaller ($p<0.01$; Cohen's $d=$ medium) and more centralized ($p<0.001$; Cohen's $d=$ large) compared to the two other clusters. The mean score of

Table 3. Validation of the Taxonomy (Second-Level Analysis).

Sample 1: Main Sample (Ward's Method – Mahalanobis)	Clusters			Tukey's Pairwise Comparison			Cohen's *d* Factor		
	(1)	(2)	(3)	1–2	1–3	2–3	1–2	1–3	2–3
PMS design									
Financial	4.87	5.37	6.13	***	***	***	0.68	1.70	1.02
Internal processes	3.30	4.90	5.75	***	***	***	1.58	2.41	0.84
Innovation and learning	2.53	2.94	4.40	*	***	***	0.41	1.90	1.49
Customer	3.79	4.66	5.51	***	***	***	0.99	1.96	0.97
PMS use									
Monitoring	4.72	5.68	6.24	***	***	***	1.28	2.03	0.75
Attention focusing	3.87	5.04	5.89	***	***	***	1.58	2.71	1.13
Strategic decision-making	3.81	4.75	5.52	***	***	***	1.22	2.23	1.02
Legitimization	3.81	4.75	5.45	***	***	***	1.27	2.22	0.95
PMS revision	2.55	3.76	4.27	***	***	**	1.11	1.58	0.47
Organizational size	2.41	2.63	2.61	**	**	n.s.	0.54	0.49	0.05
Strategy	4.46	4.36	4.76	n.s.	*	***	0.12	0.38	0.50
Environmental uncertainty	3.30	3.38	3.35	n.s.	n.s.	n.s.	0.07	0.03	0.04
Decentralization	3.50	4.63	5.33	***	***	***	1.31	2.11	0.81
Number of cases (*N*)	52	155	93						

Note: All the constructs, except for size, are measured using a seven-point Likert-type scale. Size is measured using the natural log of the number of employees. $^{*}p<.05$; $^{**}p<.01$; $^{***}p<.001$.

strategy suggest that firms in cluster 1 place less emphasis on differentiation than firms in cluster 3 ($p<0.05$; Cohen's $d=$ small). No significant difference is observed for the PEU compared to the other clusters. Firms belonging to cluster 2 (management support tool) are significantly less decentralized ($p<0.001$; Cohen's $d=$ large) and place more emphasis on cost leadership ($p<0.001$; Cohen's $d=$ medium) than firms belonging to cluster 3. No significant difference is observed for the PEU. Lastly, the results suggest that firms belonging to cluster 3 (institutionalized organizational process) are significantly more decentralized and more oriented toward differentiation compared to the firms belonging to clusters 1 and 2. Firms in cluster 3 are significantly larger than those in cluster 1 ($p<0.01$; Cohen's $d=$ medium) but no difference is observed with any cluster for PEU.

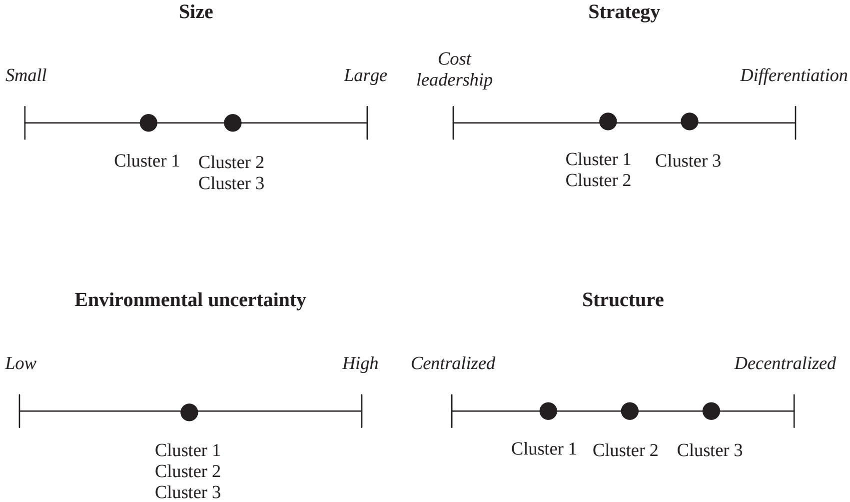


Fig. 2. Synthesis of the Relationships among PMS Groups and Organizational Factors. *Notes:* (1) Cluster 1, Outcomes Surveillance Mechanism; Cluster 2, Management Support Tool; Cluster 3, Institutionalized Organizational Process. (2) The Graphs are Not Intended to be Scaled But Merely Illustrative of the Differences between Groups for Each Organizational Factor.

In sum, the examination of organizational factors reveals significant differences among the groups of PMS for three out of the four factors. The three factors reflecting differences are internal factors (size, strategy, decentralization), while the only external factor (PEU) does not show any difference among the three groups. Fig. 2 provides a synthesis of the relationships between PMS groups and organizational factors. The remainder of this section discusses in more detail the observed results and their links with past studies examining the relationships between PMS and organizational factors.

Size and Decentralization

Small and centralized firms are described by Moores and Yuen (2001) as organizations in which minimal amounts of information are used for decision-making. They use mostly informal controls but when formal control systems are developed, they are generally based on simple and narrowly defined measures. As size and decentralization increases, accounting and control processes tend to become more specialized and sophisticated (Bruns & Waterhouse, 1975; Merchant, 1981, 1984; Khandwalla, 1977). The need to stimulate effective communication flows becomes more apparent (Burns & Stalker, 1961; Chandler, 1962). As a result, managers devote greater effort to collecting and processing information, and decision-making tends to focus on a broad array of factors which demands a larger amount of information (Moores & Yuen, 2001).

In the current study, cluster 1 (outcomes surveillance mechanism) is associated with small and centralized firms. As previously mentioned, PMS in this group are used occasionally by managers (mainly for monitoring) and they are not developed to a great extent (mostly financial measures). As size and decentralization increase from cluster 1 to clusters 2 and 3, PMS are used more intensively as communication devices, decision-making tools, and legitimization devices. The scope of information becomes broader to include various types of nonfinancial measures and the performance measures are revised more regularly. Furthermore, our results are in line with those obtained by other researchers who have examined empirically the relationships among size, structure, and PMS (mostly in terms of a mix of financial and nonfinancial information). For instance, evidence has been provided to support the influence of decentralization on the greater use of nonfinancial indicators (e.g., Abernethy & Lillis, 2001; Baines & Langfield-Smith, 2003; Dixon et al., 1990; Gosselin, 2005). Similarly, some authors document the positive influence of size on the mix of financial and nonfinancial measures (e.g., Hoque & James, 2000; Khandwalla, 1977; Moores & Yuen, 2001).

Strategy

The generic strategies developed by Porter (1980) and Miles and Snow (1978), which have notable similarities according to Miller (1986), present different control requirements. According to those classifications, the differentiators/prospectors have (i) broad and tentative planning practices occurring after feedback from action, (ii) decentralized and horizontal information systems, and (iii) performance measures focused on effectiveness in exploiting opportunities, used to revise plans. In contrast, cost leaders/defenders are characterized by (i) intensive, detailed, and definitive planning practices occurring before action, (ii) centralized and vertical information systems, and (iii) performance measures focused on task efficiency, used to secure conformity with plans. For firms pursuing a differentiator/prospector-type strategy, financial measures will influence the manager to pay less attention to the firm's critical success factors and competitive bases, such as price, quality, reliability, service, innovation, customization, and time. Hence, a greater emphasis on nonfinancial criteria as opposed to narrow financial criteria should be more prominent in prospector firms than in defender firms (Hoque, 2004).

In the current study, firms in cluster 3 (institutionalized organizational process) place more emphasis on differentiation than firms in the two other clusters. As previously mentioned, PMS in this group are used extensively by managers for various purposes. Financial and nonfinancial information are well balanced and supported by several indicators which are revised periodically. These results are in line with the previous arguments stating that if management wishes to stress effectiveness in innovation, developing customer satisfaction and a reasonable rate of return, PMS should be designed and used to support these arrangements. They also support past studies that adhere to the association between differentiator/prospector strategy and greater use of financial and nonfinancial measures (e.g., Abernethy & Guthrie, 1994; Boulianne, 2002; Govindarajan, 1988; Govindarajan & Fisher, 1990; Gosselin, 2005; Hoque, 2004; Said et al., 2003).

Perceived Environmental Uncertainty

The more changes that occur in the competitive environment, regulation, and technology, the more important the managers' PEU becomes (Tymon, Stout, & Shaw, 1998). Managers are then expected to process and use more information for decision-making (Ewusi-Mensah, 1981; Galbraith, 1973; Gordon & Narayanan, 1984). They also tend to use more information within the management control systems to understand uncertain situations and to cope with the complexities of the environment (e.g., Chenhall, 2003; Chenhall & Morris, 1986; Gul & Chia, 1994; Mia, 1993). More specifically, as PEU increases, the use and scope of PMS also increases (e.g., Dixon et al., 1990; Hoque et al., 2001; Gosselin, 2005; Ittner & Larcker, 1997; Said et al., 2003). Surprisingly, the current results do not reflect any difference in the level of PEU among the three PMS groups. On the one hand, it is possible that the associations captured by bivariate or multivariate analysis between PEU and control systems are incomplete. In other words, internal variables such as strategy, structure, and size may represent correlated omitted variables that align with environmental uncertainty and control systems. In the context of configurational analysis whereby all organizational factors are analyzed together, the explanatory power of uncertainty may be limited by the presence of those internal factors. This may reflect one shortcoming of a contingency approach, i.e., its ability to capture the complexity of organizational reality (Ketchen & Shook, 1996). On the other hand, the current results may be the consequence of the measurement problems mentioned earlier with the construct of PEU. Consequently, the conclusions related to PEU and PMS aspects must be taken with caution.

In sum, the analysis of validation provides reassurance that the observed associations among PMS aspects are statistically robust and meaningful. Indeed, statistically, the stability and replicability of the results provide validity to the taxonomy presented previously because the pattern of relationships among PMS aspects has been reproduced using a different data set. In terms of content, the relationships observed between PMS groups and organizational factors are globally coherent with past research and thus, support the PMS groups identified.

DISCUSSION

The main purpose of this exploratory study was to move the investigation of PMS from a cartesian form of contingency fit to a configuration form. A taxonomy approach was used to develop a classification scheme of PMS which conceptualizes the relationships among three basic aspects of PMS, namely the design, use, and revision of performance indicators. Three patterns of relationship reflecting the role and importance of PMS within the organizations emerge from the analysis of empirical data: (a) PMS as an outcomes surveillance mechanism, (b) PMS as a management support tool, and (c) PMS as an institutionalized organizational process. Those groups differ in terms of (i) scope of information (narrow to broad), (ii) frequency of use (occasionally to extensively), (iii) nature of use (focused to diversified), and (iv) review of performance indicators (occasionally to regularly). This reflects the view that PMS, as well as the management control systems (MCS) overall, have moved from a mechanistic, passive, coercive, and static view whereby performance measurement was based mainly on financial indicators and considered as a component of the planning and control cycle, to an organic, active, enabling, and dynamic view based on multiple financial and nonfinancial indicators where performance measurement acts as an independent process included in a broader set of activities (Dent, 1987; Chapman, 1997, 1998; Chenhall & Morris, 1995; Henri, 2004; Ahrens & Chapman, 2004). The current taxonomy reflects this movement of PMS from a mechanistic view (outcomes surveillance mechanism) to an organic view (institutionalized organizational process).

This study contributes to the management accounting literature in two ways. Firstly, it expands on the research on management control systems by moving from a cartesian approach to a configurational approach. The latter allows for the consideration of multiple and interrelated aspects. As discussed by (Miller, 1996): ''their predictive power [configurations] relies on the fact that most alignments are unlikely while relatively few are far more common.'' Interestingly, the configurational approach has not been used extensively in management accounting settings [notable exceptions are the work of Chenhall and Langfield-Smith (1998) and Moores and Yuen (2001). Secondly, a taxonomy approach has allowed for the discovery of reliable and conceptually significant clustering attributes, which has not been done for PMS in previous studies. Those results provide a different understanding of the various levels of integration of PMS within organizational routines. Moreover, by moving away from a basic dichotomy ''presence versus absence of PMS,'' this study provides a basis to shed new light on the understanding of the influence of PMS within organizational settings.

This study is subject to potential limitations in terms of internal and external validity. Cluster analysis is criticized for its extensive reliance on researchers' judgement. However, using a two-stage procedure, two different methods, two measures of distance, split samples, and hold out samples, the replicability and stability of the results have been demonstrated. Moreover, this article examines three crucial aspects of PMS but does not integrate: (i) all dimensions of those aspects (e.g., the aspect of design does not include the mix of current and future information, level of aggregation, etc.), (ii) all the potential attributes of PMS (e.g., quality of measurement, alignment with strategic priorities, cause-effect linkages), and (iii) other management control systems (e.g., budget, incentives). For instance, even carefully designed PMS used for attention-focusing and revised periodically may fail in coordinating employee's efforts and their decision-making if not linked to the budgeting and incentive systems. Also, the analysis is conducted at the category level and not the item level. Furthermore, using the survey method to collect data creates the potential for bias due to common-response. Lastly, considering differences in the design and use of control systems among firms depending on their size and industry, results cannot be generalized outside the scope of the current sample (i.e., small-to-medium size manufacturing firms).

A number of directions for further research emerge from this study. Future research could investigate the relationship between PMS life cycle and organizational life cycle. Do PMS and organizations evolve simultaneously or does a delay occur between the two? Moreover, research could build on the work of Miller (1996) and attempt to determine how the relationships among PMS aspects vary depending on the degree of configuration throughout the firm (low or high) and the presence of central themes. Also, future research is needed to examine this classification in other industries, especially service firms. Lastly, other attributes of PMS not included in this study and other organizational factors could be investigated and linked to the current taxonomy.

NOTES

1. Following other upper echelon studies (e.g., Carpenter & Fredrickson, 2001), top management team is defined as the top two tiers of an organization's management, which include CEO/general manager, chief operating officer (COO), chief financial officer (CFO), and the next highest management tier of a firm (senior vice-presidents).

2. The response rate was calculated as the percentage of the number of usable returned questionnaires to the number of questionnaires sent, after adjusting for the firms which had closed, ended manufacturing activities or moved, or for which the contact person had left the organization.

3. The Ward's method is a hierarchical clustering procedure in which the similarity used to join clusters is calculated as the sum of squares between the two clusters summed over all variables. The average linkage method is an agglomerative method that represents similarity as the average distance from all objects in one cluster to all objects in another (Hair et al., 1998). The Ward's method is used as the main method in this article, whereas the average linkage method is used to validate the results.

4. The squared Euclidean distance is the most commonly used measure of the similarity between two objects, whereas the Mahalanobis distance is a standardized form of Euclidean distance which adjusts for intercorrelations among the variables (Hair et al., 1998). As noted by Ketchen and Shook (1996), a high correlation among clustering variables can be problematic because it may overweigh one or more underlying construct. Considering the potential multicollinearity between dimensions of PMS or between contextual factors, the Mahalanobis distance is used as a main distance measure in the analysis. The square Euclidean distance is used to validate the results.

5. The category level has been preferred to the item level as unit of analysis for various reasons. First, the number of observations in the various samples is not large enough to conduct cluster analysis with 51 items (i.e., design, 20 items; use, 27 items; revision, 4 items). Furthermore, the reliability as well as the convergent and discriminant validity of the various categories has been supported in the previous section. In other words, the various items are varying together and they are reflective of the same underlying category. Lastly, multicollinearity acts as a weighting process in cluster analysis in favor of the set of variables having more items. Thus, as suggested by Hair et al. (1998), the variables should be reduced to similar numbers in each set to compensate. In this study, we observe multicollinearity among the items of the PMS aspects. Hence, considering the difference in the number of items among

PMS aspects, the analysis could not simultaneously contain one PMS aspect at the category level and the other at the item level. Hence, four categories have been used for each of the two PMS aspects that reflect the largest number of items (i.e., design and use).

6. Strategy refers to the choices made by managers to position their organization in particular environments (Chenhall, 2003). The strategy differs depending on the development of competitive advantages based on lowest cost or differentiation (Porter, 1980). Strategy is measured using the instrument developed by Govindarajan (1988). Respondents are asked to position their products relative to those of leading competitors in six areas. The higher the score, the more the firm follows a differentiation strategy. Inversely, the lower the score, the more the firm follows a low-cost strategy. Among the various dimensions of structure, decentralization has been used extensively in management accounting studies as a proxy for organizational structure. Decentralization refers to the distribution of power in an organization (Miller, Dröge, & Toulouse, 1988). A three-item instrument developed by Miller and Dröge (1986) is used to measure decentralization. A higher score indicates a greater extent of decentralization. Lastly, PEU refers to the top managers' perceived inability to predict an organization's external environment accurately (Milliken, 1987). Govindarajan's (1984) instrument is used to assess perceived environmental uncertainty. The respondents are asked to assess the predictability or unpredictability of eight environmental factors. The higher the score, the more uncertain the firm's perceived environment is. Size is measured using the natural log of the number of employees. As for the PMS aspects, several tests have been conducted to assess construct validity and they generally reflect satisfactory results (see Appendix B). Specifically, CFAs were conducted and Cronbach α values were calculated. For every construct, all factor loadings were significant ($p < .01$) and the Cronbach α coefficients respect the common cut-off level of 0.70 (Nunnally, 1967). The goodness-of-fit indices related to the CFA respect the recommended threshold values, except for the CFI of environmental uncertainty, which almost reaches the threshold.

7. Factor d is an indice measuring the magnitude of a treatment effect. Unlike significance tests, this indice is independent of sample size. Cohen (1988) defined d as the difference between the means, M_1–M_2, divided by the pooled standard deviation. Cohen suggests interpreting the standardized difference between two means as small ($d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$).

8. The robustness of this solution is supported by the results of a MANOVA and discriminant analysis. The results of the MANOVA show that the three clusters are significantly different for the dimensions of PMS as well as contextual factors ($p < .001$), except for PEU. The discriminant analysis reveals that the two discriminant functions are statistically significant ($p < .001$), the group centroids for each cluster differ substantially (graphically positioned in three different quadrants), and 95% of the original grouped cases are correctly classified. Furthermore, the two-stage procedure was also conducted on samples 1a and 1b (half sample), and for all samples using a combination of Ward's and average linkage method, as well as Euclidean and Mahalanobis distance. Similar patterns of relationships and conclusions are generally reflected by those analyses that provide evidence of the stability of the results. Details of these tests are available from the author upon

request. The second-level analysis was not conducted on the hold out sample (sample 2, $n = 83$) because the sample size is too small compared to the number of clustering variables.

ACKNOWLEDGMENTS

I would like to thank Murray Lindsay and Nicole Thibodeau for their insightful comments and suggestions. I would also like to acknowledge the useful contributions of the participants at the 2005 CAAA conference and EIASM 3rd Workshop on Performance Measurement and Management Control.

REFERENCES

Abernethy, M. A., & Guthrie, C. H. (1994). An empirical assessment of the "fit" between strategy and management information system design. *Accounting and Finance*, *34*(2), 49–66.

Abernethy, M. A., & Lillis, A. M. (2001). Interdependencies in organization design: A test in hospitals. *Journal of Management Accounting Research*, *13*, 107–129.

Ahrens, T., & Chapman, C. S. (2004). Accounting for flexibility and efficiency: A field study of management control systems in a restaurant chain. *Contemporary Accounting Research*, *21*(2), 271–301.

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis*. Beverly Hills, CA: Sage.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423.

Ansari, S., & Euske, K. J. (1987). Rational, rationalizing, and reifying uses of accounting data in organizations. *Accounting, Organizations and Society*, *12*(6), 549–570.

Atkinson, A. A., Waterhouse, J. H., & Wells, R. B. (1997). A stakeholder approach to strategic performance measurement. *Sloan Management Review*, *38*(3), 25–37.

Baines, A., & Langfield-Smith, K. (2003). Antecedents to management accounting change: A structural equation approach. *Accounting, Organizations and Society*, *28*(7–8), 675–698.

Banker, R. D., Chang, H., & Pizzini, M. J. (2004). The balanced scorecard: Judgemental effects of performance measures linked to strategy. *The Accounting Review*, *79*(1), 1–24.

Bhimani, A., & Langfield-Smith, K. (2007). Structure, formality and the importance of financial and non-financial information in strategy development and implementation. *Management Accounting Research*, *18*, 3–31.

Bisbe, J., & Otley, D. (2004). The effects of the interactive use of management control systems on product innovation. *Accounting, Organizations and Society*, *29*(8), 709–737.

Bititci, U. S., Turner, T., & Begemann, C. (2000). Dynamics of performance measurement systems. *International Journal of Operations and Production Management*, *20*(6), 692–704.

Boland, R. J., & Pondy, L. R. (1983). Accounting in organizations: A union of natural and rational perspectives. *Accounting, Organizations and Society*, *8*(2/3), 223–234.

Boulianne, E. (2002). Linking strategic choices and management accounting systems scope: An empirical analysis. In: M. J. Epstein & J. Manzoni (Eds), *Performance measurement and management control: A compendium of research*. Kidlington, Oxford: Elsevier Science.

Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing and updating performance measurement systems. *International Journal of Operations and Production Management*, *20*(7), 754–771.

Bouwens, J., & Abernethy, M. A. (2000). The consequences of customization on management accounting system design. *Accounting, Organizations and Society*, *25*, 221–241.

Brockmann, E. N., & Simmonds, P. G. (1997). Strategic decision making: The influence of CEO experience and use of tacit knowledge. *Journal of Managerial Issues*, *IX*(4), 454–467.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In: K. A. Bollen & J. Scott Long (Eds), *Testing structural equation models*. Newburry Park, CA: Sage.

Bruns, W. J., & Waterhouse, J. H. (1975). Budgetary control and organization structure. *Journal of Accounting Research*, *13*(2), 177–203.

Burchell, S., Clubb, C., Hopwood, A., & Hughes, J. (1980). The roles of accounting in organizations and society. *Accounting, Organization and Society*, *5*(1), 5–27.

Burns, T., & Stalker, G. M. (1961). *The management of innovation*. London: Tavistock Publications.

Carpenter, M. A., & Fredrickson, J. W. (2001). Top management teams, global strategic posture, and the moderating role of uncertainty. *Academy of Management Journal*, *44*(3), 533–545.

Cavalluzzo, K. S., & Ittner, C. D. (2004). Implementing performance measurement innovations: Evidence from government. *Accounting, Organizations and Society*, *29*(3–4), 243–267.

Chandler, A. D. (1962). *Strategy and structure*. Cambridge: MIT Press.

Chapman, C. S. (1997). Reflections on a contingent view of accounting. *Accounting, Organizations and Society*, *22*(2), 189–205.

Chapman, C. S. (1998). Accountants in organisational networks. *Accounting, Organizations and Society*, *23*(8), 737–766.

Chenhall, R. H. (2003). Management control 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. H., & Langfield-Smith, K. (1998). The relationship between strategic priorities, management techniques and management accounting: An empirical investigation using a systems approach. *Accounting, Organizations and Society*, *23*(3), 243–264.

Chenhall, R. H., & Morris, D. (1986). The impact of structure, environment and interdependence on the perceived usefulness of management accounting systems. *The Accounting Review*, *61*(1), 16–35.

Chenhall, R. H., & Morris, D. (1995). Organic decision and communication processes and management accounting systems in entrepreneurial and conservative business organizations. *Omega, International Journal of Management Science*, *23*(5), 485–497.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Earlbaum Associates.

deHaas, M., & Kleingeld, A. (1999). Multilevel design of performance measurement systems: Enhancing strategic dialogue throughout the organization. *Management Accounting Research*, *10*, 233–261.

Dent, J. F. (1987). Tensions in the design of formal control systems: A field study in a computer company. In: W. J. Bruns & R. S. Kaplan (Eds), *Accounting and management: Field study perspectives*. Boston, MA: Harvard Business School Press.

Dermer, J. (1990). The strategic agenda: Accounting for issues and support. *Accounting, Organizations and Society*, *15*(1–2), 67–76.

Dixon, J. R., Nanni, A. J., & Vollman, T. E. (1990). *The new performance challenge-measuring operations for world-class competition (Dow Jones-Irwin ed.)*. Homewood, IL: Dow Jones-Irwin.

Ewusi-Mensah, K. (1981). The external organizational environment and its impact on management information systems. *Accounting, Organizations and Society*, *6*(4), 301–316.

Feldman, M. S., & March, J. G. (1981). Information in organizations as signal and symbol. *Administrative Science Quarterly*, *26*, 171–186.

Fisher, J. (1992). Use of nonfinancial performance measures. *Journal of cost management*, *6*(1), 31–38.

Fitzgerald, L., Johnston, R., Brignall, S., Silvestro, R., & Voss, C. (1991). *Performance measurement in service businesses*. London: The Chartered Institute of Management Accountants.

Forza, C., & Salvador, F. (2000). Assessing some distinctive dimensions of performance feedback information in high performing plants. *International Journal of Operations and Production Management*, *20*(3), 359–385.

Galbraith, J. (1973). *Designing complex organizations*. Reading, MA: Addison-Wesley.

Gerdin, J., & Greve, J. (2004). Forms of contingency fit in management accounting research – A critical review. *Accounting, Organizations and Society*, *29*, 303–326.

Gerdin, J., & Greve, J. (2008). The appropriateness of statistical methods for testing contingency hypotheses in management accounting research. *Accounting, Organizations and Society*, *33*(7/8), 995.

Gordon, L. A., & Narayanan, V. K. (1984). Management accounting systems, perceived environmental uncertainty and organization structure: An empirical investigation. *Accounting, Organizations and Society*, *9*(1), 33–47.

Gosselin, M. (2005). An empirical study of performance measurement in manufacturing firms. *The International Journal of Productivity and Performance Management*, *54*(5/6), 419.

Govindarajan, V. (1984). Appropriateness of accounting data in performance evaluation: An empirical examination of environmental uncertainty as an intervening variable. *Accounting, Organizations and Society*, *9*(2), 125–135.

Govindarajan, V. (1988). A contingency approach to strategy implementation at the business-unit level: Integrating administrative mechanisms with strategy. *Academy of Management Journal*, *31*(4), 826–853.

Govindarajan, V., & Fisher, J. (1990). Strategy, control systems, and resource sharing: Effects on business-unit performance. *Academy of Management Journal*, *33*(2), 259–285.

Gul, F. A., & Chia, Y. M. (1994). The effects of management accounting systems, perceived environmental uncertainty and decentralization on managerial performance: A test of three-way interaction. *Accounting, Organizations and Society*, *19*(4–5), 413–426.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis (5th ed.)*. Upper Saddle River, NJ: Prentice Hall.

Hall, M. (2008). The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance. *Accounting, Organizations and Society*, *33*, 141–163.

Henri, J.-F. (2004). Performance measurement and organizational effectiveness: Bridging the gap. *Managerial Finance*, *30*(6), 93–123.

Henri, J.-F. (2006a). Management control systems and strategy: A resource-based perspective. *Accounting, Organizations and Society*, *31*(6), 529–558.

Henri, J.-F. (2006b). Organizational culture and performance measurement systems. *Accounting, Organizations and Society*, *31*(1), 77–103.

Hofstede, G. (1978). The poverty of management control philosophy. *Academy of Management Review*, *3*(3), 450–461.

Hoque, Z. (2004). A contingency model of the association between strategy, environmental uncertainty and performance measurement: Impact on organizational performance. *International Business Review*, *13*, 485–502.

Hoque, Z., & James, W. (2000). Linking balanced scorecard measures to size and market factors: Impact on organizational performance. *Journal of Management Accounting Research*, *12*, 1–17.

Hoque, Z., Mia, L., & Alam, M. (2001). Market competition, computer-aided manufacturing and use of multiple performance measures: An empirical study. *British Accounting Review*, *33*, 23–45.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In: R. H. Doyle (Ed.), *Structural equation modeling*. Thousand Oaks, CA: Sage.

Hyvönen, J. (2007). Strategy, performance measurement techniques and information technology of the firm and their links to organizational performance. *Management Accounting Research*, *18*, 343–366.

Ittner, C. D., & Larcker, D. F. (1997). Quality strategy, strategic control systems, and organizational performance. *Accounting, Organizations and Society*, *22*(3,4), 293.

Ittner, C. D., & Larcker, D. F. (2001). Assessing empirical research in managerial accounting: A value-based management perspective. *Journal of Accounting and Economics*, *32*, 349–410.

Ittner, C. D., Larcker, D. F., & Meyer, M. W. (2003a). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review*, *78*(3), 725–758.

Ittner, C. D., Larcker, D. F., & Randall, T. (2003b). Performance implications of strategic performance measurement in financial services firms. *Accounting, Organizations and Society*, *28*(7-8), 715–741.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard – Measures that drive performance. *Harvard Business Review*, *70*(1), 71–79.

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard: Translating strategy into action (Harvard Business School Press ed.)*. Boston, MA: Harvard Business School Press.

Kennerley, M. (2003). Measuring performance in a changing business environment. *International Journal of Operations and Production Management*, *23*(2), 213–229.

Kennerley, M., & Neely, A. (2002). A framework of the factors affecting the evolution of performance measurement systems. *International Journal of Operations and Production Management*, *22*(11), 1222–1245.

Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: An analysis and critique. *Strategic Management Journal*, *17*, 441–458.

Khandwalla, P. N. (1977). *The design of organizations*. New York: Harcourt Brace Jovanovich Inc.

Kuwaiti, M. E. (2004). Performance measurement process: Definition and ownership. *International Journal of Operations and Production Management*, *24*(1), 55–78.

Lee, C., Lee, K., & Pennings, J. M. (2001). Internal capabilities, external networks, and performance: A study on technology-based ventures. *Strategic Management Journal*, *22*, 615–640.

Libby, T., Salterio, S. E., & Webb, A. (2004). The balanced scorecard: The effects of assurance and process accountability on managerial judgment. *The Accounting Review*, *79*(4), 1075–1094.

Lingle, J. H., & Schiemann, W. A. (1996). From balanced scorecard to strategic gauges: Is measurement worth it. *Management Review*, *8*(3), 56–61.

Lipe, M. G., & Salterio, S. (2002). A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society*, *27*, 531–540.

Lynch, R. L., & Cross, K. F. (1991). *Measure up – Yardsticks for continuous improvement*. Cambridge, MA: Basil Blackwell.

Malina, M. A., & Selto, F. H. (2001). Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard. *Journal of Management Accounting Research*, *13*, 47–90.

Malina, M. A., & Selto, F. H. (2004). Choice and change of measures in performance measurement models. *Management Accounting Research*, *15*, 441–469.

Markus, M. L. (1983). Power, politics and MIS implementation. *Communications of the ACM*, *26*(6), 430–444.

Merchant, K. A. (1981). The design of the corporate budgeting system: Influences on managerial behavior and performance. *The Accounting Review*, *56*(4), 813–829.

Merchant, K. A. (1984). Influences on departmental budgeting: An empirical examination of a contingency model. *Accounting, Organizations and Society*, *9*(3/4), 291–307.

Meyer, A. D., Tsui, A. S., & Hinings, C. R. (1993). Configurational approaches to organizational analysis. *Academy of Management Journal*, *36*(6), 1175–1195.

Mia, L. (1993). The role of MAS information in organisations: An empirical study. *The British Accounting Review*, *25*(3), 269–285.

Miles, R. E., & Snow, C. C. (1978). *Organizational strategy, structure and process*. New York: McGraw Hill.

Miller, D. (1986). Configurations of strategy and structure: Towards a synthesis. *Strategic Management Journal*, *7*, 233–249.

Miller, D. (1987). The structural environmental correlates of business strategy. *Strategic Management Journal*, *8*(1), 55–76.

Miller, D. (1996). Configurations revisited. *Strategic Management Journal*, *17*, 505–512.

Miller, D., & Dröge, C. (1986). Psychological and traditional determinants of structure. *Administrative Science Quarterly*, *31*, 539–560.

Miller, D., Dröge, C., & Toulouse, J.-M. (1988). Strategic process and content as mediators between organizational context and structure. *Academy of Management Journal*, *31*(3), 544–569.

Miller, D., & Friesen, P. (1984). *Organizations: A quantum view*. Englewood Cliffs, NJ: Prentice-Hall.

Miller, D., & Shamsie, J. (1996). The resource-based view of the firm in two environments: The Hollywood film studios from 1936 to 1965. *Academy of Management Journal*, *39*(3), 519–543.

Milliken, F. J. (1987). Three types of perceived uncertainty about the environment: State, effect and response uncertainty. *Academy of Management Review*, *12*, 133–143.

Moores, K., & Yuen, S. (2001). Management accounting systems and organizational configuration: A life-cycle perspective. *Accounting, Organizations and Society*, *26*, 351–389.

Nanni, A. J., Dixon, R., & Vollmann, T. E. (1992). Integrated performance measurement: Management accounting to support the new manufacturing realities. *Journal of Management Accounting Research*, *4*(Fall), 1–19.

Naranjo-Gil, D., & Hartmann, F. (2006). How top management teams use management accounting systems to implement strategy. *Journal of Management Accounting Research*, *18*, 21.

Naranjo-Gil, D., & Hartmann, F. (2007). Management accounting systems, top management team heterogeneity and strategic change. *Accounting, Organizations and Society*, *32*(7–8), 735–756.

Neely, A., Adams, C., & Kennerley, M. (2002). *The performance prism: The scorecard for measuring and managing business success*. Hemel Hempstead: Prentice-Hall.

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Porter, M. E. (1980). *Competitive strategy: Techniques for analyzing industries and competitors*. New York: The Free Press.

Punj, G., & Stewart, D. W. (1983). Cluster analysis in marketing research: A review and suggestions for application. *Journal of Marketing Research*, *20*, 134–148.

Said, A. A., Elnaby, H. R. H., & Wier, B. (2003). An empirical investigation of the performance consequences of nonfinancial measures. *Journal of Management Accounting Research*, *15*, 193–223.

Scott, T. W., & Tiesen, P. (1999). Performance measurement and managerial teams. *Accounting, Organizations and Society*, *24*, 263–285.

Simon, H. A., Guetzkow, H., Kozmetsky, G., & Tyndall, G. (1954). *Centralization vs decentralization in organizing the controllers' department*. New York: Controllership Foundation Inc.

Simons, R. (1990). The role of management control systems in creating competitive advantage: New perspectives. *Accounting, Organizations and Society*, *15*(1/2), 127–143.

Simons, R. (1995). *Levers of control: How managers use innovative control systems to drive strategic renewal*. Boston: Harvard Business School Press.

Simons, R. (2000). *Performance measurement and control systems for implementing strategy (Prentice Hall ed.)*. Upper Saddle River, NJ: Prentice Hall.

Spanos, Y. E., & Lioukas, S. (2001). An examination into the causal logic of rent generation: Contrasting Porter's competitive strategy framework and the resource-based perspective. *Strategic Management Journal*, *22*, 907–934.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics (4th ed.)*. Needham Heights, MA: Allyn & Bacon.

Tymon, W. G., Stout, D. E., & Shaw, K. N. (1998). Critical analysis and recommendations regarding the role of perceived environmental uncertainty in behavioral accounting research. *Behavioral Research in Accounting*, *10*, 23–46.

Vandenbosch, B. (1999). An empirical analysis of the association between the use of executive support systems and perceived organizational competitiveness. *Accounting Organizations and Society*, *24*, 77–92.

Van der Stede, W. A., Chow, C. W., & Lin, T. W. (2006). Strategy, choice of performance measures, and performance. *Behavioral Research in Accounting*, *18*, 185–205.

Vitale, M. R., & Mavrinac, S. C. (1995). How effective is your performance measurement system? *Management Accounting*, *77*(2), 43–55.

Waggoner, D. B., Neely, A. D., & Kennerley, M. P. (1999). The forces that shape organisational performance measurement systems: An interdisciplinary review. *International Journal of Production Economics*, *60–61*, 53–60.

Widener, S. K. (2006). Association between strategic resource importance and performance measure use: The impact on firm performance. *Management Accounting Research*, *17*, 433–457.

Wisner, J. D., & Fawcett, S. E. (1991). Linking firm strategy to operating decisions through performance measurement. *Production and Inventory Management Journal*, *32*(3), 5–11.

APPENDIX A. PROFILE OF THE RESPONDENTS

Position	%	Experience within the firm (average in years)
CEO/General manager	29	18.4
COO	29	14.6
Senior Vice-presidents	28	13.2
CFO/Vice-president finance	11	10.5
Other	3	10.0
Average		14.7

Number of Employees	%
Fewer than 499	66
Between 500 and 999	18
Between 1,000 and 4,999	13
Between 5,000 and 9,999	2
Between 10,000 and 19,999	1
Average	796

Industry Classification	%
Food and kindred products	8.4
Tobacco manufactures	0.3
Textile mill products	3.1
Apparel and other textile products	4.2
Lumber and wood products	10.4
Furniture and fixture	4.2
Paper and allied products	8.1
Printing and publishing	1.8
Chemicals and allied products	4.4
Petroleum and coal products	1.6

APPENDIX A. (*Continued*)

Industry Classification	%
Rubber and miscellaneous plastics products	3.9
Leather and leather products	1.3
Stone, clay, glass, and concrete products	3.1
Primary metal industries	6.0
Fabricated metal products	10.4
Industrial machinery and equipment	10.4
Electrical and electronic equipment	7.3
Transportation equipment	7.3
Instrument and related products	2.3
Miscellaneous manufacturing industries	1.3

APPENDIX B. QUESTIONNAIRE ITEMS AND STATISTICS OF MEASUREMENT ANALYSIS

1. PMS Design

Please rate the extent to which each of the following measures is used by your top management team.

Scale: 1 = not at all to 7 = to a very great extent.

Constructs and Items	Confirmatory Factor Analysis	Cronbach α
Financial		0.70
Operating income	0.892**	
Sales growth	0.439**	
Return-on-investment (ROI)	0.697**	
Return-on-equity (ROE)[a]	0.879**	
Net cash flows[a]	0.416**	
Costs per unit produced[a]	0.808**	
Internal processes[b]		0.75
Materials efficiency variance	1.084**	
Manufacturing lead time	1.005**	
Rate of material scrap loss	1.086**	
Labor efficiency variance	1.043**	

Constructs and Items	Confirmatory Factor Analysis	Cronbach α
Innovation and learning		0.73
Number of new patents	0.809**	
Number of new product launches	0.940**	
Time-to-market for new products	0.479**	
Employee satisfaction	1.002**	
Customer		0.73
Market share	1.054**	
Customer response time	1.191**	
On-time delivery	1.030**	
Number of customer complaints	0.954**	0.73
Number of warranty claims	1.047**	
Survey of customer satisfaction	0.522**	

Goodness-of-fit of the model: χ^2 (157) = 499.136; $p <$ .001; NNFI = 0.903; CFI = 0.920; RMSEA = 0.077

[a]These three items have been added to the original instrument.[b]Three items of the original instrument have been ignored (i.e., ratio of good output to total output at each production process, % of shipments returned due to poor quality, and number of overdue deliveries).

2. PMS Revision

During the last 12 months, how often have each of the following events related to your performance measurement system occurred?
Scale: 1 = never to 7 = regularly

Constructs and Items	Confirmatory Factor Analysis	Cronbach α
Performance indicators were deleted from the measurement system	0.726**	0.76
Performance indicators were added within the measurement system	1.218**	
Changes occurred in performance targets	1.209**	
Changes occurred in the definition of performance indicators	1.308**	

Goodness-of-fit of the model: χ^2 (2) = 0.687; $p >$ .001; NNFI = 0.998; CFI = 1.0; RMSEA = 0.0001

3. PMS Use

Please rate the extent to which your top management team currently uses performance measures.
Scale: 1 = not at all to 7 = to a great extent

Constructs and Items	Confirmatory Factor Analysis	Cronbach α
Monitoring		0.79
Track progress towards goals	0.985**	
Monitor results	0.801**	
Compare outcomes to expectations	0.879**	
Review key measures	0.837**	
Attention-focusing		0.87
Enable discussion in meetings of superiors, subordinates, and peers[a]	0.964**	
Enable continual challenge and debate underlying data, assumptions, and action plans[a]	0.929**	
Provide a common view of the organization	1.084**	
Tie the organization together	0.986**	
Enable the organization to focus on common issues	1.016**	
Enable the organization to focus on critical success factors	0.794**	
Develop a common vocabulary in the organization	1.058**	
Strategic decision-making		0.86
Make strategic decisions once the need for a decision is identified, and an immediate response is required	0.840**	
Make strategic decisions once the need for a decision is identified, and an immediate response is not required	0.844**	
Make decisions when it is difficult to differentiate among plausible solutions to a problem because each has good arguments	1.004**	
Make decisions when encountering a problem that is unstructured and has not been encountered before	0.979**	
Make decisions when you have been recently faced with a similar decision	1.005**	
Anticipate the future direction of the company, as opposed to responding to an identifiable problem	0.958**	
Make a final decision on a strategic issue of major importance	0.783**	

Constructs and Items	Confirmatory Factor Analysis	Cronbach α
Legitimization		0.87
Confirm your understanding of the business	0.889**	
Justify decisions	0.623**	
Verify assumptions	0.915**	
Maintain your perspectives	0.924**	
Support your actions	0.968**	
Reinforce your beliefs	0.902**	
Stay close to the business	0.901**	
Increase your focus	0.954**	
Validate your point of view	0.993**	

Goodness-of-fit of the model: χ^2 (318) = 1206.621; $p<$ .001; NNFI = 0.955; CFI = 0.959; RMSEA = 0.090

[a]These two items have been added to the original instrument.

4. Strategy

Please position your main products relative to those of leading competitors in the following six areas.

Scale: 1 = significantly lower to 7 = significantly higher

Constructs and Items	Confirmatory Factor Analysis	Cronbach α
Product selling price	0.400**	0.71
Percentage of sales spent on research and development	0.569**	
Percentage of sales spent on marketing expenses	0.515**	
Product quality	0.687**	
Brand image	0.889**	
Product features	0.642**	

Goodness-of-fit of the model: χ^2 (8) = 35.96; $p<$.001; NNFI = 0.900; CFI = 0.947; RMSEA = 0.096

5. Perceived Environmental Uncertainty (PEU)

Please indicate the extent that each of the following factors is predictable or unpredictable in the context of your main business.
Scale: 1 = highly predictable to 7 = highly unpredictable

Items (First Order Construct)	Confirmatory Factor Analysis	Cronbach α
Manufacturing technology	0.643**	0.71
Competitors' actions	0.805**	
Market demand	0.562**	
Product attributes/design	0.540**	
Raw material availability	0.740**	
Raw material price	0.645**	
Government regulation	0.691**	
Labor union actions	0.483**	

Goodness-of-fit of the model: χ^2 (17) = 57.794; $p < .001$; NNFI = 0.90; CFI = 0.939; RMSEA = 0.079

6. Decentralization

To what extent is decision-making at top levels in your firm characterized by the use of integrative mechanisms (committees, task forces, liaison personnel) to decide the following classes of decisions.
Scale: 1 = used rarely to 7 = used very frequently

Items (First Order Construct)	Confirmatory Factor Analysis	Cronbach α
Products or service decisions (e.g., production, marketing, R&D strategies)	0.661**	0.70
Capital budget decisions (e.g., the selection and financing of long-term investments)	1.318**	
Long-term strategies (growth, diversification, etc.) and decisions related to changes in the firm's operating philosophy.	1.282**	

Goodness-of-fit of the model: χ^2 (0) = 0.0; $p = 1.0$; NNFI = 1.0; CFI = 1.0; RMSEA = 0.0 (saturated model)

Note: *Significant at the .05 level; **Significant at the .01 level.

APPENDIX C. DESCRIPTIVE STATISTICS

	Mean	SD	Correlation Matrix (Pearson)								
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Financial	5.48	.92	1								
(2) Internal processes	4.79	1.36	.298**	1							
(3) Innovation/learning	3.27	1.26	.241**	.388**	1						
(4) Customer	4.67	1.13	.274**	.505**	.525**	1					
(5) Monitoring	5.63	.98	.454**	.378**	.162**	.316**	1				
(6) Attention-focusing	5.07	1.05	.405**	.423**	.308**	.386**	.674**	1			
(7) Strategic decision-making	4.79	.99	.386**	.411**	.310**	.353**	.534**	.674**	1		
(8) Legitimization	4.79	.96	.378**	.396**	.305**	.334**	.603**	.751**	.743**	1	
(9) Revision	3.70	1.27	.177**	.255**	.296**	.277**	.244**	.282**	.269**	.228**	1

Note: *Significant at the 0.05 level; **Significant at the 0.01 level.

ACCOUNTING CHOICE AND OPTIMAL INCENTIVE CONTRACTS: A ROLE OF FINANCIAL REPORTING IN MANAGEMENT PERFORMANCE EVALUATION

Parunchana Pacharn

ABSTRACT

Three structural properties of accounting commonly embedded in Generally Accepted Accounting Principles are examined in a two-period principal-agent model. These structural properties are conservation of income, consistency, and selective recognition. The article illustrates that these properties are essential for the use of accounting information in management performance evaluation: they are necessary conditions for an accounting mechanism to be more efficient than a direct revelation mechanism. The trade-off between the gain from the information revelation and the incentive cost of discretion determines whether contracting is more efficient under the accounting mechanism or under the direct revelation mechanism.

Advances in Management Accounting, Volume 17, 289–316
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017012

1. INTRODUCTION

The use of accrual accounting numbers in management performance evaluation has been viewed as problematic because some practices allowed by Generally Accepted Accounting Principles (GAAP) may discourage goal-congruent or encourage self-serving behavior. Examples of such practices include charging for sunk costs and letting management select among several accounting method choices. In optimal contracting literature, the shareholders shall design optimal performance measures (such as cash flows, stock prices, reports from the managers, etc.) and corresponding contracts to resolve the agency problem without relying on accrual accounting numbers. In practice, however, accounting numbers are commonly used in combination with other measures for management evaluation. Moreover, many blue-chip firms have adopted earnings-based compensation plans such as Economic Value Added (EVA) or Economic Profit Plan (EPP). With this observation, few recent papers have analyzed the role of accounting numbers in management performance evaluation. For example, Dutta and Reichelstein (2005a) examine a multi-period principal-agent model in which optimal performance measures rely on all sources of information, including accrual accounting numbers such as net income and book value. Dutta and Reichelstein (2005b) develop alternative accrual accounting rules from an incentive and control perspective.

This article studies structural properties of GAAP accrual accounting. These properties are traditionally viewed as important for financial reporting systems to generate high quality earnings for valuation purposes. The article illustrates that they are also essential for the use of accounting numbers for management performance evaluation: they are necessary conditions for an accounting mechanism to be more efficient than a direct revelation mechanism. Three structural properties of accrual accounting are examined in this article: conservation of income, consistency, and selective recognition.

Conservation of income requires that the sum of earnings over the life of the firm does not depend on accounting methods – higher earnings in the current period mean an offsetting reduction in earnings in future periods (Sunder, 1997). Consistency is defined as "conformity from period to period with unchanging policies and procedures" (FASB Concepts Statement No. 2). That is, the accounting methods should be used consistently across periods. Selective recognition means that accounting methods produce accounting reports as a function of selected (yet unknown) events. An accounting method can be viewed as a set of recognition rules, which are conditioned on a set of selected information. Revenues and expenses are determined based on this set of information only; all other available information outside of this set is excluded. For example, the sum-of-the-years' digits and the double-declining balance depreciation methods use different sets of information to determine the depreciation expense. The latter excludes the salvage value information from the set.[1]

These three properties limit the manager's ability to manipulate earnings. The manager may choose from many accounting methods. However, the manager has an incentive to commit to the choice he has made because an accounting principle change is costly[2] (consistency). The chosen method may be conservative, neutral, or aggressive. However, the total earnings always add up to the same amount. An income-increasing or decreasing accrual must reverse itself in the future periods (conservation of income). Selective recognition restricts the information used in determining recognition rules. For example, the successful efforts method uses the outcome of the exploration to determine whether to capitalize or to expense exploration costs. The discretion is limited in that the manager cannot use other information available (such as the magnitude of the costs) to allocate the costs across periods in exact amounts that he may wish. Note that selective recognition is also a characteristic of a broader class of accounting choices, which includes situations in which there is no promulgated choice but management can achieve a desired result by altering structure, timing, or placement in financial statements of transactions. Examples include accounting for business combinations and accounting for software development cost.[3]

The three structural properties are incorporated in a stylized two-period principal-agent model where the agent possesses private information about his inputs and about future productivity. The model contains the following features. First, the agent can use discretion only through an accounting method choice, which must be chosen before the productivity parameter is learned. Second, the manager cannot switch to another accounting method after the choice has been made.[4] Third, if the manager chooses an accounting method that "borrows" or "lends" earnings, the reversal occurs in the second period. Fourth, the accounting methods available prescribe recognition rules based on a restricted information set whose single element is the future productivity parameter.

The manager's private information is his productivity in the second period. The article examines three mechanisms to communicate this information: (1) a direct revelation mechanism that allows the principal to observe the outputs directly and allows the manager to report the productivity parameter, (2) an accounting mechanism under which the manager, using discretion through an accounting method choice as described above, provides earnings reports to the principal, and (3) an accounting mechanism with no discretion – the manager provides earnings reports according to an accounting method pre-specified by the principal. The most efficient mechanism is the accounting mechanism with no discretion. The accounting mechanism with discretion dominates (is dominated by) the direct revelation mechanism when the hidden information problem in the second period is more (less) severe than the hidden action problem in the first period.

The intuition behind the above results is as follows. With the direct revelation mechanism, the agent receives an information rent in the second period because of the hidden information problem. This information rent is awarded to motivate the agent to work and reveal his private information. The agent also receives a bonus to motivate him to work in the first period. The accounting mechanism with a pre-specified accounting method is more efficient than the direct revelation mechanism for two reasons. First, the accounting reports partly reveal the agent's private information and thus reduces the information rent. Second, the lower information rent in the second period has a spill-over effect that helps motivate the agent to work in the first period and thus allows the principal to lower the bonus as well. When compared to the direct revelation mechanism, the accounting mechanism with discretion does not always dominate. The principal benefits from the rent reduction and the spill-over effect but incurs the cost of motivating the agent to adopt an appropriate method choice at the beginning of the first period. This mechanism is more efficient than the direct revelation mechanism if the benefit is more than the cost.

This article extends existing studies on the incentive aspects of delegating accounting method choices. Demski, Patell, and Wolfson (1984) develop a single-period agency model of decentralized accounting choice where the principal can benefit by delegating the accounting system selection to the agent who has superior information about the environment. Suh (1990) shows that delegation of accounting method choice is an alternative Pareto-equivalent mechanism to direct communication that can be used to achieve consumption smoothing by the agent. In a multi-period model, Christensen and Demski (1995) link an accrual measure to a monitor of project choice and use the accrual measure to discipline other sources of information.

This article contributes to debates on recognition by examining the incentive effects of accounting methods. Previous models study various aspects of recognition (Antle & Demski, 1989; Antle, Demski, & Ryan, 1994).

More closely related to this model is Liang (2000), which analyzes the choice of the optimal recognition rules in the presence of accounting and non-accounting information sources and shows that delaying accounting recognition may be optimal when the disciplining role of accounting reports is considered.

This study is related to the literature on contracting under alternative specifications of accounting structures. Dye and Verrecchia (1995) study the effects of altering GAAP on the agency problems between current shareholders and their manager and between current and prospective shareholders. Kirschenheiter (1999) analyzes optimal contracting under historical cost and market value accounting. Sankar and Subramanyam (2001) examine the optimal use of GAAP reporting discretion in a model where there is no communication channel for the agent to voluntarily submit private information and income smoothing arises from consumption smoothing incentives. Finally, Liang (2004) shows that earnings management activities that facilitate the efficient allocation of compensation risk across periods can arise from economic trade-offs among managers, shareholders, and regulators.

This study contributes to the literature on endogenous earnings management.[5] Earnings management may arise as rational equilibrium behavior from the principal's inability to commit to how the information will be used (Fudenberg & Tirole, 1995; Arya, Glover, & Sunder, 1998; Christensen, Demski, & Frimor, 2002) or from restrictions on the information flow (e.g., Dye, 1988; Demski, 1998). In my model, earnings management arises endogenously not because of restrictions on information flow but because the agent's message space is state-dependent (as in Evans & Sridhar, 1996).[6]

Conservation of income has been shown to be a key element in endogenous earnings management models. For example, Demski (1998) shows that shareholders can prefer managed to unmanaged earnings under the conservation of income. This article is similar to Demski (1998) in that both address recognition issues under intertemporal accounting structures. The conservation of income restricts the manager's message space in both papers. However, other elements of Demski's paper differ from this article significantly in three key aspects. First, in Demski's model, the manager is only able to manage earnings effectively when he exerts high productive effort. Second, the manager's reporting choice is determined ex post, after a potential early read. Third, communication is restricted – the manager is not allowed to forecast the second period outcome directly. In this article, the manager's ability to manipulate earnings is not related to the input choice; the reporting strategy is determined ex ante; and full communication is allowed. These differences underscore the complexity of the interaction between accounting structural properties and optimal incentive contracts. Endogenous earnings management may arise from two considerably different combinations of accounting properties.

The remainder of the article is organized as follows. Section 2 describes the model. Section 3 presents the optimal contracts. Section 4 discusses the roles of the three structural properties. Section 5 concludes the article.

2. MODEL

2.1. Structure

Consider a two-period principal-agent model, $t = 1, 2$. The principal owns a production technology that requires the agent's inputs, $a_t \in A = \{H, L\}$, personally costly for the agent. a_t determines the probability distribution over period t's outputs, $R_t \in X = \{x_0, x_1, x_2\}$. H is more productive than L; however, it is also more costly. The agent's personal cost function is $c(\cdot)$, where $c(H) = h > 0$ and $c(L) = 0$. Assume that the principal always prefers high inputs from the agent.

At the beginning of the first period, the principal offers a contract to the agent, specifying a compensation function, $W(\cdot)$. The agent decides to accept or reject the contract based on $W(\cdot)$ and his reservation utility, which is arbitrarily set to zero. If the agent rejects the contract, the game ends. If the agent accepts the contract, he privately chooses the first input, a_1. At this time, the accounting method $\omega \in \Omega$ is chosen (by the principal or the agent depending on the mechanism). Ω is defined in the following subsection.

At the end of the first period, the agent observes the first period output, R_1, which is a function of a_1 and a random state of nature. Let $P(\cdot)$ denote the probability of an event. Let $P(R_1 = x_i | a_1 = H) = p_i$ and $P(R_1 = x_i | a_1 = L) = q_i$, $i = 0, 1, 2$. Assume $p_0 < p_1 < p_2$ and $q_2 < q_1 < q_0$. This implies that the probability distribution satisfies the Monotone Likelihood Ratio Property.[7] For tractability, I assume $p_1 \geq q_1$ and $q_2(1 - q_2) \geq q_1^2$. The agent privately learns the second period productivity parameter, $\theta \in \Theta = \{b, g\}$. For simplicity, let $P(\theta = b) = P(\theta = g) = 0.5$. The principal knows the distribution but not the realization of θ. Under the direct revelation mechanism, the principal observes R_1 and receives a report $\hat{\theta} \in \Theta$. Under the accounting mechanisms, the principal receives a report $\hat{R}_1$, which is determined according to ω. Then, the agent chooses his second input, a_2.

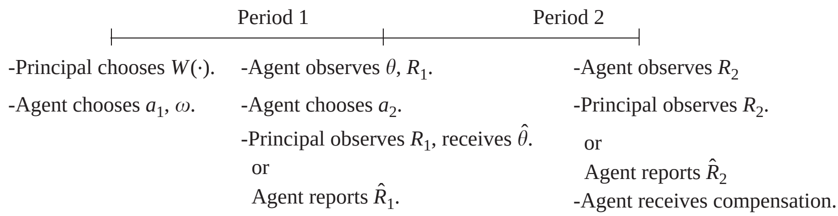


Fig. 1. Timeline.

At the end of the second period, the agent observes the second period output, R_2, which is independent of R_1. The support of the distribution of R_2 depends on θ and a_2. Let $P(R_2 = x_1|a_2 = H, \theta = b) = P(R_2 = x_0|a_2 = L, \theta = b) = 1$ and $P(R_2 = x_2|a_2 = H, \theta = g) = P(R_2 = x_1|a_2 = L, \theta = g) = 1$.[8] Under the direct revelation mechanism, the principal observes R_2. Under the accounting mechanisms, the principal receives another report $\hat{R}_2$. Finally, the agent is compensated and the game ends.

Fig. 1 summarizes the sequence of events.

Each player behaves as if she or he maximizes the expected value of a von Neuman–Morgenstern utility function. The players are indifferent regarding the timing of compensation. The agent is compensated using a single payment at the end of the game. Both players are risk-neutral.[9] The principal's utility is given by $U_p = R_1 + R_2 - W$. The agent's utility is given by $U_a(a_1, a_2, W) = W - c(a_1) - c(a_2)$.

2.2. Accounting Method

Let $\{$*truth*, *borrow*, *lend*$\}$ be the set of available recognition rules. *truth* requires that the agent's reports be the same as the actual outputs. *borrow* requires that the agent maximize the first period report. Finally, *lend* requires that the agent minimize the first period report. When the agent's reports deviate from the output, the reports are subject to two additional constraints, $\hat{R}_t \in \{x_0, x_1, x_2\}$ and $\hat{R}_1 + \hat{R}_2 = R_1 + R_2$, to guarantee that the agent's reports are not obviously false (Table 1).

For simplicity, let $x_2 - x_1 \neq x_1 - x_0$. This assumption implies that other recognition rules such as partial borrowing (where $\hat{R}_1 < \max R_1$) or smoothing (where $\hat{R}_1 = (R_1 + R_2)/2$) are not possible because they violate either the constraint $\hat{R}_2 = R_1 + R_2 - \hat{R}_1$ or the constraint $\hat{R}_t \in \{x_0, x_1, x_2\}$. Given this assumption, the above set of recognition rules is exhaustive.

Table 1. Recognition Rules.

Truth	*Borrow*	*Lend*
$\hat{R}_t = R_t$	$\hat{R}_1 = \max R_t$ s. t. $\hat{R}_2 = R_1 + R_2 - \hat{R}_1$ $\hat{R}_t \in \{x_0, x_1, x_2\}$	$\hat{R}_1 = \min R_t$ s. t. $\hat{R}_2 = R_1 + R_2 - \hat{R}_1$ $\hat{R}_t \in \{x_0, x_1, x_2\}$

This assumption is not critical to the results and is imposed only to simplify the optimization problems (see Appendix A for more details).

An accounting method, $\omega \in \Omega$, specifies recognition rules as a function of selected future events. Let $\omega = (\omega_b,\ \omega_g)$ denote an accounting method, where $\omega_\theta \in \{truth,\ borrow,\ lend\}$ is the recognition rule used when the future productivity is θ. While the recognition rules are mappings from all information to accounting reports, Ω are mappings from Θ to the set of recognition rules $\{truth,\ borrow,\ lend\}$.

3. OPTIMAL CONTRACTS

3.1. Direct Revelation Mechanism

The direct revelation mechanism allows the agent to directly report $\hat{\theta} \in \{b, g\}$ to the principal and allows the principal to directly observe and contract on the output R_t. That is, the agent must report truthfully, i.e., $\hat{R}_t = R_t$. Let $\bar{\omega} = (truth, truth)$. Let $W = w_{i,j}^{\hat{\theta}}$ denote the compensation when $\hat{R}_1 = x_i$, $\hat{R}_2 = x_j$, and $\hat{\theta} \in \{b, g\}$. Define H $(\cdot)$ in A, such that $a_2 = H$, $\forall(\theta,\ R_1)$. The principal solves the following optimization problem:

$$\min_{W(\cdot)} E(W(\hat{R}_1, \hat{R}_2, \hat{\theta}) \qquad \text{[P-1]}$$

s.t.

$$EU(\bar{\omega}, H, H(\cdot), W(\cdot)) \geq 0 \qquad \text{(IR)}$$

$$EU(\bar{\omega}, H, H(\cdot), W(\cdot)) \geq EU(\bar{\omega}, a_1, a_2(\cdot), W(\cdot)), \forall(a_1, a_2(\cdot)) \qquad \text{(IC)}$$

$$W(\cdot) \geq 0, \forall(\hat{R}_1, \hat{R}_2, \hat{\theta}) \qquad \text{(BC)}$$

The principal minimizes the expected compensation to the agent subject to the individual rationality constraint (IR), the incentive compatibility

Table 2. An Optimal Contract Under the Direct Revelation Mechanism.

R_1	(θ, a_2)			
	(b, L)	(b, H)	(g, L)	(g, H)
x_0	$w^b_{0,0} = 0$	$w^b_{0,1} = h$	$w^g_{0,1} = h$	$w^g_{0,2} = 2h$
x_1	$w^b_{1,0} = 0$	$w^b_{1,1} = h$	$w^g_{1,1} = h$	$w^g_{1,2} = 2h$
x_2	$w^b_{2,0} = 0$	$w^b_{2,1} = h$	$w^g_{2,1} = h$	$w^g_{2,2} = 2h + (2h/(p_2 - q_2))$
Expected compensation		$\left(\frac{3}{2} + \frac{p_2}{p_2 - q_2}\right)h$		

constraint (IC), and the bankruptcy constraint (BC). (IR) requires that the agent's expected utility, given high inputs and $\bar{\omega}$, be at least the agent's reservation utility, arbitrarily set to zero. (IC) requires that the agent prefer high inputs to any other combinations. (BC) guarantees the agent a non-negative compensation.[10]

Table 2 presents an optimal contract. All wages not presented in Table 2 are zero (see Appendix B for more details).

The intuition behind the above contract is as follows. Consider the first column (b, L). The principal sets $w^b_{i,0}$ to zero because he infers that the agent does not work in the second period. The wages $w^b_{i,1}$ in the second column (b, H) then are set at h to motivate the agent to work when $\theta = b$. The wages $w^g_{i,1}$ in the third column (g, L) are set at h instead of zero even though the principal infers that the agent does not work in the second period because the agent is able to lie and report $w^b_{i,1}$. The agent receives the information rent because the principal is not able to distinguish between (b, H) and (g, L) when the report is $w^b_{i,1}$. The wages $w^g_{i,2}$ in the last column must be at least $2h$ to motivate the agent to work when $\theta = g$. Note that the agent earns the information rent even with full communication of θ because the agent needs to be motivate to report $\hat{\theta}$ truthfully (or untruthfully). To motivate the agent to work in the first period, $w^g_{2,2}$ is set at $2h + (2h/(p_2 - q_2))$.

3.2. Accounting Mechanism with No Discretion

Consider a mechanism under which the principal can contract on the accounting method to be used.[11] The principal solves an optimization

Table 3. An Optimal Contract under the Accounting Mechanism with no Discretion.

R_1	(θ, a_2)			
	(b, L)	(b, H)	(g, L)	(g, H)
x_0	$\hat{w}_{0,0} = 0$	$\hat{w}_{0,1} = h$	$\hat{w}_{1,0} = 0$	$\hat{w}_{2,0} = h$
x_1	$\hat{w}_{1,0} = 0$	$\hat{w}_{1,1} = h$	$\hat{w}_{1,1} = h$	$\hat{w}_{2,1} = 2h$
x_2	$\hat{w}_{2,0} = h$	$\hat{w}_{2,1} = 2h$	$\hat{w}_{2,1} = 2h$	$\hat{w}_{2,2} = ((2 - (p_1 - q_1))/(p_2 - q_2))h$
Expected compensation	$\left(\frac{3}{2} + \frac{p_2}{p_2 - q_2}\right)h - \left(1 - \frac{(p_2 q_1 - p_1 q_2)}{p_2 - q_2}\right)\frac{h}{2}$			

problem similar to [P-1] except that the principal can specify the agent to use any $\omega \in \Omega$.

I present an optimal contract with $\omega = (truth, borrow)$ as the expected compensation needed to motivate high inputs from the agent is lowest with this accounting method (see Appendix C).[12]

Let $W = \hat{w}_{i,j}$ denote the compensation when $\hat{R}_1 = x_i$, $\hat{R}_2 = x_j$. Table 3 presents an optimal contract with $\omega = (truth, borrow)$ (see Appendix C for more details).

The agent receives the information rent when the report is $\hat{w}_{1,1}$ and $\hat{w}_{2,1}$ in the same fashion as that under the direct revelation mechanism. However, the principal can distinguish between (b, H) and (g, L) when $R_1 = x_0$. (The report is $\hat{w}_{0,1}$ for (b, H) and $\hat{w}_{1,0}$ for (g, L).) This affects the incentive cost in two ways. First, there is no information rent when $R_1 = x_0$ and $\hat{w}_{2,0}$ is set at h. Second, setting $\hat{w}_{2,0}$ at h leads to setting $\hat{w}_{2,1}$ at $2h$ to motivate the agent to work when $\theta = b$ and $R_1 = x_2$. This may appear to increase the information rent. However, this payment helps motivate the agent to work in the first period (unlike the rent awarded when $R_1 = x_0$ which reduces the incentive to work in the first period). Finally $\hat{w}_{2,2}$ is set to motivate the agent to work in the first period but the amount is lower than that under the direct revelation mechanism because the rent has a spill-over effect: elimination of the rent when $R_1 = x_0$ helps create an incentive to work and thus avoid $R_1 = x_0$. The accounting mechanism with no discretion is strictly more efficient than the direct revelation mechanism and, as we will see, the accounting mechanism with discretion as well.

3.3. Accounting Mechanism with Discretion

Consider another accounting mechanism under which the principal provides incentives for the agent to choose an accounting method from the set of accepted accounting methods Ω. The principal's optimization problem is as follows:

$$\min_{W(\cdot),\omega} E(W(\hat{R}_1, \hat{R}_2)) \tag{P-2}$$

s.t.

$$EU(\omega, H, H(\cdot), W(\cdot)) \geq 0 \tag{ir}$$

$$EU(\omega, H, H(\cdot), W(\cdot)) \geq EU(\omega', a_1, a_2(\cdot), W(\cdot)), \forall(\omega', a_1, a_2(\cdot)), \omega' \in \Omega \tag{ic}$$

$$W(\hat{R}_1, \hat{R}_2) \geq 0, \forall(\hat{R}_1, \hat{R}_2) \tag{bc}$$

The individual rationality constraint (ir), the incentive compatibility constraint (ic), and the bankruptcy constraint (bc) are counterparts of (IR), (IC), and (BC), respectively. [P-2] differs from [P-1] in that the agent has to be motivated to choose ω ex ante. Therefore, (ic) is more demanding than (IC). The agent's choices ex post (at the end of the first period) are the same in both programs.

The incentive cost under this mechanism is lower than that under the direct revelation mechanism only if ω helps reveal θ. Partial revelation of θ helps reduce the information rent because the agent's opportunity to shirk is limited. This condition is not met when $\omega = $ (*truth*, *truth*), (*borrow*, *borrow*), or (*lend*, *lend*). As the lower incentive costs are driven by the elimination of the information rent, it is necessary that the optimal accounting method prescribes different recognition rules for different θ so that the reports provide information about the realization of θ. This property has an important implication on the incentive cost as stated formally in the following proposition.

Proposition 1. If the principal motivates an accounting method with $\omega_b = \omega_g$, the direct revelation mechanism weakly dominates the accounting mechanism with discretion.

Proof. All proofs are provided in the Appendices D–F.

As illustrated in previous subsections, the incentive cost under the direct revelation mechanism [i.e., with (*truth*, *truth*)] is higher than that under the accounting mechanism with no discretion when the principal contracts on the accounting method (*truth*, *borrow*). Similarly, the incentive costs is higher when the principal contracts on (*borrow*, *borrow*) or (*lend*, *lend*) than that when he contracts on (*truth*, *borrow*). This is because the agent's reports given (*truth*, *borrow*) partially reveal θ beyond what is provided by (*borrow*, *borrow*) or (*lend*, *lend*).

Proposition 1 suggests that a choice such as the successful efforts method may be preferred to the full cost method for contracting purposes. The difference in recognition rules helps reveal information about future productivity, which allows the principal to contract more efficiently.[13]

3.4. Numerical Example

The following numerical example illustrates that the incentive cost under the accounting mechanism with discretion may still be lower than that under the direct revelation mechanism even with the additional cost of motivating the method. Let $h = 10$, $p_0 = q_2 = (1/6)$, $p_1 = q_1 = (1/3)$, and $p_2 = q_0 = (1/2)$. Optimal contracts under the direct revelation mechanism and under the accounting mechanism with no discretion are presented in Tables 4 and 5. Table 6 illustrates an optimal contract under the accounting mechanism with discretion.

Under the direct revelation mechanism, the agent receives the information rent. Because $\hat{w}_{i,1} = 10$, $\hat{w}_{i,2}$ must be at least 20. To motivate the agent to work in the first period, $\hat{w}_{2,2}$ is set at 80. Under the accounting mechanism with no discretion, the principal contracts on (*truth*, *borrow*) and therefore can distinguish between (b, H) and (g, L) when $R_1 = x_0$. $\hat{w}_{2,0}$ is set at 10. The spill-over effect from eliminating the information rent when $R_1 = x_0$ helps lower $\hat{w}_{2,2}$ which is set at 60 to motivate the agent to work in the first period. Note that $\hat{w}_{2,1}$ is set at 20 but this payment helps motivate the agent to work

Table 4. An Optimal Contract under the Direct Revelation Mechanism.

R_1	(θ, a_2)			
	(b, L)	(b, H)	(g, L)	(g, H)
x_0	$w^b_{0,0} = 0$	$w^b_{0,1} = 10$	$w^g_{0,1} = 10$	$w^g_{0,2} = 20$
x_1	$w^b_{1,0} = 0$	$w^b_{1,1} = 10$	$w^g_{1,1} = 10$	$w^g_{1,2} = 20$
x_2	$w^b_{2,0} = 0$	$w^b_{2,1} = 10$	$w^g_{2,1} = 10$	$w^g_{2,2} = 80$
Expected compensation		30		

Table 5. An Optimal Contract under the Accounting Mechanism with no Discretion.

R_1	(θ, a_2)			
	(b, L)	(b, H)	(g, L)	(g, H)
x_0	$\hat{w}_{0,0} = 0$	$\hat{w}_{0,1} = 10$	$\hat{w}_{1,0} = 0$	$\hat{w}_{2,0} = 10$
x_1	$\hat{w}_{1,0} = 0$	$\hat{w}_{1,1} = 10$	$\hat{w}_{1,1} = 10$	$\hat{w}_{2,1} = 20$
x_2	$\hat{w}_{2,0} = 10$	$\hat{w}_{2,1} = 20$	$\hat{w}_{2,1} = 20$	$\hat{w}_{2,2} = 60$
Expected compensation		26.67		

Table 6. An Optimal Contract under the Accounting Mechanism with Discretion.

R_1	(θ, a_2)			
	(b, L)	(b, H)	(g, L)	(g, H)
x_0	$\hat{w}_{0,0} = 0$	$\hat{w}_{0,1} = 10$	$\hat{w}_{1,0} = 0$	$\hat{w}_{2,0} = 10$
x_1	$\hat{w}_{1,0} = 0$	$\hat{w}_{1,1} = 10$	$\hat{w}_{1,1} = 10$	$\hat{w}_{2,1} = 35$
x_2	$\hat{w}_{2,0} = 10$	$\hat{w}_{2,1} = 35$	$\hat{w}_{2,1} = 35$	$\hat{w}_{2,2} = 45$
Expected compensation		29.17		

in the first period as well. The expected compensation is 26.67, which is 3.33 less than the expected compensation under the direct revelation mechanism. Out of the 3.33 difference, 0.83 is from the elimination of information rent directly. The remaining 2.5 comes indirectly from the lower bonuses needed to motivate the agent to work in the first period.

Under the accounting mechanism with discretion, the principal prefers to motivate $\omega = (truth, borrow)$ to all other possible methods.[14] The optimal contract is similar to that under the accounting mechanism with no discretion in that there is no information rent when $R_1 = x_0$, while the information rent of 10 is needed when $R_1 = x_1$ or $R_1 = x_2$. The principal reduces $\hat{w}_{2,2}$ (the bonus paid to motivate the first period input) to 45, even lower than that under the accounting mechanism with no discretion. However, $\hat{w}_{2,1}$ is set at 35, higher than that under the accounting mechanism with no discretion. This allocation of the bonuses (from $\hat{w}_{2,2}$ to $\hat{w}_{2,1}$) results from the additional ex ante constraints to guarantee that the agent prefers $\omega = (truth, borrow)$ and high inputs to all other possible combinations.

It increases the incentive costs since the first period input is motivated most efficiently by rewarding the agent only when $R_1 = x_2$. (The agent receives $\hat{w}_{2,1}$ when $R_1 = x_1$ if $\theta = g$.) Even with this additional cost, the expected compensation under the accounting mechanism with discretion is lower than that under the direct revelation mechanism. This result is stated in the following proposition.

Proposition 2. There exists a non-empty set of parameter values where the principal strictly prefers the accounting mechanism with discretion to the direct revelation mechanism.

(Optimal contracts under the accounting mechanism with discretion and conditions where the principal prefers the accounting mechanism with discretion or the direct revelation mechanism are provided in Appendix E.)

The solution to the principal's optimization problem under the accounting mechanism with discretion is given in the Appendix E, as well as the conditions where the principal prefers the accounting mechanism with discretion to the direct revelation mechanism. The following discussion is based on the optimal contracts in the Appendix E with some restrictions placed on the parameters to graphically illustrate the trade-off of constraints and the optimality of the mechanisms.

Let $h = 10$, $p_0 = q_2$, $p_1 = q_1 = (1/3)$, and $p_2 = q_0$. Fig. 2 presents the expected compensation under the direct revelation mechanism, the accounting mechanism with no discretion, and the accounting mechanism with discretion.

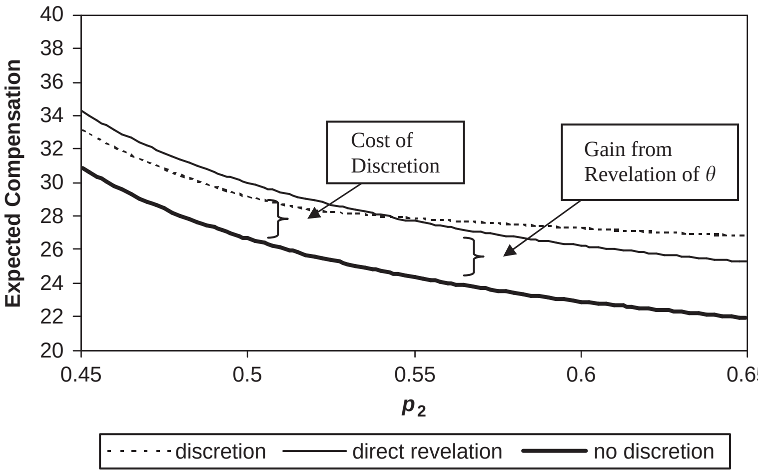


Fig. 2. Expected Compensation.

The solid line shows the expected compensation under the direct revelation mechanism. The bold solid line shows the expected compensation under the accounting mechanism with no discretion. Both lines are decreasing because the incentive problem in the first period becomes less severe as p_2 and q_0 increase. The expected compensation is always lower under the accounting mechanism with no discretion. The difference between the two lines is $(((1 - p_1)h)/2) = 3.33$ (as discussed in the numerical example), which can be disaggregated into $(p_0 h)/2$ from the direct elimination of the information rent and $(p_2 h)/2$ from the lower bonuses. Because p_1 in this case is fixed, the amount is constant. As p_2 and q_0 increase, p_0 and q_2 in turns decrease and the higher expected compensation is offset by the smaller gain from saving on the bonuses. Similarly, when p_2 and q_0 are smaller, there is less gain from saving the bonuses, but there is more gain from saving the rent.

The expected compensation under the accounting mechanism with discretion is shown by the dotted line. Similar to the optimal contract under the direct revelation mechanism, the line is decreasing because the incentive problem in the first period becomes less severe as p_2 and q_0 get larger. The difference between the dotted line and the bold solid line represents the cost of discretion, i.e., the cost of motivating the agent to adopt (*truth*, *borrow*). The cost of discretion increases as p_2 and q_0 increase. The intuition is as follows. This cost arises from the additional constraints imposing to motivate the agent to choose $\omega =$ (*truth*, *borrow*) at the beginning of the game. When p_2 and q_0 are small the first period incentive problem is relatively more severe (than the second period incentive problem), the bonus necessary to motivate the agent to work in the first period without the additional constraints is already large. The effect of adding ex ante constraints is relatively small because the original incentives were enough to accommodate the additional constraints with only a small fraction of the bonuses reallocated. On the other hand, when p_2 and q_0 are large the first period incentive problem is less severe and the bonus without the additional constraints is small. The effect of adding ex ante constraints is relatively large. When p_2 and q_0 are small enough, the principal prefers contracting under the accounting mechanism with discretion to contracting under the direct revelation mechanism.

4. ROLES OF ACCOUNTING STRUCTURAL PROPERTIES

The three structural properties play critical roles in the model. If one of the properties is absent, the incentive cost under the accounting mechanism with discretion would increase because the principal must provide incentives to the agent to adopt the appropriate method and actions. More importantly, the properties complement one another and must be present concurrently. The following example shows that absence of selective recognition or consistency results in the same outcome. Consider an agent who has chosen a successful efforts method for exploration costs at the beginning. If the current revenue is low and the agent is reluctant to report more expenses, the agent may capitalize some exploration cost related to dry wells (violation of selective recognition[15]) or the agent may switch to the full cost method (violation of consistency). The two scenarios are equivalent in the model: the agent's recognition rules are not determined by θ only, but also by R_1. The implication of this outcome is presented in the following proposition.

Proposition 3. If recognition rules are determined by (θ, R_1), the principal prefers the direct revelation mechanism to the accounting mechanism with discretion.

Selective recognition prevents the agent from conditioning his choice of recognition rules on R_1. This is important because it limits the cost of imposing additional ex ante constraints under the accounting mechanism with discretion. The principal prefers the accounting mechanism with discretion to the direct revelation mechanism when the gain from partial revelation of θ (relaxing ex post constraints) is more than the cost of discretion (imposing ex ante constraints). Allowing the agent to condition his choice on R_1 imposes a set of ex ante constraints that always dominates the relaxed ex post constraints. Therefore, the principal prefers the direct revelation mechanism to the accounting mechanism with discretion if the agent can condition his choice on R_1. For selective recognition to effectively restrict the information used, the timing of the choice is critical. The manager must decide on his accounting method choice before he learns about the future productivity. Together, the timing of the choice and selective recognition creates an ex ante commitment from the manager that limits his ability to fully manipulate the reports. For example, the manager may desire an accounting method that allocates more (less) expenses to the second period when the second period productivity is good (bad). Unless he prefers this strategy regardless of the first period outcome, he would have to commit himself ex ante to a method that he may find undesirable ex post in some state. This ex ante commitment is crucial in permitting partial revelation of the manager's private information in the model.

For the same reason, consistency is critical. The model assumes that the agent's accounting choice must be consistent. The agent may not switch to another accounting method ex post. Allowing the agent to switch to another (although accepted) accounting method choice after R_1 is known is equivalent to allowing him to condition his choice on both θ and R_1. Therefore, the principal prefers the direct revelation mechanism to the accounting mechanism with discretion if the accounting principle change is allowed. Consistency complements selective recognition in an important way. If the manager is able to change the reporting strategy through an accounting principle change when he learns new information, the ex ante commitment to the recognition rules cannot be maintained.

5. CONCLUSION

The article evaluates the effect of accounting method choices for control purpose, which is often overlooked in a cost-and-benefit analysis of an accounting choice. I examine three structural accounting properties typically found in GAAP. With these structural properties, discretionary numbers (accounting earnings) can be better than directly observed outputs (cash flows) for performance evaluation. The trade-off between the gain from information revelation and the incentive cost of discretion determines whether contracting is more efficient using observed outputs or using accounting numbers.

As accounting numbers are multi-purpose, discretion may be preferred for other reasons not present in the model. Examples include the use of audited financial statements to monitor debt contracts (Watts & Zimmerman, 1986),[16] implementation costs, and the use of accounting numbers for valuation purpose.

NOTES

1. The sum-of-the-years' digits method uses the acquisition cost, the salvage value, the expected useful life, and the number of periods remaining. The double-declining balance method uses the same set excluded the salvage value. Other information may be available but not used in determining the period's expense, e.g., the probability of impairment, current replacement cost of the assets, etc.

2. An accounting principle change is costly because the Securities and Exchange Commission (SEC) imposes considerable information requirements when a method change is made, and the financial community tends to be suspicious of such changes.

3. In accounting for business combination, if the excess of the purchase price over the fair value is assigned to goodwill, the choice decreases current earnings when

impairment occurs and increases current earnings otherwise. If the value is assigned to in-process R&D, the choice decreases earnings regardless of the impairment. However, other information (e.g., the assets' fair value) is not used in determining the effect of the choice. In accounting for software development cost, postponing the point of technological feasibility decreases current earnings while advancing it neutralizes current and future earnings (given straight-line amortization).

4. The result does not change qualitatively with an alternative model in which the manager incurs a large personal cost if he switches to a new method.

5. The literature focuses on frictions that prevent the application of the Revelation Principle to their setting. The Revelation Principle states that, in games of private information with costless communication, any equilibrium allocation which involves non-truthful communication, reporting discretion included, can also be supported in an equilibrium where truthful communication is induced (see Myerson, 1979).

6. This can be viewed as a restriction on communication. Although the agent can communicate his private information freely, the agent is not allowed to report his message space (the set of all allowable messages). A state-dependent message space does not automatically rule out the Revelation Principle. Green and Laffont (1986) show that the Nested Range Condition is necessary and sufficient for the revelation argument.

7. See Mas-Colell, Whinston, and Green (1995, p. 485).

8. Because of moving support, the second period problem can be viewed as one of hidden information alone (no hidden action). See Demski and Sappington (1984).

9. When the agent is assumed to be risk-averse with a constant absolute risk aversion (CARA) utility function, similar results obtain. Propositions and proofs are available from the author.

10. Without BC, a first-best contract is feasible by selling the firm to the agent, therefore imposing all the risk on him (Harris & Raviv, 1979). For examples of models with bankruptcy (limited liability) constraints, see Sappington (1983) and Innes (1990).

11. Examples of scenarios where the manager must use a pre-specified accounting method includes accounting for research and development costs. The Statements of Financial Accounting Standards (SFAS) No. 2 requires that R&D costs be charged to expense when incurred.

12. Examples of accounting methods that may be viewed as consistent with $\omega = (truth, borrow)$ include accounting for goodwill (SFAS 142) where the cost is expensed when the future productivity is bad (impairment) and is capitalized otherwise (no impairment).

13. Similar to the direct revelation mechanism, full communication between the principal and the agent is allowed under the accounting mechanism with discretion. However, communication of θ does not have value for the principal. The incentive cost of motivating the appropriate report of θ directly is more than the incentive cost of motivating an accounting method that helps reveal θ only partially. To see this, suppose the principal provides a contract with $\hat{w}^b_{i,j}(\hat{w}^g_{i,j})$ when the agent reports $\hat{\theta} = b(\hat{\theta} = g)$ and $\hat{R}_1 = i, \hat{R}_2 = j$. The agent always manages the direct report of θ to receive the greater of the two. Therefore, there is another contract with $\hat{w}_{i,j} = \mathrm{Max}\{\hat{w}^b_{i,j}, \hat{w}^g_{i,j}\}$.

14. The expected compensation needed to motivate high inputs from the agent is lowest when $\omega = $ (*truth*, *borrow*). See the Appendix. Note that optimal accounting method choices depend on the agent's utility function. For example, when the agent is risk averse with a CARA utility function, the principal may prefer $\omega = $ (*lend*, *borrow*) to (*truth*, *borrow*).

15. This is not a switch to the full cost method because the exploration cost related to dry wells would still be expensed if the current revenue were high.

16. Discretion reduces the probability of a breach for contracts, given that a breach is defined in terms of accounting numbers.

ACKNOWLEDGMENTS

I thank Jonathan Glover, Yuji Ijiri, Pierre Liang, Anil Arya, Hemantha Herath, Adam Koch, Evelyn Patterson, Nishi Sinha, Shyam Sunder, and workshop participants at Brock University, Carnegie Mellon University, the Ohio State University, State University of New York at Buffalo, and Yale University for their helpful comments.

REFERENCES

Antle, R., & Demski, J. S. (1989). Revenue recognition. *Contemporary Accounting Research*, *5*, 423–451.

Antle, R., Demski, J. S., & Ryan, S. (1994). Multiple sources of information, valuation, and accounting earnings. *Journal of Accounting, Auditing, and Finance*, *9*, 675–696.

Arya, A., Glover, J., & Sunder, S. (1998). Earnings management and the revelation principle. *Review of Accounting Studies*, *3*, 7–34.

Christensen, P., & Demski, J. S. (1995). Project selection and audited accrual measurement in a multi-task setting. *European Accounting Review*, *4*(3), 405–432.

Christensen, P., Demski, J. S., & Frimor, H. (2002). Auditing in agencies with moral hazard and renegotiation. *Journal of Accounting Research*, *40*, 1071–1090.

Demski, J. S. (1998). Performance measure manipulation. *Contemporary Accounting Research*, *15*, 261–285.

Demski, J. S., Patell, J. M., & Wolfson, M. A. (1984). Decentralized choice of monitoring systems. *The Accounting Review*, *59*, 16–34.

Demski, J. S., & Sappington, D. E. M. (1984). Optimal incentive contracts with multiple agents. *Journal of Economic Theory*, *33*, 152–171.

Dutta, S., & Reichelstein, S. (2005a). Stock price, earnings, and book value in managerial performance measures. *Accounting Review*, *80*(4), 1069–1100.

Dutta, S., & Reichelstein, S. (2005b). Accrual accounting for performance evaluation. *Review of Accounting Studies*, *10*, 527–552.

Dye, R. (1988). Earnings management in an overlapping generations model. *Journal of Accounting Research*, *26*, 195–235.
Dye, R., & Verrecchia, R. E. (1995). Discretion vs. uniformity: Choices among GAAP. *The Accounting Review*, *70*, 389–415.
Evans, J. H., III., & Sridhar, S. (1996). Multiple control systems, accrual accounting, and earnings management. *Journal of Accounting Research*, *34*, 45–65.
Fudenberg, D., & Tirole, J. (1995). A theory of income and dividend smoothing based on incumbency rents. *Journal of Political Economy*, *103*, 75–93.
Green, R. J., & Laffont, J. (1986). Partial verifiable information and mechanism design. *Review of Economic Studies*, *LIII*, 447–456.
Harris, M., & Raviv, A. (1979). Optimal incentive contracts with imperfect information. *Journal of Economic Theory*, *20*, 231–259.
Innes, R. (1990). Limited liability and incentive contracting with ex ante action choices. *Journal of Economic Theory*, *52*, 45–96.
Kirschenheiter, M. (1999). Optimal contracting, accounting standards, and market structures. *Contemporary Accounting Research*, *16*, 243–276.
Liang, P. J. (2000). Accounting recognition, moral hazard, and communication. *Contemporary Accounting Research*, *17*, 457–490.
Liang, P. J. (2004). Equilibrium earnings management, incentive contracts, and accounting standards. *Contemporary Accounting Research*, *21*, 685–717.
Mas-Colell, A., Whinston, M., & Green, J. (1995). *Microeconomic theory*. Oxford: Oxford University Press.
Myerson, R. (1979). Incentive compatibility and the bargaining problem. *Econometrica*, *47*, 61–74.
Sankar, M. R., & Subramanyam, K. R. (2001). Reporting discretion and private information communication through earnings. *Journal of Accounting Research*, *39*, 365–386.
Sappington, D. (1983). Limited liability contracts between principal and agent. *Journal of Economic Theory*, *29*, 1–21.
Suh, Y. S. (1990). Communication and income smoothing through accounting method choice. *Management Science*, *36*, 704–723.
Sunder, S. (1997). *Theory of accounting and control*. Cincinnati, OH: South-Western Publishing.
Watts, R. L., & Zimmerman, J. L. (1986). *Positive accounting theory*. Englewood cliffs, NJ: Prentice-Hall.

APPENDIX A. RELAXING THE ASSUMPTION $x_2 - x_1 \neq x_1 - x_0$

When the assumption $x_2 - x_1 \neq x_1 - x_0$ is relaxed, the set of accounting choices expands. The set of recognition rules is {*truth*, *borrow*, *lend*, *smooth*, *desmooth*}. "*Smooth*" allows the agent to report $\hat{R}_1 = \hat{R}_2 = x_1$ when $R_1 = x_0$, $R_2 = x_2$ and when $R_1 = x_2$, $R_2 = x_0$. "*Desmooth*" allows the agent to report $\hat{R}_1 = x_0, \hat{R}_2 = x_2$ or $\hat{R}_1 = x_2, \hat{R}_2 = x_0$ when $R_1 = R_2 = x_1$.

The principal faces more incentive compatibility constraints under the accounting mechanism with discretion because the agent chooses from 25 instead of nine accounting choices. However, the additional constraints are not so costly that the direct revelation mechanism is always preferred to the accounting mechanism with discretion. For example, Table 6 in this article shows that $\hat{w}_{2,0} = \hat{w}_{1,1} = 10$ in the optimal contract under the accounting mechanism with discretion. Thus, the optimal contract in the numerical example does not change even if the agent is allowed smoothing or desmoothing.

APPENDIX B. OPTIMAL CONTRACTS UNDER THE DIRECT REVELATION MECHANISM

The principal solves the following optimization problem:

$$\underset{W(\cdot)}{\text{Min}} \frac{1}{2}\left[\sum_{i=0}^{2} p_i w_{i,1}^{b} + \sum_{j=0}^{2} p_j w_{j,2}^{g}\right]$$

s.t.

(1) $\frac{1}{2}\left[\sum_{i=0}^{2} p_i w_{i,1}^{b} - h + \sum_{j=0}^{2} p_j w_{j,2}^{g} - h\right] - h \geq 0$

(2) $\frac{1}{2}\left[\sum_{i=0}^{2} p_i w_{i,1}^{b} - h + \sum_{j=0}^{2} p_j w_{j,2}^{g} - h\right] - h \geq \frac{1}{2}\left[\sum_{i=0}^{2} q_i w_{i,1}^{\hat{\theta}} - h + \sum_{j=0}^{2} q_j w_{j,2}^{\hat{\theta}} - h\right]$

(3) $w_{i,1}^{b} - h \geq w_{i,0}^{\hat{\theta}}$

(4) $w_{i,1}^{b} - h \geq w_{i,1}^{g} - h$

(5) $w_{i,2}^{g} - h \geq w_{i,1}^{\hat{\theta}}$

(6) $w_{i,2}^{g} - h \geq w_{i,2}^{b} - h$

(7) $w_{i,j}^{\hat{\theta}} \geq 0$

Solution: (1) is dominated by (2) and (7). $w_{i,0}^{\hat{\theta}}, w_{i,1}^{g}, w_{i,2}^{b}$ can be set to zero since they are only on the right-hand-side of the constraints. Therefore, $w_{i,1}^{b} \geq h$ and $w_{i,2}^{g} \geq w_{i,1}^{\hat{\theta}} + h$. From monotone likelihood ratio property (MLRP), $w_{2,1}^{b}$ and $w_{2,2}^{g}$ are the most efficient to satisfy (2). So we have $w_{0,1}^{b} = w_{1,1}^{b} = h$, $w_{0,2}^{g} = w_{1,2}^{g} = 2h$, any combination of $w_{2,1}^{b}$ and $w_{2,2}^{g}$ such that $w_{2,1}^{b} \geq h$, $w_{2,2}^{g} \geq w_{2,1}^{b} + h$, and $w_{2,1}^{b} + w_{2,2}^{g} = 3h + (2h/(p_2 - q_2))$.

In the above program, the principal motivates the agent to report the productivity parameter truthfully. Alternatively, the principal may motivate the agent to 1) always lie, 2) lie when the productivity is bad, tell the truth when the productivity is good, and 3) tell the truth when the productivity is bad, lie when the productivity is good.

When the agent always lies, the principal gets the same information as when he always tell the truth. The optimization problem and optimal contracts are the same as the above with reversed subscripts. When the agent always says the productivity is good (case 2) or always says the productivity is bad (case 3), the principal gets no information about the productivity. The optimization program and optimal contracts are as follows.

Let $W = w_{i,j}$ denote the compensation when $R_1 = x_i$, $R_2 = x_j$.

The principal solves the following optimization problem:

$$\underset{W(\cdot)}{\text{Min}}\frac{1}{2}\left[\sum_{i=0}^{2} p_i w_{i,1} + \sum_{j=0}^{2} p_j w_{j,2}\right]$$

s.t.

$$(1)\quad \frac{1}{2}\left[\sum_{i=0}^{2} p_i w_{i,1} - h + \sum_{j=0}^{2} p_j w_{j,2} - h\right] - h \geq 0$$

$$(2)\quad \frac{1}{2}\left[\sum_{i=0}^{2} p_i w_{i,1} - h + \sum_{j=0}^{2} p_j w_{j,2} - h\right] - h \geq \frac{1}{2}\left[\sum_{i=0}^{2} q_i w_{i,1} - h + \sum_{j=0}^{2} q_j w_{j,2} - h\right]$$

$$(3)\quad w_{i,1} - h \geq w_{i,0}$$

$$(4)\quad w_{i,2} - h \geq w_{i,1}$$

$$(5)\quad w_{i,j} \geq 0$$

Solution: (1) is dominated by (2) and (5). $w_{i,0}$ can be set to zero since they are only on the right-hand-side of the constraints. Therefore, $w_{i,1} \geq h$ and $w_{i,2} \geq w_{i,1} + h$. From MLRP, $w_{2,1}$ and $w_{2,2}$ are the most efficient to satisfy (2). So we have $w_{0,1} = w_{1,1} = h$, $w_{0,2} = w_{1,2} = 2h$, any combination of $w_{2,1}$ and $w_{2,2}$ such that $w_{2,1} \geq h$, $w_{2,2} \geq w_{2,1} + h$, and $w_{2,1} + w_{2,2} = 3h + (2h/(p_2 - q_2))$. This is the equivalent to the contract where the agent always tells the truth.

APPENDIX C. OPTIMAL CONTRACT UNDER THE ACCOUNTING MECHANISM WITH NO DISCRETION (*TRUTH*, *BORROW*) IS OPTIMAL

The contract specifies $\omega \in \Omega$. The agent receives no compensation if he does not use the specified accounting method. There are nine possible accounting methods.

For a contract that specifies $\omega = (truth, truth)$, the principal solves the following problem:

$$\underset{W(\cdot)}{\text{Min}} \frac{1}{2}[p_0(\hat{w}_{0,1} + \hat{w}_{0,2}) + p_1(\hat{w}_{1,1} + \hat{w}_{1,2}) + p_2(\hat{w}_{2,1} + \hat{w}_{2,2})]$$

s.t.

(1) $\frac{1}{2}[p_0(\hat{w}_{0,1} + \hat{w}_{0,2}) + p_1(\hat{w}_{1,1} + \hat{w}_{1,2}) + p_2(\hat{w}_{2,1} + \hat{w}_{2,2})] - 2h \geq 0$

(2) $\frac{1}{2}[p_0(\hat{w}_{0,1} + \hat{w}_{0,2}) + p_1(\hat{w}_{1,1} + \hat{w}_{1,2}) + p_2(\hat{w}_{2,1} + \hat{w}_{2,2})] - h$
$\geq \frac{1}{2}[q_0(\hat{w}_{0,1} + \hat{w}_{0,2}) + q_1(\hat{w}_{1,1} + \hat{w}_{1,2}) + q_2(\hat{w}_{2,1} + \hat{w}_{2,2})]$

(3) $\hat{w}_{0,1} - h \geq \hat{w}_{0,0}$

(4) $\hat{w}_{1,1} - h \geq \hat{w}_{1,0}$

(5) $\hat{w}_{2,1} - h \geq \hat{w}_{2,0}$

(6) $\hat{w}_{0,2} - h \geq \hat{w}_{0,1}$

(7) $\hat{w}_{1,2} - h \geq \hat{w}_{1,1}$

(8) $\hat{w}_{2,2} - h \geq \hat{w}_{2,1}$

(9) $\hat{w}_{i,j} \geq 0$

Solution: (1) is dominated by (2) and (9). $\hat{w}_{i,0}$ can be set to zero. Therefore, $\hat{w}_{i,1} \geq h$ and $\hat{w}_{i,2} \geq \hat{w}_{i,1} + h$. From MLRP, $\hat{w}_{2,1}$ and $\hat{w}_{2,2}$ are the most efficient to satisfy (2). So we have $w^b_{0,1} = w^b_{1,1} = h$, $\hat{w}_{0,2} = \hat{w}_{1,2} = 2h$, any combination of $\hat{w}_{2,1}$ and $\hat{w}_{2,2}$ such that $\hat{w}_{2,1} \geq h$, $\hat{w}_{2,2} \geq \hat{w}_{2,1} + h$, and $\hat{w}_{2,1} + \hat{w}_{2,2} = 3h + (2h/(p_2 - q_2))$. Note that the last condition is to satisfy IC. The expected compensation under this contract is $((3/2) + (p_2/(p_2 - q_2))h$.

Given $\omega = (truth, borrow)$, the principal solves the following problem:

$$\underset{W(\cdot)}{\text{Min}} \frac{1}{2}[p_0(\hat{w}_{0,1} + \hat{w}_{2,0}) + p_1(\hat{w}_{1,1} + \hat{w}_{2,1}) + p_2(\hat{w}_{2,1} + \hat{w}_{2,2})]$$

s.t.

(1) $\frac{1}{2}[p_0(\hat{w}_{0,1} + \hat{w}_{2,0}) + p_1(\hat{w}_{1,1} + \hat{w}_{2,1}) + p_2(\hat{w}_{2,1} + \hat{w}_{2,2})] - 2h \geq 0$

(2) $\frac{1}{2}[p_0(\hat{w}_{0,1} + \hat{w}_{2,0}) + p_1(\hat{w}_{1,1} + \hat{w}_{2,1}) + p_2(\hat{w}_{2,1} + \hat{w}_{2,2})] - h$
$\geq \frac{1}{2}[q_0(\hat{w}_{0,1} + \hat{w}_{2,0}) + q_1(\hat{w}_{1,1} + \hat{w}_{2,1}) + q_2(\hat{w}_{2,1} + \hat{w}_{2,2})]$

(3) $\hat{w}_{0,1} - h \geq \hat{w}_{0,0}$

(4) $\hat{w}_{1,1} - h \geq \hat{w}_{1,0}$

(5) $\hat{w}_{2,1} - h \geq \hat{w}_{2,0}$

(6) $\hat{w}_{2,0} - h \geq \hat{w}_{1,0}$

(7) $\hat{w}_{2,1} - h \geq \hat{w}_{1,1}$
(8) $\hat{w}_{2,2} - h \geq \hat{w}_{2,1}$
(9) $\hat{w}_{i,j} \geq 0$

Solution: (1) is dominated by (2) and (9). $\hat{w}_{0,0}, \hat{w}_{1,0}, \hat{w}_{0,2}, \hat{w}_{1,2}$ can be set to zero. From MLRP, $(p_2/q_2) > ((p_1 + p_2)/(q_1 + q_2))$. Therefore, $\hat{w}_{2,2}$ is the most efficient to satisfy (2). From the ex post constraints, $\hat{w}_{0,1} = \hat{w}_{1,1} = \hat{w}_{2,0} = h$ and $\hat{w}_{2,1} = 2h$. Finally, $\hat{w}_{2,2} = 3h + B$, where $B = (h/(p_2 - q_2))(2 - (p_1 - q_1) - 3(p_2 - q_2)) \geq 0$ or $B = 0$ otherwise.

The expected compensation under this contract is $(h/(2(p_2 - q_2)))(4p_2 - 2q_2 + p_2q_1 - p_1q_2)$ when $B \geq 0$ and $(h/2)(2 + p_1 + 3p_2)$ when $B = 0$. The expected compensation is greater when $B \geq 0$ and $(h/(2(p_2 - q_2)))(4p_2 - 2q_2 + p_2q_1 - p_1q_2)$ is less than $((3/2) + (p_2/(p_2 - q_2)))h$. Therefore, $\omega = (\textit{truth}, \textit{borrow})$ dominates $\omega = (\textit{truth}, \textit{truth})$.

Solving similar programs for the remaining seven possible accounting choices shows that $\omega = (\textit{truth}, \textit{borrow})$ is optimal. The expected compensation when $\omega = (\textit{truth}, \textit{lend})$ is the same as that when $\omega = (\textit{truth}, \textit{truth})$. Therefore, $\omega = (\textit{truth}, \textit{lend})$ is dominated. Then expected compensation when $\omega = (\textit{borrow}, \textit{truth})$ or $\omega = (\textit{lend}, \textit{truth})$ is $(h/(2(p_2 - q_2)))(4p_2 - 2q_2 + 3p_2q_1 - 3p_1q_2)$ which is greater than the expected compensation when $\omega = (\textit{truth}, \textit{borrow})$. The expected compensation when $\omega = (\textit{borrow}, \textit{borrow})$ or $\omega = (\textit{lend}, \textit{lend})$ is $(h/(2(p_2 - q_2)))(5p_2 - 3q_2 + 2p_2q_1 - 2p_1q_2)$ if $2 - 2\delta_1 - 4\delta_2 \geq 0$ and is $(h/2)(3 + 2p_1 + 4p_2)$ otherwise. It can be shown that this is greater than the expected compensation when $\omega = (\textit{truth}, \textit{borrow})$. Finally, the expected compensation when $\omega = (\textit{borrow}, \textit{lend})$ or $\omega = (\textit{lend}, \textit{borrow})$ is $(h/(2(p_2 - q_2)))(4p_2 - 2q_2 + 3p_2q_1 - 3p_1q_2)$ if $2 - 3\delta_1 - 3\delta_2 \geq 0$ and is $(h/2)(2{+}3p_1{+}3p_2)$ otherwise. It can be shown that this is greater than the expected compensation when $\omega = (\textit{truth}, \textit{borrow})$.

APPENDIX D

Proof of Proposition 1. Let $EW_{\mathrm{D}}\ (\omega)$ denote the expected compensation under the accounting mechanism with discretion.

$EW_{\mathrm{D}}(\omega) \geq EW(\bar{\omega})$ because the optimization problem under the accounting mechanism with discretion ω can be viewed as [P-1] with additional ex ante constraints.

$EW_{\mathrm{D}}\ (\textit{truth}, \textit{truth}) \geq EW\ (\textit{truth}, \textit{truth})$.

From Appendix B, $EW(borrow,\ borrow) = EW(lend,\ lend) \geq EW(truth,\ truth)$ $EW_D(borrow,\ borrow) \geq EW(borrow,\ borrow) \geq EW(truth,\ truth)$. ■

APPENDIX E

Proof of Proposition 2. The numerical example suffices. ■

Optimal contracts and conditions where the discretion or the direct revelation mechanism is preferred

Define. $\Omega^1 = \{(truth,\ borrow)\}$; $\Omega^2 = \{(truth,\ truth),\ (truth,\ lend)\}$; $\Omega^3 = \Omega/(\Omega^1 \cup \Omega^2)$.

Let [D (ω_b, ω_g)] denote the principal's optimization problem under the accounting mechanism with discretion $\omega' = (\omega_b, \omega_g)$.

Lemma. If $\omega \in \Omega^k$ and $\tilde{\omega} \in \Omega^k$, the players are indifferent between the choice of ω and $\tilde{\omega}$.

Proof. [D (*truth*, *truth*)] is equivalent to [D (*truth*, *lend*)]. The programs have the same set of undominated ex post constraints except that the former has the constraint $\hat{w}_{2,2} - h \geq \hat{w}_{2,1}$, whereas the latter has the constraint $\hat{w}_{2,2} - h \geq \hat{w}_{1,2}$. The programs have the same set of undominated ex ante constraints with one exception. The former has the constraint $EU\ ((truth,\ truth),\ H,\ H(\cdot),\ W(\cdot)) \geq EU\ ((truth,\ lend),\ H,\ a_2(\cdot),\ W(\cdot))$, which is equivalent to $\hat{w}_{2,2} - h \geq \hat{w}_{1,2}$. The latter has the constraint $EU\ ((truth,\ lend),\ H,\ H(\cdot),\ W(\cdot)) \geq EU\ ((truth,\ truth),\ H,\ a_2(\cdot),\ W(\cdot))$, which is equivalent to $\hat{w}_{2,2} - h \geq \hat{w}_{2,1}$. Therefore, the two programs are equivalent.

[D (*lend*, *lend*)] and [D (*borrow*, *borrow*)] are equivalent, with reversed subscripts. Also, [D (*lend*, *truth*)] is equivalent to [D (*lend*, *lend*)] except that the set of undominated ex post constraints in the former does not include $\hat{w}_{2,2} - h \geq \hat{w}_{1,2}$. However, an ex ante constraint $EU\ ((lend,\ truth),\ H,\ H(\cdot),\ W(\cdot)) \geq EU\ ((lend,\ lend),\ H,\ a_2(\cdot),\ W(\cdot))$ imposes the conditions. Similarly. [D (*borrow*, *truth*)], [D (*borrow*, *lend*)], and [D (*lend*, *borrow*)] are equivalent, with some subscripts reversed.

Finally, [D (*borrow*, *truth*)] is equivalent to [D (*lend*, *lend*)] except that the set of undominated ex post constraints in the former does not include $\hat{w}_{0,2} - h \geq \hat{w}_{0,1}$ and $\hat{w}_{2,2} - h \geq \hat{w}_{1,2}$. Theses conditions are imposed by

$EU((borrow,\ truth), H, H(\cdot), W(\cdot)) \geq EU((borrow,\ lend), H, a_2(\cdot), W(\cdot))$ and $EU((borrow,\ truth), H, H(\cdot), W(\cdot)) \geq EU((borrow,\ borrow), H, a_2(\cdot), W(\cdot))$. ■

Optimal contract under the accounting mechanism with discretion

By Lemma and Proposition 1, the principal prefers contracting under the direct revelation mechanism to the accounting mechanism with discretion if $\omega \neq (truth,\ borrow)$. Consider [D $(truth,\ borrow)$]. Define $\delta_i = p_i - q_i$. The ex post constraints are similar to those in the principal's optimization problem under the direct revelation mechanism. Therefore, $\hat{w}_{0,1} = h$. Given (bc) is satisfied, the right-hand-side of (ic)≥ 0. (ic) weakly dominates (ir).

The optimization problem reduces to the following:

$$\underset{\hat{v}_{2j}}{\text{Min}} \frac{1}{2}(p_0(\hat{w}_{2,0} + h) + p_1(\hat{w}_{2,1} + \hat{w}_{1,1}) + p_2(\hat{w}_{2,1} + \hat{w}_{2,2})).$$

s.t.

(1) $\hat{w}_{2,1} - h \geq \hat{w}_{2,0}$

(2) $\hat{w}_{2,0} - h \geq 0$

(3) $\hat{w}_{2,2} - h \geq \hat{w}_{2,1}$

(4) $p_0(\hat{w}_{2,0} - h) + p_1(\hat{w}_{2,1} - h) \geq p_0 h + p_1 \hat{w}_{1,1}$

(5) $$\frac{1}{2}[p_1(\hat{w}_{1,1} - h) + p_2(\hat{w}_{2,1} - h) + p_0(\hat{w}_{2,0} - h) + p_1(\hat{w}_{2,1} - h) + p_2(\hat{w}_{2,2} - h)] - h \geq \frac{1}{2}[q_1(\hat{w}_{1,1} - h) + q_2(\hat{w}_{2,1} - h) + q_0(\hat{w}_{2,0} - h) + q_1(\hat{w}_{2,1} - h) + q_2(\hat{w}_{2,2} - h)]$$

(6) $$\frac{1}{2}[p_1(\hat{w}_{1,1} - h) + p_2(\hat{w}_{2,1} - h) + p_0(\hat{w}_{2,0} - h) + p_1(\hat{w}_{2,1} - h) + p_2(\hat{w}_{2,2} - h)] - h \geq \frac{1}{2}[q_1(\hat{w}_{1,1} - h) + q_2(\hat{w}_{2,1} - h) + q_0 h + q_1 \hat{w}_{1,1} + q_2(\hat{w}_{2,2} - h)]$$

(7) $$\frac{1}{2}[p_1(\hat{w}_{1,1} - h) + p_2(\hat{w}_{2,1} - h) + p_0(\hat{w}_{2,0} - h) + p_1(\hat{w}_{2,1} - h) + p_2(\hat{w}_{2,2} - h)] - h \geq \frac{1}{2}[q_1 h + q_0 h + q_1 \hat{w}_{1,1} + q_2(\hat{w}_{2,2} - h)]$$

(8) $$\frac{1}{2}[p_1(\hat{w}_{1,1} - h) + p_2(\hat{w}_{2,1} - h) + p_0(\hat{w}_{2,0} - h) + p_1(\hat{w}_{2,1} - h) + p_2(\hat{w}_{2,2} - h)] - h \geq \frac{1}{2}[q_1 h + q_0(\hat{w}_{2,0} - h) + q_1(\hat{w}_{2,1} - h) + q_2(\hat{w}_{2,2} - h)]$$

(9) $\hat{w}_{1,1} - h \geq 0$

(10) $\hat{w}_{2,1} - h \geq \hat{w}_{1,1}$

The following table presents the conditions under which a set of constraints determines the optimal contract:

Case	Binding Constraints	Conditions
[134]	(1), (3), (4), (9)	$p_0 \geq \gamma_1$ and $p_0 \geq \gamma_2$
[136]	(1), (3), (6), (9)	$p_0 \leq \gamma_1$, $p_0 \leq \gamma_4$, and $p_0 \geq \gamma_3$
[137]	(1), (3), (7), (9)	$p_0 \leq \gamma_2$, $p_0 \leq \gamma_3$, and $p_0 \leq \gamma_5$
[256]	(2), (5), (6), (9)	$p_0 \geq \gamma_6$
[356]	(3), (5), (6), (9)	$p_0 \geq \gamma_4$, $p_0 \geq \gamma_7$, and $p_0 \leq \gamma_6$
[378]	(3), (7), (8), (9)	$p_0 \leq \gamma_7$, $p_0 \leq \gamma_8$, and $p_0 \geq \gamma_5$

$$\gamma_1 = \frac{(1-p_2)(1-\delta_2)}{\delta_2} \qquad \gamma_2 = \frac{(1-p_2)(3-2p_2-q_0)}{p_2+\delta_2}$$

$$\gamma_3 = \frac{(1+p_2)(1-q_0)}{q_2} - 2(2+q_1) \qquad \gamma_4 = 2(q_0+\delta_2-1) - \frac{q_0(1-p_2)}{(1-q_2)}$$

$$\gamma_5 = 2(p_2+q_0) - 3 + \frac{p_2 q_0}{(1-q_2)} \qquad \gamma_6 = 2(\delta_2-1) + \frac{q_0(3-p_2)}{(1-q_2)}$$

$$\gamma_7 = \frac{q_0((1+p_2)(1-q_0) - 2q_2(2+q_1))}{(1-q_2)(q_1-q_2)} \qquad \gamma_8 = 2(p_2+q_0) - 3 + \frac{q_0(3-p_2-q_0)}{(1-q_2)}$$

The conditions for the optimality of the discretion or the direct revelation mechanism are as follows:

a) If the contract under the accounting mechanism with discretion is case [1349], the accounting mechanism with discretion is weakly dominated by the direct revelation mechanism.
b) If the contract under the accounting mechanism with discretion is cases [1369], [2569], or [3569], the direct revelation mechanism is weakly dominated by the accounting mechanism with discretion.
c) If the contract under the accounting mechanism with discretion is case [1379] or [3789], the accounting mechanism with discretion is weakly dominated by the direct revelation mechanism, unless the following condition holds. In case [137], $p_0 \leq \bar{p}_a$. In case [3789], $p_b \leq p_0 \leq \bar{p}_b$.

$\bar{p}_a$, $\bar{p}_c$, and p_c are defined as follows:

$$\bar{p}_a = 2p_2 + \frac{2}{\delta_2} - \frac{(1+p_2)(1-q_0)}{q_2},$$

$$\bar{p}_b = 1 - p_1 - \frac{q_2(3 - q_0)}{(1 - q_0)}, \text{ and}$$

$$\bar{p}_b = \frac{q_0}{1 - q_2}\left(1 + \frac{p_2\delta_2(q_1 - q_2)}{\delta_2(1 - q_0) - 2q_2}\right). \quad \blacksquare$$

APPENDIX F

Proof of Proposition 3. Suppose the recognition rules are determined by (θ, R_1). The principal's optimization problem under the accounting mechanism with discretion is one similar to [P-2] but with the set of ex ante constraints replaced by a set of more demanding ex post constraints: $U(\omega, H, H, W(\cdot)) \geq U(\omega', H, a_2, W(\cdot))$, $\forall$ (ω', R_1, θ). Therefore, by Lemma and Proposition 1, the accounting mechanism with discretion with $\omega \in \Omega^2$ or $\omega \in \Omega^3$ is weakly dominated by the direct revelation mechanism.

Consider $\omega = (truth, borrow)$. With the original set of ex ante constraints, the program relaxes the constraint $\hat{w}_{0,2} - h \geq \hat{w}_{0,1}$. After the substitution, the condition $\hat{w}_{2,0} - h \geq \hat{w}_{0,1}$ is imposed by the constraint $U((truth, borrow), H, H, W(\cdot)) \geq U((truth, truth), H, L, W(\cdot))$. Therefore, the accounting mechanism with discretion with $\omega = (truth, borrow)$ is also weakly dominated by the direct revelation mechanism. $\blacksquare$

A NOTE ON AN ALTERNATIVE COST ALLOCATION SYSTEM FOR MUNICIPAL SERVICE UNITS

Mohamed E. Bayou

ABSTRACT

The deciding factor for operating a governmental project either as an independent, self-supporting municipal enterprise insulated from political influence or as a special revenue fund financed by tax levies is whether the amount of revenue generated covers the operating costs of the project. Cost allocation issues play an important role in this decision since the development of an acceptable user charge requires calculations of the full-cost per unit of service. If not properly understood and applied, these issues can produce unfair rates, which in turn may lead to wars between the city government and the communities it serves.

To understand the role of cost allocation in developing fair rates for a municipal enterprise's services, this article selects the most common public unity, the municipal water and sewer services, in particular, the Detroit Water and Sewer Department (DWSD). DWSD is one of the largest municipal enterprises in the United States and many of its pricing practices are typical of those followed by many cities in the United States. After presenting and illustrating the current DWSD's cost allocation and pricing procedures to highlight the unfair pricing incidents, the article proposes a modification to the current system that avoids these unfair pricing issues.

Advances in Management Accounting, Volume 17, 317–346
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017013

State and local governments often operate such municipal enterprises as public utilities, transportation systems, airports, ports, hospitals, toll bridges, parking lots, parking garages, liquor stores and public housing projects. Similar to their counterparts in the investor-owned going-concern businesses, they are expected to operate as self-supporting and even profit making entities – although not to the same extent (Hay & Wilson, 1995, p. 283). Therefore, these municipal enterprises face pricing and cost allocation issues similar to those being faced by private-sector businesses. However, these issues are more critical to municipal enterprises. As Martin (1982, p. 173) explains, the deciding factor in operating a project either as an enterprise fund, self-sustaining and isolated from political influences, or as a special revenue fund is the amount of user charge monies used to finance the project: if the amount is substantive and covers all costs, the project is operated as an enterprise; otherwise, it is operated as a special fund financed through tax levies. Therefore, for municipal enterprises' managers, proper application of a price determination process for their products and services is essential for operating in a fiscally independent, self-supporting manner (Berne & Schramm, 1986, p. 351). Complicating this deciding factor, as Stumm (1997, p. 3) argues, are the details of this revenue determination mechanism: "the true amount of revenue received as a result of a municipal activity may not be evident. Thus, municipal decisions involving these activities may be based upon incomplete or erroneous data. The results for cities in an era of fiscal stress may be disastrous."

The objective of this article is to examine the problems of determining a fair, yet acceptable user charge in one of the most common municipal enterprises, the water and sewer business. It is in the best interest of taxpayers (Hay & Wilson, 1995, p. 283) and to the city's advantage (Martin, 1982, p. 180) that the water and sewer enterprise continue to operate as a self-supporting entity. With the absence of a reliable market price, the pricing and cost allocation issues are closely related in a firm occupying a monopolistic position in its market. Since water and sewer enterprises enjoy a predominantly monopolistic advantage (with minor competition from such sources as wells and water streams), they need an intricate cost allocation mechanism to determine a fair unit price for their product. Accordingly, in a municipal enterprise, an improper cost allocation system can lead to unfair user charges.

To empirically examine how cost allocation systems in municipal service units can influence the fairness of their prices they charge to the public, the article focuses on the price determination issues in the Detroit Water and Sewer Department (DWSD). The first section of the article presents the current price issues the DWSD is struggling to solve. The second section examines and illustrates the current water rate-setting process. The purpose of this section is to highlight how the current cost allocation system can create instances of unfair prices to some communities served by the municipal enterprise. The third section presents a proposed cost allocation system that avoids the price fairness problem. Finally, summary and conclusions are presented.

FAIRNESS ISSUES OF THE CURRENT MUNICIPAL PRICE

One essential criterion of cost allocation is fairness (Horngren, Datar, & Foster, 2006, p. 495). In its mission statement, the DWSD considers establishing a fair rate for water services to about 4.3 million people in 86 wholesale communities in southeast Michigan in addition to the City of Detroit, a primary goal (DWSD Water Master Plan Technical Advisory Committee (TAC), 2003, April 28). However, the fairness of charging different rates to different communities in this region has become a serious public issue occupying major headlines in local media. For example, Gray and McDiarmid (2006, p. 1A) report that in 2006 the DWSD charged Commerce Township (#13) more than four times what it charged Ecorse (#17) (Table 1), leading customers to grumble about how the rates are set. Table 1 shows the water volume consumption and water rates for each of the 87 communities in 2006 and 2007. Ashenfelter (2007, p. 1A) explains that since the late 1970s, the debate over the water rate system has escalated to a "war" between suburban leaders and Detroit officials in which the former have accused the latter of gouging them on rates and "trying to hijack a system the city built and extended into the region."

In response to the inquiries about water rates, DWSD has offered several public presentations and issued extensive literature to explain its system, which incidentally is similar to the one several other U.S. cities use (Gray & McDiarmid, 2006, p. 14A), although its magnitude in terms of revenue collected and customers served is far more significant than others. For example, in 2005, DWSD recorded $274 million of expected revenue from four million customers (DWSD Rates, January 2006). However, while DWSD system "provides the fifth lowest rates of city-owned systems across the country" (Ashenfelter, 2007, p. 1A), the system raises several costing issues:

1. The system unitizes the annual fixed costs. Therefore, the unit fixed cost behaves as a unit variable cost. When customers decrease their demand

Table 1. Water Volumes and Water Rates statistics for 2006 and 2007.

Sl. No.	Community	Prior Year Billable Volume (2006)	Current Year Billable Volume (2007)	Change %	Water Rate		Change %
					2006	2007	
1	Allen Park	201,700	180,000	−11	7.82	8.76	12
2	Ash Township	45,700	45,000	−2	9.16	9.73	6
3	Auburn Hills	147,300	150,000	2	12.74	13.28	4
4	Belleville	22,400	21,000	−6	12.02	12.54	4
5	Berlin Township	27,300	32,000	17	11.55	12.85	11
6	Bloomfield Hills	76,700	75,000	−2	15.64	15.63	0
7	Bloomfield Township	326,300	315,000	−3	17.5	18.01	3
8	Brownstown Township	146,800	163,000	11	11.08	11.64	5
9	Canton Township	456,400	462,000	1	16.64	16.59	0
10	Center Line	46,600	45,000	−3	8	8.67	8
11	Chesterfield Township	217,600	250,000	15	10.78	11.28	5
12	Clinton Township	488,800	600,000	23	7.65	8.06	5
13	Commerce Township	100,800	115,000	14	23.29	24.28	4
14	Dearborn	1,026,300	889,000	−13	6.54	7.42	13
15	Dearborn Heights	282,500	271,000	−4	7.47	8.39	12
16	Eastpointe	165,400	161,000	−3	6.42	7.31	14
17	Ecorse	121,900	130,000	7	4.97	5.5	11
18	Farmington	65,300	67,000	3	12.43	12.5	1
19	Farmington Hills	541,200	513,000	−5	14.46	15.01	4
20	Ferndale	110,500	103,000	−7	6.45	7.66	19
21	Flat Rock	59,400	68,000	14	10.19	10.29	1
22	Flint	1,435,000	1,500,000	5	10.56	11.09	5
23	Fraser	91,400	92,000	1	8.21	8.49	3
24	Garden City	123,400	108,000	−12	8.9	10.62	19
25	Gibraltar	19,700	21,000	7	9.42	10.37	10
26	Greater Lapeer C.U.A.	122,700	118,000	−4	12.32	12.83	4

27	Grosse Ile Township	63,000	65,000	3	9.48	10.06	6
28	Grosse Pt. Park	76,900	75,000	−2	7.19	8.44	17
29	Grosse Pt. Shores	30,800	30,000	−3	10.14	10.62	5
30	Grosse Pt. Woods	100,700	100,000	−1	8.04	9.33	16
31	Hamtramck	102,100	85,000	−17	5.52	6.17	12
32	Harper Woods	88,400	78,000	−12	7.54	8.48	12
33	Harrison Township	110,900	118,000	6	8.87	9.4	6
34	Hazel Park	79,500	72,000	−9	6.92	7.84	13
35	Huron Township	67,800	71,000	5	9.92	11.33	14
36	Inkster	141,200	143,000	1	7.14	7.69	8
37	Keego Harbor	15,500	16,000	3	11.95	12.45	4
38	Lenox Twp.	16,500	16,000	−3	11.73	11.51	−2
39	Lincoln Park	182,600	195,000	7	6.66	7.53	13
40	Livonia	659,700	660,000	0	10.1	11.06	10
41	Macomb Township	431,500	460,000	7	12.3	12.5	2
42	Madison Heights	216,500	220,000	2	7.03	7.65	9
43	Melvindale	64,400	62,000	−4	7.34	7.77	6
44	Village of New Haven	14,700	15,000	2	11.72	11.04	−6
45	Northville	43,200	36,000	−17	17.42	15.67	−10
46	Northville Township	140,800	170,000	21	19.73	19.53	−1
47	Novi	303,800	305,000	0	18.11	18.86	4
48	Oak Park	146,300	143,000	−2	10.02	9.31	−7
49	Oakland County Drain Commission	3,600	4,000	11	5.24	5.66	8
50	Orion Twp (b)	164,800	165,000	0	19.42	17.95	−8
51	Plymouth	51,500	50,000	−3	10.32	11.34	10
52	Plymouth Township	159,600	151,000	−5	13.68	14.2	4
53	Pontiac	434,400	450,000	4	14.31	12.98	−9
54	Redford Township	304,200	267,000	−12	8.91	9.76	10
55	River Rouge	81,100	87,000	7	5.33	5.88	10
56	Riverview	69,500	75,000	8	8.52	9.16	8
57	Rochester Hills	406,600	433,000	6	16.66	17.18	3

Table 1. (*Continued*)

Sl. No.	Community	Prior Year Billable Volume (2006)	Current Year Billable Volume (2007)	Change %	Water Rate		Change %
					2006	2007	
58	Rockwood	15,100	16,000	6	10.17	11.08	9
59	Romeo	12,700	13,000	2	11.34	11.86	5
60	Romulus	249,600	265,000	6	8.3	8.91	7
61	Roseville	249,800	240,000	−4	7.17	7.74	8
62	Royal Oak Township	20,200	17,000	−16	8.23	8.56	4
63	SEOCWA	1,440,700	1,430,000	−1	7.59	8.21	8
64	Shelby Township	504,600	470,000	−7	14.38	14.93	4
65	South Rockwood	5,400	5,000	−7	9.87	11.67	18
66	Southgate	157,600	160,000	2	8.00	9.29	16
67	Sterling Heights	848,300	850,000	0	10.47	11.08	6
68	St. Clair County-Greenwood (b)	17,200	39,800	131	16.26	7.47	−54
69	St. Clair County-Burtchville Township	10,700	10,000	−7	14.57	15.53	7
70	St. Clair Shores	281,500	320,000	14	7.02	7.55	8
71	Sumpter Township	47,000	44,000	−6	9.67	11.13	15
72	Sylvan Lake	8,600	9,000	5	16.05	17.08	6
73	Taylor	362,200	358,000	−1	7.34	8.64	18
74	Trenton	126,000	127,000	1	7.48	8.79	18

75	Troy	639,400	630,000	−1	14.8	14.92	1
76	Utica	32,200	30,000	−7	7.86	8.46	8
77	Van Buren Township	142,700	160,000	12	12.48	13.88	11
78	Walled Lake	44,200	42,000	−5	13.86	14.29	3
79	Warren	1,028,100	1,000,000	−3	7.07	7.99	13
80	Washington Township	59,300	74,400	25	19.17	19.6	2
81	Wayne	135,500	150,000	11	8.88	9.45	6
82	West Bloomfield Township	456,400	410,000	−10	18.15	18.66	3
83	Westland	381,300	420,000	10	9.92	9.84	−1
84	Woodhaven	77,700	78,000	0	10.01	11.79	18
85	Ypsilanti Community Utilities Authority	675,200	675,000	0	8.86	9.86	11
86	Wixom	114,200	109,000	−5	15.75	15.54	−1
	Subtotal	19,384,600	19,463,200	0			
	Detroit Retail	5,525,000	5,000,000	−10	12.63	12.69	0
	Grand Total	24,909,600	24,463,200	−1.79			

Source: The Foster Group (2005, 2007).

for water, DWSD's idle capacity rises and fixed cost per unit increases, which may trigger an increase in water rates to reflect this increase.
2. Using an *annual* master (static) budget system based on the cash basis of accounting in rate setting is inconsistent with the nature of the bulk of the fixed costs that remain fixed for several years.[1] This is a problem of using a short-term instrument to allocate a long-term cost.
3. Every year, the system allocates the preceding year's *actual* variable costs of delivering water to the wholesale communities.[2] This practice can cause two unfair cost allocation issues. First, the actual cost of water delivery can change significantly from one year to the next. A deficit may arise from this practice, and the deficit may be distributed unfairly among the communities. Second, the actual cost incurred by the water supplier may be excessive due to inefficiency and lack of proper cost controls. Therefore, by increasing water rates, the costs of the inefficiencies are passed on to the wholesale communities.[3]
4. The system uses a single rate approach for computing the water rates. This approach combines fixed and variable costs and divides the lump sum by the volume of water demanded, which produces one single rate. This article explains how this approach can lead to unfair charges in certain situations.

To summarize, the current DWSD's system creates instances of unfair cost allocation among the wholesale communities. When the annual static budgeted water volume consumed by different communities exceeds the *actual* water consumption, the discrepancy may create a significant revenue deficiency. The Foster Group's report (2005, p. 9) expects the deficiency for the fiscal year 2006–2007 to reach $15.8 millions. As the DWSD (2006, January, p. 2) explains, "if the annual revenue of the water system is insufficient to meet requirements, future rates may have to be increased to make up the difference." However, when this deficiency of one year is added to the budgeted cost of the following year, some customers may end up paying more to cover the deficiency even when they do not change their volume of water consumption from one year to another. That is, their water bills subsidize other communities' water charges when the latter reduce their water consumption. For example, Table 1 shows that Livonia (#40), Novi (#47), Orion Township (#50) and Woodhaven (#84) have relatively stable water consumption over the 2006–2007 period; yet the changes in their 2007 water rates are significantly different at +10, +4, −8 and +18%, respectively.

This article applies a cost allocation system for setting water rates that do not suffer from these cost shortcomings. The proposed system, often called "dual rate," is a common cost allocation method covered in basic cost accounting textbooks (see, e.g., Garrison, Noreen, & Brewer, 2008; Horngren et al., 2006). After a brief history of DWSD's costing system, the following section explains and illustrates the current water rate-setting process. Next, the article provides an alternative system based on the dual rate approach. By illustrating the mechanism of this approach, the article shows how this approach can mitigate the fairness issues of the current rate system.

THE CURRENT WATER RATE-SETTING PROCESS

A brief history of the DWSD's water system should help explain the complexity of developing water rates.[4] In the early 1700s, the system drew water from the Detroit River by dipping pails and casks and delivered it by horse and wagon. In 1827, the system installed the first water distribution process to serve 1,500 residents at a cost of $10 per year. Currently, the system provides an average of 650 million gallons of water daily to four million residents in Southeast Michigan. Serving an area of 1,215 square miles, the DWSD system provides water to 86 communities, extending from Flint in the north to Monroe County in the south and to Ypsilanti in the west.

The DWSD, the largest municipal system in the United States, has grown into a substantial non-profit organization employing more than 2,000 employees. The City of Detroit owns the DWSD, which receives no subsidies from property tax revenues. Drawing water from the Detroit River and Lake Huron, the organization has developed three intakes to supply water to DWSD's five water treatment plants (WTPs). These plants use 22 pumping stations and 17 reservoirs to distribute water through 3,400 miles of water pipes. The magnitude and variety of these factors have made the development of water rates a complex undertaking, as explained below.

Steps of the Municipal Rate-Setting Process

DWSD supplies water to suburban communities, its wholesale customers. In turn, each community operates its own system of distributing water to its residents (the retail customers) and bills them accordingly. The water rate-setting process uses an annual static master budget, which predicts how much water will be produced and delivered to four million residents in the coming year. The system follows a "base-extra capacity" method in accordance with the approved American Water Works Association (AWWA) standards. This method recognizes the need for (a) a base service during average conditions and (b) an extra capacity to meet customer demand during peak usage periods. During the peak period, demand is substantial. For example, during the hottest day of the summer, water consumption exceeds a billion gallons, almost double a normal day's demand.

In computing the predetermined water rates for next year, the DWSD follows three steps:

Step 1: Estimating the total cost of operating, maintaining and improving the system.

Step 2: Calculating the expected volume of water for each wholesale customer. This volume is called "Units of Service."

Step 3: Calculating the unit cost of these Units of Service, which provides the basis for calculating the overall rate for each wholesale customer.

In Step 1, the DWSD divides the total cost into two categories. The first category, the Operations and Maintenance (O&M) cost, covers specific activities necessary to meet customers' demand and maintain existing system capabilities. The DWSD reviews and budgets these costs annually. The second category, Capital Improvement Program (CIP), is a long-term committed fixed cost. Reviewed and adjusted every five years, CIP covers equipment, facilities and rehabilitation projects required to meet regulatory standards and provide the needed capacity to satisfy customer demands. Most of the funds for CIP are financed through the sale of revenue bonds.

In Step 2, the DWSD applies its philosophy that "each customer should be charged according to its individual use of the system" (DWSD, 2006, p. 4). This is a cause-and-effect criterion deemed the best criterion for cost allocation in the cost accounting literature (Horngren et al., 2006, p. 495). The central element in this philosophy is the use of "Units of Service." In brief, the Units of Service consumed by a customer multiplied by their corresponding rates produce the total amount to be charged to the customer. The computation of the Units of Service and their costs for suburban wholesale customers utilize 11 Cost Functions as summarized in Table 2.[5]

The annual application of the 11 Cost Functions and Units of Service uses the following measures:

- Amount of water used in the prior year,
- Average daily water usage,
- Amount of water used in times of peak demand in prior years,

Table 2. Cost Functions the DWSD Uses to Develop Water Rates.

Sl. No.	Cost Functions	Explanation
1	Base	The amount of water used by a customer on an average day. It is calculated by dividing the estimated annual sales by 365 days.
2	Maximum day increment	The average amount of water used by a customer on the day that DWSD distributes the highest amount of water to all of its customers. This day always occurs in the summer.
3	Peak hour increment	The amount of water used by a customer over and above the maximum day usage.
4	Base distance	The distance in miles from the customer's connection(s) to the DWSD system. Since there are five water plants each capable of supplying water to most customers in the system, the distance is the average measured distance (in miles) between the customer connection(s) and each of the five WTPs.
5	Maximum day distance	Using the same distance as base distance, this cost function determines the Units of Service by multiplying the distance factor by the maximum day increment.
6	Peak hour distance	The DWSD must provide peak capacity through the distribution network to the customer. Thus, the Units of Service are computed by multiplying the distance factor by the peak hour increment.
7	Base distance–elevation	Elevation is the difference between the customer connection(s) and the average elevation of each of the five WTPs.
8	Maximum day distance–elevation	The DWSD must provide extra capacity through the distribution network to get the water to the customer at the customer's elevation. The Units of Service are computed by multiplying the distance factor by the elevation factor by the maximum day increment.
9	Peak hour distance–elevation	The DWSD must provide peak hour demand through the distribution network by supplying water at the customer's elevation. The Units of Service are computed by multiplying the distance factor by the elevation factor by the peak hour increment.
10	Customer A	The DWSD allocates the cost of a portion of its Commercial Division to this cost function. The allocation is based on the costs associated with meter reading and the provision of customer service.
11	Customer B	The DWSD allocates the *actual* cost of meter service to this cost function based on the costs associated with installation and maintenance of meters.

Source: DWSD (2006, pp. 5–7).

- Fixed measurement of the average distance from the five water plants to the community,
- Fixed measurements of the elevation average differential between a community and the water plants,
- Cost of providing customer service,
- Meter maintenance costs.

To sum, the DWSD allocates its annual budgeted costs on the basis of expected (1) base volume of water consumption, (2) peak demand and (3) other services. The next section illustrates the mechanism of this system.

Finally, in Step 3, the DWSD calculates the predetermined rate for each of the 11 Cost Functions of Table 2. First, the system adds up all of the budgeted costs of serving all of the communities. Second, it adds up the budgeted Units of Service to be supplied to all of the communities. Third, dividing the total cost by the total Units of Service produces the budgeted rate for each of the 11 Cost Functions. Fourth, the DWSD applies these rates for each of the communities by multiplying these rates by the Units of Service they demand. The result is the overall water bill for the community. Let us see how this rate system works in the following section.

The Mechanism of the Current DWSD System – An Illustration

This illustration shows how the revenue deficit of one year can affect cost allocation of the following year. Accordingly, in a community, some customers' water bills may increase to subsidize other customers' bills in other communities even if the former do not change the volume of their water consumption. This practice violates the fairness criterion of cost allocation mentioned above.

The illustration assumes that the City of Detroit's DWSD serves four communities, A, B, C and D. To simplify the analysis, the illustration ignores the maximum day and peak hour increments. Excluding these increments does not distort the implication of the current cost allocation system. The analysis covers two years, 2006 and 2007. Table 3 shows DWSD's budgeted and actual input data for 2006. The annual budgeted fixed production cost is \$10 millions. The annual variable costs consist of production cost (O&M) of \$2 millions and combined distance–elevation cost of \$5.7 millions. Several steps were undertaken to compute the total cost of transporting the water from the five WTPs to a community. First, the community's distance from the average location of the five WTPs is computed. Second, the

Table 3. 2006 Data: Production, Delivery Costs and Water Consumption.

	Budgeted	Actual	Per Unit ($)
Water volume (units)	2,000,000	1,900,000	
Costs:			
Fixed production	$10,000,000		5.00
Variable production	$2,000,000		1.00
Distance–elevation		$5,700,000[a]	3.00

Community	(1)	(2)	(3)	(4)	(5)	(7)	(8)	(9)
	Budgeted Water Consumption (Units)	Actual Water Consumption (Units)	Distance From WTP (Miles)	Elevation (Feet)	WTP Elevation (Feet)	Elevation Differential	Divide by 10.56 Conversion Factor	Add (3)+(8)
A	200,000	190,000	21.00	602	610	0[b]	0	21.00
B	600,000	600,000	31.60	738	610	128	12.12	43.72
C	800,000	700,000	23.60	617	610	7	.66	24.26
D	400,000	410,000	52.00	866	610	256	24.24	76.24
Total	2,000,000	1,900,000						

Note: One unit of water = mcf = 1,000 cubic feet, approximately 7,500 gallons.

[a]DWSD mixes past costs with budgeted figures in computing its water rates. The average of the combined distance–elevation cost of the past three years is divided by the average actual water consumption of the past three years to yield the unit cost of $3.00 per unit/mile. The actual quality of water and delivery costs of 2006 and 2005 are assumed to be equal.

[b]Community A's elevation is lower than MTP's average elevation. Since the differential cannot be a negative quantity, Community A's differential is zero.

community's differential elevation is computed by deducting the average elevation of the five WTPs from the community's elevation. Third, the differential elevation is converted from a vertical to a horizontal quantity. Finally, the community's distance and differential elevation are added to form the combined distance–elevation in miles. Table 3 also shows the budgeted and actual volumes of water consumed in 2006, distance in miles from the average WTP, elevation level in feet of each community and the average elevation of the WTPs. For example, Community A's distance from the WTP is 21 miles, its elevation is 602 feet and the WTP's elevation is 610 feet. Community A expects to demand 200,000 units (one unit is also called "mcf," which is equal to 1,000 cubic feet or approximately 7,500 gallons of water), but actually consumed only 190,000 units in 2006. Table 3 shows that the total budgeted and actual consumption by all communities is 2,000,000 units and 1,900,000 units, respectively.

The water volumes in Table 3 are assumed to reflect the existence of unaccounted-for water (UFW). TFG (2005, p. 26) defines UFW as "the difference between the volume of water measured at some predetermined delivery point(s) to the system and the volume of water measured by the various customer meters." UFW has several causes such as the inevitable loss of water in transit and limits on the accuracy of water meters and other estimating techniques. For fairness, DWSD traces these UFW losses to the communities and charges them accordingly.

Calculating the Distance–Elevation Weighted Consumption in 2006

The farther away a community area is from the water source, the higher the cost incurred to transport the water to the community. For example, Community A's 200,000 units must travel 21 miles from the source to reach the community. Similarly, the higher the elevation of the area of a community above the area of the water source (WTP), the more energy is needed to pump the water up to reach the community. In order to compute the total cost of transporting the water from WTP to a community, the community's distance and elevation must be combined. In computing the combined distance–elevation metric under the current DWSD system, a community's elevation in feet, say 738 feet, is compared to the WTP's average elevation (610 feet) to yield an increment of 128 feet. This increment is then divided by 10.56, a factor to convert a vertical to a horizontal distance to yield 12.12 (128/10.56). The converted quantity is then added to the community's distance in miles from the average WTP. Next, DWSD

Table 4. Computation of the Budgeted and Actual Distance–Elevation Weighted Water Consumption (2006).

Community	(1)	(2)	(3)	(1) × (3)	(2) × (3)
	Budgeted Water Consumption (Units)	Actual Water Consumption (Units)	Distance–Elevation (Miles)	Budgeted Weighted Consumption	Actual Weighted Consumption
A	200,000	1900,000	21.00	4,200,000	3,990,000
B	600,000	600,000	43.72	26,232,727	26,232,727
C	800,000	700,000	24.26	19,410,303	16,984,015
D	400,000	410,000	76.24	30,496,970	31,259,394
Total	2,000,000	1,900,000		80,340,000	78,466,136

divides the average of the preceding three years' total distance–elevation cost by the total distance–elevation miles to compute the unit cost for the current year, given as $3.00 per mile in Table 3. In other words, DWSD mixes past costs with budgeted figures in computing its water rates. Due to the recently significant fluctuations of energy costs, this inconsistency can lead to a significant cost misallocation. Table 4 shows the budgeted and actual distance–elevation weighted water consumption.

Calculating the Communities' Water Bills in 2006

Table 5, panel A shows the computations of each community's budgeted water bill and the total budgeted revenue of $253,020,000 for DWSD in 2006. Panel B uses the actual data from Table 4 and shows actual water charges of $246,798,409 the DWSD collected from the four communities in 2006. Since DWSD's actual revenue is less than the total budgeted revenue, Table 5 shows a total deficit of $6,221,591 in DWSD's records. This deficit must be covered by charging the communities in 2007 a higher water rate, as shown in Table 5.

Calculating the 2007 DWSD's Costs to be Allocated to the Communities

For simplicity, Table 6 shows that the budgeted fixed and variable production costs and the water transportation costs per unit in 2007 did not change from 2006. Table 6 also shows the deficit of $6,221,591 carried over from 2006.

Table 5. Budgeted and Actual Cost Allocation for the Year 2006.

Panel A: Budgeted Cost Allocation

Community	(1)	(2)	(3)	(4)	(2)+(4)
	Budgeted Water Consumption (Units)	Production Costs ($5.00+$1.00)	Budgeted Weighted Consumption	Distance–Elevation Cost (At $3.00)	Total Budgeted Cost Allocated
A	200,000	1,200,000	4,200,000	12,600,000	13,800,000
B	600,000	3,600,000	26,232,727	78,698,182	82,298,182
C	800,000	4,800,000	19,410,303	58,230,909	63,030,909
D	400,000	2,400,000	30,496,970	91,490,909	93,890,909
Total	2,000,000	$12,000,000	80,340,000	$241,020,000	$253,020,000

Panel B: Actual Cost Allocation

Community	(1)	(2)	(3)	(4)	(2)+(4)
	Actual Water Consumption (Units)	Production Costs ($5.00+$1.00)	Actual Weighted Consumption	Distance–Elevation Cost (At $3.00)	Total Cost Allocated
A	1900,000	1,140,000	3,990,000	11,970,000	13,110,000
B	600,000	3,600,000	26,232,727	78,698,182	82,298,182
C	700,000	4,200,000	16,984,015	50,952,045	55,152,045
D	410,000	2,460,000	31,259,394	93,778,182	96,238,182
Total	1,900,000	$11,400,000	78,466,136	$235,398,409	$246,798,409

Note: Total deficit in 2006 = $253,020,000−$246,798,409 = $6,221,591

For simplicity, the budgeted water consumption in 2007 is the same as in 2006. But the communities' actual water consumption in 2007 is different except for Community B's water consumption, which remains at 600,000 units each year. Communities A and D decreased their actual water consumption from 190,000 units and 410,000 units in 2006, respectively, to 180,000 units and 395,000 units in 2007, respectively. Conversely, Community C increased its actual water consumption from 700,000 units in 2006 to 780,000 units in 2007.

Calculating the Weighted Water Consumption and the Communities' Water Bills in 2007

Table 7 shows the budgeted and actual weighed water consumption in 2007. The calculation follows the same procedures used in Table 4.

Table 6. 2007 Data: Production and Transportation Costs and Water Consumption.

	Budgeted	Actual	Per Unit ($)
Water volume (units)	2,000,000	1,955,000	
Costs:			
Revenue deficit	$6,221,591		3.11
Fixed production	$10,000,000		5.00
Variable production	$2,000,000		1.00
Distance–elevation		$5,700,000	3.00

Community	(1) Budgeted Water Consumption (Units)	(2) Actual Water Consumption (Units)	(3) Distance (Miles)	(4) Elevation (Feet)	(5) MTP Elevation (Feet)	(7) Elevation Differential	(8) Divide by 10.56 conversion Factor	(9) Add (3)+(8)
A	200,000	180,000	21.00	602	610	0*	0.00	21.00
B	600,000	600,000	31.60	738	610	128	12.12	43.72
C	800,000	780,000	23.60	617	610	7	.66	24.26
D	400,000	395,000	52.00	866	610	256	24.24	76.24
Total	2,000,000	1,955,000						

Note: * Elevation differential is limited to zero as a minimum (TFG 2007, p.30).

Table 7. Computation of the Budgeted and Actual Distance and Elevation Weighted Water Consumption (2007).

Community	(1)	(2)	(3)	(1) × (3)	(2) × (3)
	Budgeted Water Consumption (Units)	Actual Water Consumption (Units)	Distance–Elevation (Miles)	Budgeted Weighted Consumption	Actual Weighted Consumption
A	200,000	1,800,000	21.00	4,200,000	3,780,000
B	600,000	600,000	43.72	26,232,727	26,232,727
C	800,000	780,000	24.26	19,410,303	18,927,045
D	400,000	395,000	76.24	30,496,970	30,115,758
Total	2,000,000	1,955,000		80,340,000	79,053,530

Table 8, panel A shows the budgeted costs allocated to the four communities. Panel B shows the actual cost allocated to the four communities. Table 8 reports that the deficit has reached $4,269,395 in 2007, which will be carried over to 2008. In addition, Community B's 2007 water bill of $84,164,659 is higher than its 2006 bill of $82,298,182 (Table 5), showing an increase of $1,866,477 even though its actual water consumption remained at 600,000 units each year:

Year	Actual Water Consumption (units)	Community B's Water bill ($)
2006	600,000	82,298,182
2007	600,000	84,164,659
Change	0	+ 1,866,477

With no changes in DWSD's production and delivery costs, Community B pays nearly $1.9 million more as a subsidy to other communities' water charges. The following section presents an alternative cost allocation system that overcomes this subsidy issue.

A PROPOSED COST ALLOCATION SYSTEM

The proposed system differs from the current DWSD's system in two respects:

1. Cost allocation: Using a dual-rate approach, the proposed system allocates fixed costs on the basis of long-term demand for water. To

Table 8. Budgeted and Actual Cost Allocation for the Year 2007.

Panel A: Budgeted Cost Allocation

Community	(1)	(2)	(3)	(4)	(2)+(4)
	Budgeted Water Consumption (Units)	Production Costs ($5.00+$1.00+$3.11)	Budgeted Weighted Consumption	Distance–Elevation Cost (At $3.00)	Total Budgeted Cost Allocated
A	200,000	1,822,159	4,200,000	12,600,000	14,422,159
B	600,000	5,466,477	26,232,727	78,698,182	84,164,659
C	800,000	7,288,636	19,410,303	58,230,909	65,519,545
D	400,000	3,644,318	30,496,970	91,490,909	95,135,227
Total	2,000,000	$18,221,591	80,340,000	$241,020,000	$259,241,591

Panel B: Actual Cost Allocation

Community	(1)	(2)	(3)	(4)	(2)+(4)
	Actual Water Consumption (Units)	Production Costs ($5.00+$1.00+$3.11)	Actual Weighted Consumption	Distance–Elevation Cost (At $3.00)	Total Cost Allocated
A	180,000	1,639,943	3,780,000	11,340,000	12,979,943
B	600,000	5,466,477	26,232,727	78,698,182	84,164,659
C	780,000	7,106,420	18,925,045	56,775,136	63,881,557
D	395,000	3,598,764	30,115,758	90,347,273	93,946,037
Total	1,955,000	$17,811,605	79,053,530	$237,160,591	$254,972,196

Note: Total deficit in 2007 = $259,241,591−$254,972,196 = $4,269,395 allocate variable cost, the system develops a budgeted cost per unit and multiplies it by the actual number of units demanded by a community.

2. Budgeted delivery cost: Rather than using a historical cost of water delivery, the proposed system uses a budgeted cost. This can be accomplished by adjusting the delivery cost per unit/mile for specific price-level changes for energy.

Table 9 shows the 2006 input data, long-term budgeted water demand, actual water consumption in 2006, distance and elevation in miles. For simplicity, the long-term and annual budgeted demand (columns 1 and 2 in Table 9) are identical. This needs not be always the case because weather, rainfall and UFW fluctuations in some years may cause the entries in these two columns to differ.

The allocation of budgeted fixed cost follows two steps. First, the average annual budgeted long-term fixed cost must be determined. Table 9 shows that this cost is about $10 million annually. Second, the long-term consumption structure (column 4 of Table 10) is computed. This structure

Table 9. The Proposed Cost Allocation Approach 2006 Data: Production and Delivery Costs and Water Consumption.

	Budgeted	Actual	Per Unit ($)
Water volume (units)	2,000,000	1,900,000	
Costs:			
Fixed production	$10,000,000		5.00
Variable production	$2,000,000		1.00
Distance-elevation		$5,700,000	3.00

Community	(1) Long-term Budgeted Consumption (Units)	(2) Annual Budgeted Consumption (Units)[a]	(3) Actual Consumption (Units)	(4) Combined Distance-Elevation (Miles)[b]
A	200,000	200,000	190,000	21.00
B	600,000	600,000	600,000	43.72
C	800,000	800,000	700,000	24.26
D	400,000	400,000	410,000	76.24
Total	2,000,000	2,000,000	1,900,000	

[a]For simplicity, the budgeted long-term and budgeted annual water consumption are assumed equal.
[b]Taken from Table 6.

Table 10. Using the Dual Rate Method to Allocate Fixed Costs for the Year 2006.

Community	(1)	(2)	(3) = (1) × (2)	(4): (3) ÷ 80,340,000	(5): (4) × 10,000,000
	Long-term Budgeted Consumption (Units)	Combined Distance–Elevation (Miles)	Long-Term Weighted Consumption	Long-Term Consumption Structure (%)	Fixed Cost Allocated ($)
A	200,000	21.00	4,200,000	5	522,778
B	600,000	43.72	26,232,727	33	3,265,214
C	800,000	24.26	19,410,303	24	2,416,020
D	400,000	76.24	30,496,970	38	3,795,988
Total	2,000,000		80,340,000	100	10,000,000

(column 4) is computed first by multiplying the long-term budgeted water consumption (column 1) by the combined distance–elevation miles (column 2) to yield the weighted water consumption (column 3). Dividing each community's weighted long-term consumption by the total amount of column 3 yields a community's share of DWSD's total fixed cost. For example, by dividing Community A's weighted long-term average water consumption of 4,200,000 units by the total amount of 80,340,000 units yields 5%. That is, Community A's share is 5% of DWSD's total budgeted fixed cost. Similarly, Communities B, C and D's shares are 33, 24 and 38%, respectively. The rationale for this treatment is explained below.

A service provider's capacity decisions often depend on the expected volume of production needed to satisfy customers' demands in the long run. Thus, when DWSD decides on the size of water plants to build and equip them with machinery, reservoirs, pumps, pipes, administrative office facilities and other infrastructure resources, it predicts the annual volume of water that it will pump to different communities during the next 10 or more years. The larger this volume over the next 10 or more years, the larger the capacity it needs to build. The fixed costs of this infrastructure should be charged to the communities regardless of their annual actual use of this capacity since the cost is committed on their behalf in advance to serve them. That is, DWSD's fixed cost is allocated on the basis of service availability rather than the actual usage of the service (Garrison et al., 2008, p. 552). The long-term demand structure (column 4 in Table 10) provides the basis for allocating the long-term (fixed) cost of DWSD's capacity. This is a cause-and-effect rationale for cost allocation, as emphasized in DWSD's mission statement, as mentioned above. For example, 38% of DWSD's long-term fixed cost is due to Community D's long-run water demand. Thus, it is fair to charge Community D by 38% of DWSD's long-term fixed cost.

The variable cost allocation follows two steps as Table 11 illustrates. First, the water delivery cost is computed by multiplying the actual water consumption by the combined distance–elevation miles to determine the actual weighted consumption (column 3, Table 11). This quantity is then multiplied by the delivery cost per unit/mile of $3.00. For example, Community A's water delivery cost is $11,970,000 (column 4, Table 11). Second, the variable production cost is computed by multiplying the actual water consumption by the variable production cost of $1.00 (column 5). Third, the total variable cost charged to each community is found by adding the delivery cost and production cost. Thus, Community A's total variable cost is $12,160,000 (column 6). Finally, the total cost results from adding the total variable cost and the allocated fixed cost (column 8).

Cost Allocation in 2007

Table 11 shows the same input data as that of Table 9. In addition, Table 12 shows the actual water consumption in 2007.

Table 13 follows similar procedures as those of Table 11 to determine the total cost charged to each community in 2007. Let us see how the results of this proposed system differ from those of the current DWSD's system, illustrated above, as explained in following section.

Analysis and Implications of Results

The proposed dual-rate system produces water rates that differ from those of the DWSD's system in several respects. First, as illustrated above, the dual-rate system reports no deficit. Hence the fairness issue raised by the current DWSD's system is avoided.[6] Table 14 compares these results. Because Community B's water consumption (600,000 units) stayed the same in each year, its allocated cost also stays the same each year at $82,563,396.

Second, the proposed system relies heavily on the communities' estimated long-run budgeted water demand in allocating fixed cost. For DWSD, getting accurate demand estimates from the communities is crucial for developing fair water rates. Developing these estimates is often contentious because a community's understated demand leads to a lower charge to that

Table 11. Using the Dual Rate Method to Allocate Variable Costs for the Year 2006.

Community	(1)	(2)	(3): (1) × (2)	(4): (3) × \$3.00	(5): (1) × \$1.00	(6): (4)+(5)	(7)	(8): (6)+(7)
	Actual Consumption Units	Combined Distance–Elevation (Miles)	Actual Weighted Consumption	Water Delivery Cost (At \$3.00)	Variable Production Cost (At \$1.00)	Total Variable Cost	Total Fixed Cost[a]	Total Cost Allocated
A	190,000	21.00	3,990,000	11,970,000	190,000	12,160,000	522,778	12,682,778
B	600,000	43.72	26,232,727	78,698,182	600,000	79,298,182	3,265,214	82,563,396
C	700,000	24.26	16,984,015	50,952,045	700,000	51,652,045	2,416,020	54,068,065
D	410,000	76.24	31,259,395	93,778,182	410,000	94,188,182	3,795,988	97,984,170
Total	1,900,000		78,466,136	\$235,398,409	\$1,900,000	\$237,298,409	\$10,000,000	\$247,298,409

[a] Taken from Table 10.

Table 12. The Proposed Cost allocation Approach 2007 Data: Production and Delivery Costs and Water Consumption[a].

	Budgeted	Actual	Per Unit ($)
Water volume (units)	2,000,000	1,955,000	
Costs:			
Fixed production	$10,000,000		5.00
Variable production	$2,000,000		1.00
Distance–elevation		$5,700,000	3.00

Community	(1) Long-term Budgeted Consumption (Units)	(2) Annual Budgeted Consumption (Units)	(3) Actual Consumption (Units)
A	200,000	200,000	180,000
B	600,000	600,000	600,000
C	800,000	800,000	780,000
D	400,000	400,000	395,000
Total	2,000,000	2,000,000	1,955,000

[a]To highlight the difference between the proposed costing system and the DWSD's current system, as illustrated above, the production fixed and variable costs and the water delivery cost in total and per unit are assumed to be the same in 2007 as in 2006.

community. If the communities are not equally biased in developing these estimates, some communities end up bearing unfairly more fixed costs than others. How can DWSD curb this problem? One solution is to follow a penalty approach by charging a significantly higher rate for the quantity that exceeds a community's pre-estimated demand.[7]

Third, the illustration of the proposed method can be easily modified to account for the *budgeted* water delivery cost. For example, in 2007, instead of using the actual cost of last year of $3.00 per unit/mile, DWSD can use a budgeted rate. By multiplying the $3.00 per unit/mile actual cost of 2006 by the change in the specific price index for energy cost, DWSD can convert the actual rate into a budgeted rate. For example, assuming that the annual change in this specific price index averages 7%, the budgeted rate per unit/mile becomes $3.21 ($3.00 × 107%).

Finally, DWSD's deficit has two kinds of sources. The first kind, resulting from uncontrollable sources, lies beyond DWSD's ability to influence such factors as inevitable losses of water in transit, inaccuracy of water meters and unexpected changes in maintenance costs. In addition, a substantial

Table 13. Using the Dual Rate Method to Allocate Variable Costs for the Year 2007.

Community	(1)	(2)	(3): (1) × (2)	(4): (3) × $3.00	(5): (1) × $1.00	(6): (4)+(5)	(7)	(8): (6)+(7)
	Actual Consumption (Units)	Combined Distance-Elevation (Miles)	Actual Weighted Consumption	Water Delivery Cost (At $3.00)	Variable Production Cost (At $1.00)	Total Variable Cost	Total Fixed Cost[a]	Total Cost Allocated
A	180,000	21.00	3,780,000	11,340,000	180,000	11,520,000	522,778	12,042,778
B	600,000	43.72	26,232,727	78,698,182	600,000	79,298,182	3,265,214	82,563,396
C	780,000	24.26	18,925,045	56,775,136	780,000	57,555,136	2,416,020	59,971,156
D	395,000	76.24	30,115,758	90,347,273	395,000	90,742,273	3,795,988	94,538,261
Total	1,955,000		79,053,530	$237,160,591	$1,955,000	$239,115,591	$10,000,000	$249,115,591

[a]Taken from Table 10.

Table 14. A Comparison between the Current DWSD's System and the Proposed System in Allocating Water Costs to Community B.

	Water Consumption (Units)	Current DWSD's System Allocated Cost[a] ($)	The Proposed System Allocated Cost[b] ($)
2006	600,000	82,298,182	82,563,396
2007	600,000	84,164,659	82,563,396
Difference	0	1,866,477	0

[a]Taken from Tables 5 and 8.
[b]Taken from Tables 11 and 13.

decrease in population of the region can increase the deficit. Witsil (2007, p. 1B) reports that between 2005 and 2006, more than 10,000 people left the Detroit Metropolitan area due to the struggling auto industry. With a decreasing number of consumers and stable fixed costs, water rates must increase under the single-rate system, as explained above. Moreover, Michigan has been recently ranked the number one state in home foreclosures in the country. As foreclosures increase, bankruptcies and bad debt losses usually increase. As Askari, Guest, and Turk (2007, p. A1) reports, foreclosure rates "have grown worse this year, with even wealthy communities like West Bloomfield seeing more residents facing removal from their homes." For fairness, the resulting deficit of these uncontrollable factors is chargeable to the communities since DWSD cannot influence them. The second kind of sources for the deficit is within DWSD influence. This source, in turn, has two kinds (a) avoidable operating inefficiencies and (b) misappropriations of financial resources. The operating inefficiencies are waste that should not be passed on to the customers. Similarly, the misappropriation of resources violates the cause-and-effect criterion of cost allocation and should not be passed on to the communities. Indeed, in May 2007, a judge ruled that the "City of Detroit must reimburse water and sewer customers across the region, including Detroit, $24 million for overcharges on a new radio system that is used mostly by the city police and firefighters" (Wisely, 2007, p. 3A). Such events may explain the reason for the "war" between the City of Detroit and the suburban communities that DWSD serves, as mentioned above. This article recommends the following procedures to minimize the effect of unfair cost allocation:

(1) An independent audit of DWSD's performance should be conducted annually to determine the magnitude of the avoidable inefficiencies, and

(2) If material, the amount of this avoidable inefficiency should be charged to a general account in the City of Detroit's records because it owns and manages DWSD.

These procedures may motivate DWSD's managers to exercise better control of its operations and increase accountability of the City of Detroit's financial activities.

SUMMARY AND CONCLUSIONS

Municipal enterprises provide important services to their communities while they help their cities' finances. In return, they enjoy a high degree of autonomy and insulation from political influences. However, these advantages hinge upon their ability to generate sufficient revenues to cover the costs of their products and services. Failure to maintain their financial self-supporting status threatens their autonomy and may even reduce them to the status of special-revenue funds financed by tax levies with all of the political ramifications that accompany this state of affairs.

The amount of their revenues is a function of their user charge, the price they charge the community for their services. In the absence of a reliable market price, and when a municipal enterprise holds a monopolistic position in supplying its services, the price is often a cost-plus type. Improper cost allocation systems to determine the cost for this price can result in incomplete or erroneous data, the result of which is disastrous for cities especially at times of financial distress (Stumm, 1997). To explain how cost allocation systems can play an important role in municipal enterprises' operations, this article selects the most common type of public utilities, the water and sewer enterprise, as a template for many municipal units. In particular, the DWSD is selected for this study due to the intensive conflicts regarding its water rates that for several years have occupied headline news in the local media.

Although DWSD is highly concerned with fairness of its water rates charged to the communities it serves, its cost allocation system produces instances of unfair charges to the communities. Its system uses a single rate approach by which fixed and variable costs are added to form one rate. The fixed cost part of this rate leads to treating fixed costs as if they are variable. Therefore, when the actual volume of water consumed is less than the budgeted volume demanded, a deficit ensues. When this deficit is carried forward to the following year to be covered, some communities end up paying more even though their actual water consumption is stable from one year to another. In effect, they subsidize other communities that decrease their water consumption.

This article illustrates how DWSD's system can result in such unfair charges. Accordingly, the article proposes an alternative cost allocation approach by which fixed and variable costs are allocated separately, each using a different rate. In this dual-rate method, commonly explained in basic cost accounting textbooks, the fixed cost is allocated on the basis of long-term expectations of water consumption by the communities. The variable cost is allocated to a community on the basis of a budgeted variable cost rate and its actual water demanded. Under this dual-rate system, the deficits resulting from the cost allocation mechanism are less than those of the current single-rate system.

However, deficits may arise from such other causes as inevitable loss of water in transit, inaccuracy of water meters, unexpected changes in maintenance costs, a decrease in population and increased bad debt due to an unexpected surge in home foreclosures and bankruptcies. Using accurate estimates and a normal capacity measure that allows for such unavoidable discrepancies may reduce the deficit. But a deficit may also arise because of DWSD's avoidable operating inefficiencies and inappropriate spending. The article concludes that in order to be fair, these controllable deficits should not be charged to the suburban communities. Instead, they should be charged to the City of Detroit since it owns and manages the DWSD activities. Such action may provide incentives to DWSD managers to control the operations of their water services when they realize that they cannot pass on this deficit to the consumers.

NOTES

1. The Foster Group (TFG), a consulting firm for DWSD, asserts in its 2006–2007 Report on Water Rates (p. 4): "Revenue requirements of the System are initially evaluated on a 'Cash Basis.' The reference to Cash Basis is important when discussing the capital costs. Under the Cash Basis the capital costs consist of the annual direct costs of financing the capital improvement program."

2. The Foster Group's report (2005, p. 20) declares: "*DWSD does not budget operating costs at the five water plants by function. Rather, actual expenses recorded by the DWSD accounting systems are utilized to allocate budgeted water plant cost to functional components. During development of the FY 2006–07 water rates, a review of actual expenses for FY 2002–03, 2003–04, and 2004–05 revealed that a smaller portion of water plant operating costs should be allocated to high lift pumping function than in water rate calculations in recent prior years*" (TFG's emphasis).

3. Some suburban leaders complain that DWSD restricts their access to monitor ''how the system operates, how costs are determined for maintenance, and how rates are developed'' (Detroit Free Press, 2007, p. 2B).

4. The source of the historical facts in this section is DWSD Rates (January 2006).

5. For Detroit residents, the rate model has four additional Cost Functions related to the supply of retail services (see www.michigan.gov/treasury).

6. A variance results when DWSD's budgeted and actual fixed costs differ. When the variance is *unfavorable* (because the actual cost exceeds the budgeted cost), the variance should be allocated to the communities based on their long-term demand structure as illustrated in Table 10. When this variance is *favorable*, the communities should be credited using the same procedures as that of the unfavorable variance.

7. Recently, DWSD has decided to invest $150 million to replace or update 275,000 residential and 3,000 business water meters in the region. The new program is expected to enable DWSD to read the water meters automatically and improve accuracy without having to send its personnel to homes and businesses to do the reading (Hackney, 2007, p. 1B). While this new program will definitely help increase accuracy of measuring *actual* water consumption, it may be ineffective in developing accurate estimates of communities' long-term water consumption. Taking the average of actual readings of, say, five years may help to provide an objective estimate of the long-term demand. However, DWSD has to wait for five years after the date of installing the new meters. And even then, a significant population change in one year (an outlier) may render the five-year average inadequate measure for that community.

REFERENCES

Ashenfelter, D. (2007). Can this man finally solve metro Detroit's water wars? *Detroit Free Press*, July 18, pp. 1A, 9A.

Askari, E., Guest, G., & Turk, V. (2007). More in danger of losing homes. *Detroit Free Press*, August 15, pp. 1A, 7A.

Berne, R., & Schramm, R. (1986). *The financial analysis of government*. Englewood Cliffs, NJ: Prentice-Hall.

Detroit Free Press. (2007). Westland: Mayor objects to water proposal. May 31, p. 2B.

Detroit Water and Sewage Department (DWSD) Technical Advisory Committee (TAC). (2003). *DWSD water master plan*. (April 28). Detroit, MI: DWSD.

DWSD. (2006). *DWSD rates, understand DWSD water rates*. (January). Detroit, MI: DWSD.

Garrison, R. H., Noreen, E. W., & Brewer, P. C. (2008). *Managerial accounting* (12th ed.). Boston: McGraw-Hill/Irwin.

Gray, K., & McDiarmid Jr., H. (2006). Water rates aren't as random as they seem. *Detroit Free Press*, April 2, pp. 1A, 14A.

Hackney, E. (2007). Updated water meters to be installed this fall. *Detroit Free Press*, August 15), pp. 1B, 9A.5B.

Hay, L. E., & Wilson, E. R. (1995). *Accounting for governmental and nonprofit entities* (10th ed.). Chicago, IL: Richard D. Irwin.

Horngren, C. T., Datar, S. M., & Foster, G. (2006). *Cost accounting, a managerial emphasis* (12th ed.). Upper Saddle River, NJ: Pearson/Prentice-Hall.

Martin, J. K. (1982). *Urban financial stress*. Boston, MS: Auburn House Publishing Company.

Stumm, T. J. (1997). Comparing alternative service delivery modes: Municipal enterprises require special consideration. *Journal of Urban Affairs*, *19*(3), 275–289.

The Foster Group (TFG). (2005). *Report on FY 2006–07 water rates*. (December). Detroit, MI: DWSD.

The Foster Group (TFG). (2007). *Report on FY 2007–08 water rates*. (January). Detroit, MI: DWSD.

Wisely, J. (2007). City must repay millions. *Detroit Free Press*, March 24, p. 3A.

Witsil, F. (2007). More are moving south. *Detroit Free Press*, April 5, p. 1B.

A RESEARCH NOTE ON CONTROL PRACTICE AND CULTURE AT ENRON

Clinton Free and Norman Macintosh

ABSTRACT

At the time of its demise in 2001, the Enron Corporation could boast of its comprehensive, state-of-the-art management control and governance systems. Yet these controls were rendered ineffective in the company's last few years. This article identifies the radical change in Enron's corporate culture that took place from the Lay-Kinder era (1986–1996) to the Lay-Skilling era (1997–2001). It argues that this was a major cause of neutralizing these controls, which in turn proved to be a major factor in Enron's fall into bankruptcy. The article draws on Schein's (1993, Legitimating clinical research in the study of organizational culture, Journal of Counselling and Development*, 71, 703–708; 1996, 2004) framework of cultural practice to develop our analysis. Thus, it supports Simon's (1990, 1995) urging to more meaningfully include corporate culture in management control research studies. The article contributes to the literature by drawing attention to the rich but untold story of Enron's governance and control and also extends the research linking corporate culture and control systems.*

Advances in Management Accounting, Volume 17, 347–382
Copyright © 2009 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 1474-7871/doi:10.1108/S1474-7871(2009)0000017014

> By mid-2000, Jeff Skilling had achieved his goal: Almost all the vestiges of the old Enron ... were gone. In its place, Enron had become a trading company. And with that change came a rock-em, sock-em, fast paced trading culture in which deals and 'deal flow' became the driving force behind everything Enron did. (Bryce, 2002, p. 215)

INTRODUCTION

The motivation for this article stems from the conundrum that during the period of Enron Corporation's spectacular rise during the latter part of the 1990s and early 2000s and its sudden demise, Enron had in place a comprehensive, state-of-the-art management control and governance system. In fact, *Fortune* magazine gave Enron its "No. 1 in Quality of Management" award in 1999. Yet these controls proved to be ineffective in Enron's last few years. This article attempts to explain how this breakdown happened and draw out some vital lessons for management control practices in general. While Enron's financial accounting "irregularities" and its "audacious" use of special purpose enterprises (SPEs) have been the focus of a vast number of academic (e.g. Abdel-khalik, 2002; Benston & Hartgraves, 2002; Healy & Palepu, 2003) and popular media articles, as well as sundry government investigations, little attention has been given to Enron's management controls and governance mechanisms.

What seems to have been overlooked is that during Richard Kinder's term as President from 1986 to 1996, Enron operated with a highly effective management control system. With CEO Ken Lay providing the inspirational leadership role, and Kinder closely monitoring business operations and cash flows, Enron could boast of its family-like, collegial corporate culture featuring respect for all employees and for financial results. Kinder left in 1996, when it appeared certain that he would not get the CEO appointment, a position he coveted. Instead, Lay and the board of directors picked Jeffrey Skilling to take over the reins as President and CEO. As one insider put it, this would prove to be the worst single mistake Lay made in his career (Bryce, 2002, p. 118). Enron's corporate culture, on Skilling's watch, underwent a change, one that rendered Enron's sophisticated governance and controls ineffective.

The article draws on Schein's (1996, 2004) ideas about corporate culture to inform our observations about Enron's management controls and to enrich the existing knowledge base concerned with the relationship between corporate culture and management control. For Schein, researchers of organizations have persistently, "underestimated the importance of culture in how organizations function" (1996, p. 229). Thus, our article also responds to Simons (1990, 1995, 2000) urging to include corporate culture in management accounting and control systems research.

The article makes two major contributions to the field of management accounting and control systems. First, it draws attention to the untold but rich story of management control practices at Enron. Second, it extends the literature and theory on culture based studies of these systems. In doing so, we relied on the vast database of research, articles, books, official inquiries and government investigative committee reports, available regarding the modus operandi of Enron's executives and managers during the company's somewhat short, and ultimately inglorious, history to document Enron's culture and its role in the management control process. Our assumption was that we can learn a lot from both success and failure experiences. The Enron saga illustrates both.

This article draws on a vast archival database of public information and insider accounts concerning Enron's development from its incorporation in 1985 until its demise in 2002. This includes accounts by former Enron employees (e.g. Cruver, 2003; Swartz & Watkins, 2003; Watkins, 2003a, 2003b, 2003c), Securities Exchange Commission (SEC) filings, Enron's Annual Reports, official reports (e.g. Joint Committee on Taxation, Senate Finance Committee, 2003; Permanent Subcommittee on Investigations of the Committee on Governmental Affairs, United States Senate, 2002), web sites, newspaper and magazine reports, journal articles, books, and other data sources that are available regarding the recent spectacular rise and fall of Enron. The status of the data should be treated as similar to that relied on by historians and the narrative that follows should be seen as a genealogical account of Enron's management control system and governance procedures during the existence of the company.[1]

The article is organized as follows. The next section provides an overview of recent research in the area of management control systems, noting the under-representation of culture within control system frameworks. This is followed by a review of Edward Schein's research into organizational culture, with its particular emphasis on the role of leadership. Section 3 presents the methodology used in this study. Then a detailed investigation into the control systems and prevailing organizational culture at Enron during (i) Richard Kinder's tenure as President from 1990 to 1996 and (ii) Ralph Skilling's tenure as President from 1996 to 2002 is presented. Drawing on this analysis, Section 5 discusses the important relationship between organizational culture and management control systems. The final section summarizes the major themes of the article and suggests avenues for future research.

MANAGEMENT CONTROL SYSTEMS

Management control systems, involving tools such as budgets, performance measures, standard operating procedures and performance-based remuneration and incentives, seek to elicit behavior that achieves the strategic objectives of an organization (Kaplan, 1984; Merchant, 1998). Recent research in management accounting and control systems has focused on several specific management control devices such as the balanced score card, risk management controls, the French *tableau de bord*, enterprise resource planning (ERP), materials requirement planning (MRP), economic value added (EVA), and value based management (VBM) systems, value chain analysis systems, relative performance evaluation benchmarking. Numerous distinctive elements of management control systems have been identified in the accounting literature (see Ouchi, 1979; Modell, 1995; Merchant, 1998; Whitley, 1999; Chenhall, 2003). Key management accounting control systems dimensions and their respective sources are presented in Table 1. As Table 1 reflects, however, there is limited consistency in the way that management control systems have been characterized in the prior literature and many researchers have focused on singular aspects and simple distinctions and taxonomies.

As Table 1 reflects, however, there is limited consistency in the way that management control systems have been characterized in the prior literature and many researchers have focused on singular, micro aspects, and simple distinctions and taxonomies. In contrast, a cultural approach allows the researcher to account for how these various control systems are mobilized in idiosyncratic ways when embedded in unique organizational cultures. Simons (1990, 1995) has repeatedly urged researchers to incorporate culture in their analyses of management control systems.

Table 1. Dominant Conceptualizations of Management Control Systems in the Accounting Literature.

Dichotomous Conceptualization	Source
Action vs Results controls	Merchant (1998) and Ouchi (1979)
Formal vs Informal controls	Whitley (1999), Merchant (1998), Modell (1995) and Amigoni (1978)
Tight vs Loose controls	Whitley (1999), Merchant (1998) and Amigoni (1978)
Restricted vs Flexible controls	Bisbe and Otley (2004)
Impersonal vs Interpersonal controls	Whitley (1999)
Diagnostic vs Interactive controls	Simons (1990, 1995) and Bisbe and Otley (2004)

In his levers of control framework, Simons (1990, 1995) offers an integrative framework of management control that has attracted the attention of both practitioners and academics, and has spawned a growing amount of research. Simons postulates that four complementary control systems – beliefs systems (core values), boundary systems (behavioural constraints), diagnostic control systems (monitoring mechanisms) and interactive control systems (involving top management support) – work together to benefit firms.

With respect to Enron, there have been numerous attempts to portray the firm's demise as the consequence of a few unethical, ''rogues'' or ''bad apples'' (the phrase used by President Bush) acting in the absence of any controls (Conrad, 2003). However, at the time of its demise, Enron featured many of the formal accoutrements of management control identified by Simons. Indeed, prior to 2001, the company was lauded as an excellent corporate citizen with exemplary ethical and control practices (Sims & Brinkmann, 2003). Numerous Harvard Business School case studies and Gary Hamel's (2000) popular book *Leading the Revolution* (Hamel was also a paid advisor to Enron) praised the flexibility and control of the Enron business model and commended it to others.[2] Enron's control regime comprised elements corresponding to each of Simons' levers of control, including its exacting formal code of ethics, an elaborate performance review regime and bonus regime, Risk Assessment and Control group (RAC) as well as the conventional powers held by the Enron Board and various Board committees (such as the audit and compensation committees). For example, a typical deal required approval from the finance department (for external funding), the portfolio management department (for portions of the deal that would remain on Enron's balance sheet), the risk management department (for approval of the customer's credit risk and of risks of price and interest-rate changes) and the legal department (for contacts and analysis of legal risks) (Hamel, 2000, pp. 213–214).

ORGANIZATIONAL CULTURE, LEADERSHIP AND MANAGEMENT CONTROL

We selected Schein's framework of social/cultural practice for three reasons. First, it has been drawn on to advantage in the organizational, administrative, and institutional theory fields where it has shown great promise to further these disciplines. As yet, however, it has drawn little attention in management accounting and control research. Second, Schein's framework seemed valuable for extending and fleshing out Simon's important but relatively undeveloped ideas relating to culture and control. Third, and more pervasively, it forefronts the role of leaders in the enculturation process (Sims & Brinkmann, 2003). This is important in the Enron case since both Kinder and Skilling, as Lay's alter ego leader partner, each championed a different organizational culture during their respective tenures as President.

In this study, culture is defined using Schein's (2004) widely cited definition:

> A pattern of shared basic assumptions that the group learned as it solved its problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way you perceive, think, and feel in relation to those problems.

Schein also argues that, "the term culture should be reserved for the deeper level of basic assumptions and beliefs that are shared by members of an organization, that operate unconsciously, and that define in a basic 'take-for-granted' fashion an organization's view of itself and its environment" (2004, p. 6).

Schein (2004) divides organizational culture into three levels. The first level relates to *visible artefacts*. These refer to visual organizational structures and processes. Artefacts have surface-level visibility which can be easily discerned (but are sometimes hard to understand). Artefacts include language, technology, products and styles (clothing, manners of address, myths and stories). The second level entails *espoused values*, that is, public rationalizations for behaviour. In Enron's case, these are espoused in its Code of Ethics. Espoused values are central to social validation and are typically determined by influential leaders and coalitions within organizations and may conceal underlying reasons. Schein's final level is *basic assumptions and values*. He argues that the core, or essence, of culture is represented by basic underlying assumptions and values, which are difficult to discern because they exist at a largely unconscious level. Schein defines basic assumptions as fundamental beliefs, values, and perceptions that:

> have become so taken for granted that one finds little variation within a cultural unit ... What I am calling basic assumptions are congruent with what Argyris has identified as "theories-in-use," the implicit assumptions that actually guide behavior, that tell group members how to perceive, think about, and feel about things ... Basic assumptions, like theories-in-use, tend to be nonconfrontable and nondebatable. (Schein, 2004, p. 18)

This article, then, mobilizes Schein's mechanisms of culture schema to analyse the radical shift in Enron's culture and its management control systems during Jeffrey Skilling's tenure as President from 1996 to 2001. It seeks to enrich the governance and control literature by downplaying the view that operating constraints and contextual factors are rigidly deterministic. Rather, we foreground the effect of leadership and organizational culture on control system design and use.

Leadership and Culture

For Schein (2004), leadership is critical to the creation and maintenance of culture; there is a constant interplay between culture and leadership. Leaders create the mechanisms for cultural embedding and reinforcement. Cultural norms arise and change because of what leaders tend to focus their attention on, their reactions to crises, their role modelling, and their recruitment strategies. Schein (2004, p. 231) outlines six primary mechanisms that are available to organizational leaders and dominant coalitions (the social network of individuals having the greatest influence on the selection of an organization's goals and strategies) to create and reinforce aspects of culture:

1. What leaders pay attention to, measure, and control on a regular basis.
2. How leaders react to critical incidents and organizational crises.
3. Criteria by which leaders allocate scarce resources.
4. Deliberate role modelling, teaching and coaching.
5. Criteria by which leaders allocate rewards and status.
6. Criteria by which leaders recruit, select, promote, retire and excommunicate organizational members.

Schein's theorizing clearly suggests a strong link between executive leadership actions and the nature of an organization's culture. Leaders' visions provide the substance of new organizational culture.

Schein (1996) stresses the need for research dealing with organizational culture, arguing for increased attention to cultural aspects of organizations. He also acknowledges that, even with rigorous study, researchers can only make statements about elements of culture rather than explicating culture in its entirety (Schein, 2004). According to Schein, deciphering an organization's culture is a highly interpretive and subjective process that requires insights into historical as well as contemporary activities. A number of first-hand descriptions and reports about Enron were available to the researchers including rich accounts of language, customs and traditions. Detailed and intimate accounts of life within Enron are now accessible through the vast and ever burgeoning archival database relating to the firm. In accessing this archive of primary and secondary accounts, we focused our analysis around accepted "rules of the game", the climate of group interaction, shared meanings and shared knowledge for socialization (paradigms, habits of thinking and acting). In addition, a number of secondary cues were available in the public sphere including published, publicly announced espoused values, and visible cultural artefacts such as corporate mission statements, corporate metaphors and important symbols. Through the interpretation and systematic comparison of first-hand narratives by key Enron employees (Cruver, 2003; Swartz & Watkins, 2003; Watkins, 2003a, 2003b, 2003c) and observers (such as Fortune business writers McLean & Elkind, 2004) as well from various espoused values and visible artefacts from archival sources, the aim is to make sense of the relationship between Enron's culture and its control systems.

One author and two research assistants coded over 5,000 documents into 53 nominal variables. This process highlighted key factors, events and influences in Enron's demise and permitted the construction of culture and enculturation matrices (Bernard, 2002). We did not regard individual accounts and fragments of data as indicative of cultural features, but rather interpreted them as part of a wider corpus of data. Our data analysis was characterized by a hermeneutic, iterative process of going back and forth from critical reflection to various data sources, and from part to whole, searching for key themes and patterns, and questioning, redefining, or buttressing the key themes and patterns identified with further evidence (Kets de Vries & Miller, 1987). Before addressing the management control issues, it will be helpful to situate these in the light some background regarding Enron's historical development.

BACKGROUND: THE ENRON CORPORATION

Enron Corporation came into being in the middle of a recession in 1985 when the Houston Gas Company merged with Internorth Inc. of Omaha Nebraska. Kenneth Lay engineered the merger as CEO of Houston Gas, and garnered several million dollars worth of stock options when Internorth paid what some thought was a handsome premium for Houston Natural Gas. The new company was born with $12.1 billion in assets, 15,000 employees, the nation's second largest pipeline network, and a towering amount of debt. It reported a first-year loss of $14 million.

The next two years were precarious ones for Enron as it teetered on the verge of bankruptcy. In spite of a hostile greenmail takeover attempt by Irwin Jacobs and a Moody's Investment Service debt rating of "highly speculative", Enron survived and in 1987 reported $5.9 billion in revenue and a $53.7 million profit from on-going operations. At the time Enron was a typical natural gas firm owning mainly hard assets such as debt-ridden pipelines and refining equipment. With assets in Texas, Oklahoma, California, Florida and Western Canada, Enron had all the traditional trappings of a highly leveraged, "old economy" firm competing in the regulated energy economy. This strategy did not excite the stock market.

The hard assets, nevertheless, proved to be the platform for building a revolutionary kind of energy company. Lay's strategic vision entailed transforming Enron into an exciting, new energy trading company, innovating new ways of servicing the rapidly expanding energy market, not only in the USA but also eventually worldwide. With his extensive experience and background as a senior executive in the gas business, vast knowledge of economics, energy technology and regulation, as well as his mastery of energy business politics (both local and in Washington), Lay was ideally suited to lead Enron into the new energy markets era.

By the year 2000, Enron was a totally different company, although still carrying a very large load of long- and short-term debt. In fact, it had become "the star of the New Economy", emerging as a paragon of the intellectual capital company with an enviable array of intangible resources. These included its political connections; its sophisticated organizational structuring; its highly skilled technical workforce of financial instruments traders specializing in the energy sector worldwide; its state of the art, on-line, computerized, web-based information system and its expert accounting knowledge at applying generally accepted accounting principles (GAAP) and minimizing income taxes. Enron had emerged as an archetype of fast track, new-knowledge companies operating in the global information economy.

Enron's emerging strategy centred around the idea that it would slowly divest its "hard assets" and shift resources into building the world's leading, sophisticated financial instruments "carpentry shop" platform for trading energy contracts in the gas, electricity, coal and hydropower markets. With Lay spearheading the movement in 1989, several large pipeline companies, the merchant bank Morgan Stanley, and a prestigious law firm, successfully put in place Skilling's idea of a Natural Gas Clearinghouse (NGC) as the vehicle for an organized but unregulated spot market in gas. In the years to come, Enron would build its empire around trading financial contracts of various kinds for gas and other energy sources. By the mid-1990s, Enron had "morphed" into a Wall Street-like investment bank, devising and trading in exotic financial instruments and financing large energy projects around the world. As one observer put it, "Enron was the child of deregulation" (Fox, 2003, p. 11). Even though Enron increasingly relied on the intellectual capital of its trading operations for most of its profits, hard assets still had an important place in Enron's strategy. The tactic was to get a physical presence in some business (such as broadband communication or water plants), learn the details of that particular business, and then build a trading operation around it. This was the context in which Richard Kinder's management controls played a vital role in Enron's early success.

The Lay-Kinder Era: 1990–1996

Ken Lay was a hands-off, removed manager who focused on the big picture. A master of public relations, he vigorously worked the corridors of power in Washington and in Houston.[3] As a personal friend of the Bush family and other powerful Houston politicians and industrialists, and with many influential Washington government contacts, Lay could successfully play the role of "Mr. Outside", while his alter ego, Richard Kinder, played the part of "Mr. Inside". As one observer put it, "Kinder ran the company ... While Lay gave speeches and posed for pictures" (Bryce, 2002, p. 114).

Kinder joined Enron's executive team in 1986 and immediately headed up an urgently needed cost reduction campaign aimed at improving Enron's cash poor and debt heavy financial situation. With a well-earned reputation for understanding operations and saving money, he performed this task very well. He oversaw Enron's various operations and drove employees at all levels to meet quarterly financial performance targets; "Kinder's job was making sure Enron worked, and he was good at it" (Bryce, 2002, p. 20). Kinder consistently paid attention to, and measured, financial performance. He demanded that business unit leaders meet their earnings target. When market analysts queried him about Enron's ability for sustained earnings growth, he replied, "Blood will flow in the streets of Houston before we miss our numbers" (Bryce, 2002, p. 116). Kinder kept his promise. During his term as President from 1990 to 1996, Enron's revenues rose from \$5.3 billion to \$13.38 billion, while reported earnings increased from \$202 million to \$584. Kinder was a master at "making the trains run on time" (Bryce, 2002, p. 112).

Kinder's control style was both people and numbers oriented. Every Monday morning, Kinder held a meeting in the boardroom, where every Enron business unit leader was "expected to show up, ready for a grilling" (Bryce, 2002, p. 111). Kinder insisted that they come prepared with their numbers, plans and strategies. He possessed an almost photographic memory and could instantly recall facts and figures of transactions even though they had been made years earlier. As one Enron manager remembered, "You could give him a budget number and explain where it came from and he'd say 'that's not what you told me last year.' And then he'd go to his desk and retrieve the year-earlier budget and prove you wrong. It was amazing" (Bryce, 2002, p. 112). As one business unit leader explained, "Rich didn't care if you had a great story. He wanted to know several things: How do you plan to make money? How do you secure your risk? And how do you assure your cash flow? It's a simple focus but it can encompass a lot of things. You could give him a one-page deal and he would pick out the one number you can't explain ... He was impossible to bullshit", and if managers "lied to him about their numbers, Rich would eat them for lunch" (Bryce, 2002, p. 112).

During these face-to-face meetings, the data supporting managers' budgets and strategic plans were debated and frequently challenged by Kinder. Such information formed the basis for a continuing agenda. Kinder also demanded up-to-the-minute reports from the business group heads who soon learned to gather and analyse important information before the meetings in anticipation of Kinder's questioning. "Kinder would sit in that room with his yellow pad and he knew every god-damned thing happening in that company", said one former 20-year Enron executive (Gruley & Smith, 2002, p. A1). Control was ongoing, dynamic and stringent. Kinder's leadership sought to instil a systematic, rational, results-oriented focus. He rewarded and promoted managers (and allocated resources to projects and deals) that survived his scrutiny.

Kinder, known throughout the company as "Doctor Discipline", understood intimately the details of every part of Enron from gas fields, to pipelines, to trading energy options, futures, swaps and derivatives. He focused hawk-like on expenses and cash flows and also kept a close eye on employee levels. In 1990 Enron employed nearly 7,000 people and reported about $200 million in profit. By 1996, Enron reported over $600 million in net income but had added only 500 more employees. Kinder continued to require managers to present him with details of their strategies, plans and proposals. Then he would zero in on any weak points and demand explanations.

Kinder personally handled all Enron's relationships with bankers, Wall Street analysts and bond rating agencies. And, in view of Enron's chronic heavy debt load, he closely monitored cash levels and cash flows. In fact, cash management was so important for Kinder that all business group managers were given a budget target for cash flow (as well as for profits) and the bonus system he orchestrated was tied to meeting both. As one executive recalled, "it wasn't enough just to get into a new business, you had to have a strategy that was going to be a natural outgrowth of your existing business. [Kinder] was a detail person. He wanted to know if there were growth areas, it had to be logical, thought out and have a good reason behind it. Kinder would bring business unit leaders who submitted overly optimistic proposals down to earth with his frequently used line, 'Let's not start smoking our own dope' " (McLean & Elkind, 2004).

Kinder's controls exhibit many of the theoretical characteristics of Simons' (1995) interpersonal, interactive management control style. Under this control style, top management pay a great deal of attention to the business managers' various control reports such as strategic plans, capital expenditure proposals, operating budgets, statistical reports on operating logistics, and financial results-and then challenge and debate the information. At Enron, Kinder scheduled regular face-to-face meetings with business unit managers to question, review, discuss and debate the information and assumptions underlying the data in such reports and documents. Operating managers, in turn, were motivated to schedule meetings with their managers and to gather and interpret additional information to use as ammunition in anticipation of having to respond to Kinder and defend their thinking and actions. The process engendered a lot of organizational learning and cross-functional communication, as new ideas emerged, and innovative rethinking ensued (Bisbe & Otley, 2004, p. 711).

Moreover, Lay and Kinder proved to be a complementary top management team. One senior executive observed, Lay "had the ability to take prima donnas and get them to hover around a common theme ... While Lay was inspiring the troops, Kinder kept the egos and the budgets in balance" (Bryce, 2002, p. 117). Elements of Enron's culture, however, were to undergo a radical change under Jeff Skilling's reign as President and CEO as he instantiated a very different management control style.

The Lay-Skilling Era: 1996–2002

Jeff Skilling held a reputation in high school as a scholarly, high achieving student but with a penchant for somewhat dangerous activities, a characteristic that resurfaced later at Enron. Turning down Princeton for his undergraduate education, he went to Southern Methodist University where he earned an applied science and business degree and then took a job at the First City National bank. Finding it boring, after two years he left and went to the Harvard Business School where he excelled as a top scholar, thriving on the highly competitive, tough-minded, give-and-take of the classroom case method discussions. Upon graduation in 1979 he joined the McKinsey & Company consulting firm in Houston, where his intellect and tenacity impressed many clients, including Ken Lay.

Skilling had a large influence on the evolution of Enron's strategic business model. While working in the 1980s as a consultant in the gas industry, Skilling noticed a paradox in the gas market. Although the demand for gas was strong and gas reserves were plentiful, the short-term demand and supply situation was chronically out of balance. Skilling proposed forming what he called a "Gas Bank" which was simply a trading ledger that facilitated buy and sell orders for long-term gas contracts. The Gas Bank allowed gas producers to enter into long-term selling contracts enabling them to rationalize their exploration and drilling programmes. At the same time, it enabled users, mainly public utilities and industrial plants, to enter into long-term buying contracts which secured their gas needs at a fixed price, thus enabling them to make plans for capital expenditures.

Lay liked the idea and in 1989, building on its large holdings of gas, Enron launched the Gas Bank. It proved an instant success, and by 1990 Enron had signed contracts with 35 producers and 50 large gas customers. The Gas Bank also made loans for new facilities to gas-fired plants. This resulted in an overall industry wide increase in the consumption of gas, which also brought more business to Enron's pipeline and trading operations. In 1990, Enron was selling 1.5 cubic feet of gas a day at an average price of $3.50 per 1,000 cubic feet and buying it for $1.20, thus netting a neat $2.30 spread. This success, along with the 1990 launching by the New York Mercantile Exchange (NYMEX) of a gas futures exchange, moved Lay to hire Skilling as head of its trading operation, Enron Finance Corporation (EFC). With Skilling leading the way, Enron's business model shifted in a few short years from a gas trading and pipeline company to become mainly a Wall Street-type financial engineering trading platform operating in financial commodities of all kinds. In 1996 (the year Kinder departed Enron), trading operations (wholesale and retail) accounted for 91 per cent of reported revenues, 54 per cent of income before tax and 62 per cent of identifiable assets. By 2000, trading operations accounted for 99 per cent of income, 88 per cent of income before tax and 80 per cent of identifiable assets, while reported revenue increased from $11,904 million in

1996 to nearly $100,000 million in 2000 – a tenfold increase. Enron had in large part morphed into a full-scale Wall Street trading corporation specializing in the financial engineering of derivatives, options and hedges.

Enron's business strategy evolved out of the Gas Bank idea and by 1996, a number of different groups were involved in its trading operations. The Physical Trading Desk dealt mainly in real gas contracts by working the spread between buy and sell orders. The Financial Trading Desk dealt in options, futures and swaps of financial contracts (but not in physical gas trades) often capitalizing on the Physical Trading Desk's proprietary knowledge of pricing spreads. And in 1992, Skilling created the Internal Research Group (including PhDs in mathematics and physics) who built highly sophisticated mathematical models to support the Financial Trading Desk's complex deals, and staffed it with experienced traders from Wall Street firms. While dealing in options, futures, swaps and derivatives was common on Wall Street, Enron was the first to bring these skills into the energy markets where it already had extensive experience and knowledge.

Management Control under Skilling

There have been many attempts to portray Enron's demise as the result of a few unscrupulous individuals acting in the absence of any formal controls (Conrad, 2003). However, it should be noted that Enron featured many of the formal trappings of management control that have been identified in prior research. Key elements of Enron's control system included its exacting formal code of ethics, elaborate performance review regime, bonus regime, RAC as well as the conventional powers held by the Enron Board and various committees.

Enron's Code of Ethics

Enron's code served as a behavioural control intended to prohibit a range of unethical behaviours. The Code stressed the following four key pillars:

Communication. We have an obligation to communicate. Here, we take the time to talk with one another ... and to listen.
Respect. We treat others as we would like to be treated ourselves. We do not tolerate abusive or disrespectful treatment.
Integrity. We work with customers and prospects openly, honestly and sincerely. When we say we will do something, we will do it; when we say we cannot or will not do something, then we won't do it.

Excellence. We are satisfied with nothing less than the very best in everything we do. We will continue to raise the bar for everyone. The great fun here will be for all of us to discover just how good we can really be. (Enron Values, Enron Annual Report, 2000)

The Code, which was to be signed and annually re-affirmed by every Enron employee, proved to be of wide interest-so much so that the political history division of the Smithsonian National Museum of American History acquired it for its permanent exhibit of exemplary business practices. However, the reality of Enron's business practices flew in the face of the Code. By mid-2006, some sixteen Enron accounting and finance managers, including CFO Andrew Fastow, had pleaded guilty to various criminal offences including fraudulent accounting practices and manipulating quarterly earnings reports and in May 25, 2006, a jury found both Lay and Skilling guilty of fraud and conspiracy. Enron's tax department, which operated as a profit centre, worked closely with large Wall Street banks, accounting firms and prestigious legal firms in creating nearly 900 offshore partnerships in tax havens. As President, Skilling made several well-documented ruthless public comments[4] and bribery and corruption were recurring features of Enron's global operations.[5] The juxtaposition between the Code and Enron's business practices did not go unnoticed by many employees. As one later reflected, ''the contrast between Enron's moral mantra and the behaviour of some of Enron's executives is bone chilling'' (Cruver, 2003, p. xii).

Enron's Performance Review System

Another vital link in Enron's management controls was the Peer Review Committee (PRC) system that Skilling devised, better known inside Enron as ''rank-and-yank''. PRC featured two basic motivational forces – fear and greed. Skilling wanted to keep only ''the very best'', meaning those who produced their profit and volume target – so every six months one or two out of every ten employees were dismissed. This weeding out process, and the fear factor it engendered, was at the heart of PRC and it worked as follows.

Each employee received a formal performance review every six months. While the employee got to select five co-workers, superiors or subordinates to give feedback to the committee, the employee's boss had to be one of the reviewers. In addition, anyone else could also provide feedback data and submit scores to the individual's supervisor if they so desired. Formal feedback categories included straightforward matters such as innovation, product knowledge, client relationship skills, intellectual curiosity, dependability, teamwork, communication, loyalty and especially revenue generation. All these data were collected on a web site for the PRC members to use when assigning a final mark from one to five to the employee under review (whose photo was displayed on a screen). The bottom 15 per cent, no matter how good they were, received a "1" which automatically meant redeployment to "Siberia", a special area where they had two weeks to try to find another job at Enron.[6] If they did not – and most did not – it was "out the door".

The process as it actually worked, not surprisingly, featured a lot of backroom negotiation. Employees being reviewed by the PRC made side deals with selected volunteer reviewers whereby they would give each other high scores. One manager described his conversation with another manager when he broached the subject of an imminent PRC as follows: "[I said] "I was wondering if you had a few minutes to talk some PRC". She replied, "Why – you want to cut a deal?" "Done", I said – and just like that we cut our deal" (Fox, 2003, p. 622). Business unit managers also made deals with each other, exchanging bad scores for both employees they wanted to dispose of, and for rivals they wanted to discredit. And, if one manager wanted to give a "5" to some employees in his group, and another manager wanted to keep all of her employees, then they would cut a deal to reach the combined 15 per cent required as above. Former employees Fusaro and Miller (2003, p. 52) argue that Enron's "rank-and-yank" machinations created "an environment where employees were afraid to express their opinions or to question unethical and potentially illegal business practices. Because the rank-and-yank system was both arbitrary and subjective, it was easily used by managers to reward blind loyalty and quash brewing dissent". The PRC was a powerful mechanism for preventing the emergence of subcultures running counter to the organizational tone set by Enron's hierarchy.

This punitive environment brought the consequences of dissent sharply into focus. Tourish and Vatcha (2005, p. 474) argue that this resulted in an "identification-with-the-aggressor syndrome" where those at the receiving end of aggression assume an aggressive posture themselves. In pitting employees against each other, the rank-and-yank system acted to stress the imagined weaknesses of individuals and to obfuscate organizational problems. In sum, this led to an erosion of employee confidence in their own perceptions and, most crucially, to further compliance with the organization's leaders in a way that strengthened conformist behaviour. The cut-throat nature of the performance evaluation thus became undiscussable: as Tourish and Vatcha (2005, p. 471) put it, the prevailing culture fostered by the PRC rendered "the undiscussability of the undiscussable also undiscussable".

The Bonus Regime

The competitiveness the PRC created was exacerbated by Enron's bonus regime, a key mechanism aimed at aligning individual and corporate goals. Under Skilling, the PRC database was used to determine employee bonuses by arraying all employees on a bell curve. This, along with how well the individual's business group had performed in producing profits and revenue, determined each employee's bonus for the period. Bonuses could range from 10 to 26 per cent of take-home pay. Not surprisingly, employees relied on several tactics to manipulate the system.

For example, traders had to calculate the forward price curves for business groups that originated long-term contracts. Sometimes they would change the price projections at the last minute before the contracts were signed in order to favour their short-term trades at the expense of the originators' long-term contracts. Another tactic (related to the fact that traders were competing for jobs and bonuses with those sitting next to them in the same trading group) involved sabotaging their neighbours' deals or even stealing their trades and positions when their neighbours left their posts. As one executive commented later, the bell curve "… had a hard Darwinian twist … [it] made a humongous difference on Enron by instilling a competitive streak in every employee" (Fox, 2003, p. 84).

Suspicion and ultra-competitive behaviours led to considerable secrecy and suspicion within Enron. This secrecy, which fostered the deceit with respect to the company's true financial position, extended inside the company to its various divisions, such that no one besides top management had any kind of picture of the financial health of the company as a whole. As Fowler (2002, p. 14) put it:

> "Every division and business unit was like its own silo, separate from all the other businesses," said the former CEO of one of the divisions. "It was decentralized and not heavy on teamwork, with all of the divisions in competition with each other for resources … But since most only saw their part of the business, they assumed the problems were isolated. You understood your piece of the business and maybe what the guy next to you did, but very few understood the big picture," a former broadband worker said. "That segmentation allowed us to get work done very quickly, but it isolated that institutional knowledge into the hands of very few people."

Risk Assessment and Control Group

Another integral part of Enron's management control system was RAC. RAC was responsible for approving all trading deals and for managing Enron's overall risk. Every deal put together by a business unit had to be described in detail in a Deal Approval Sheet (DASH), which included a description of the original information, economic data, a cash flow model, the deal's value, internal rate of return and net present value, risk component, a Financial Approval Sheet (FASH) and an authorization page. In principle, RAC had taken the place of Kinder's Monday morning control meetings with business group heads.

RAC analysts were required to conduct an independent assessment of each DASH, and deals required various levels of approval from numerous departments. As one manager described it, "Signatures were usually all over [the Authorization page], including legal, origination, accounting, finance, and RAC; and on top of that was an approval hierarchy that took big deals to the top level. Some deals were big enough to require additional approval, even from the board of directors" (Cruver, 2003, p. 81). Often, however, approval signatures came *after* the deal was transacted. A major reason for this was that the sundry business groups were under intense pressure from top management, especially Skilling, to push deals through. In 2000, nearly 1,000,000 trades were made through EnronOnline with a notional value of $336 billion. At its peak in 2001, Enron was making as many as 1,200 different types of trades worth billions of dollars of transactions every day. Margins on each trade, however, had narrowed dramatically by 2001 as competitors copied Enron's trading strategy. Where Enron's traders used to make about five cents on the dollar per trade, this dropped to one cent, largely because competitors were now copying Enron. Not surprisingly, then, Skilling's mantra became "volume! volume! volume!"

An important consequence of this, and one that would play a big role in Enron's demise, was that traders started to push through over-valued deals. Crucially, mark-to-market accounting as instantiated by Enron, the total revenue from a deal got reported at the time the deal was made. So in the case of a five-year electricity or gas supply contract, for example, the entire revenue for the five years was recorded in the first quarter. For Enron business managers, this meant they had to constantly increase their volume of deals in order to meet their quarterly budgets and collect bonuses. Even though very knowledgeable risk management personnel staffed the RAC Group, as time went by they became more and more reluctant to turn back projects that looked bad, especially since the corporate ethos held that the driving force of its business model was its ever-growing flows of deals. Rejecting them often meant "political death" for RAC members since the project proposal people could lose their bonuses and so would take revenge during the PRC process.[7] Moreover, they were not inclined to reject proposals for fear of certain repercussions from Skilling. Over-valued deals became the cultural norm rather than the exception. The high degree of subjectivity in the proposed deals, the increasing shortage of staff to assess the growing flow of deals, the push from the top for high volume deals and the twisted PRC process combined to render the RAC control process largely ineffective. The new culture led to the perversion of the RAC.

Enron's Shifting Corporate Culture under Skilling

Under Skilling's direction, then, Enron's corporate culture began to take on all the characteristics of a combative Wall Street trading organization and ultimately compromised formal controls and espoused values. As one executive later reported, "Traders are mercenaries. Their job is to kill. And mercenaries, by definition, don't have any loyalties ... With traders, it's rape, pillage, and plunder all the time. They don't care about the shareholders or the business strategy or the long-term interests of the company. They just wanted to make deals and get their bonuses" (Bryce, 2002, pp. 124–125). The Enron trading floor was said to mirror Skilling's personality – ultra competitive, highly individualistic and highly tolerant of risk. Recruitment practices supported this ethos.

Recruitment and Orientation Rituals

By the mid-1990s, Enron had established a reputation as an exciting, dynamic place to work. As it grew, it hired many of the gas industry's and Houston's most talented professionals. Employees saw Enron as different than the giant, sluggish companies where "some employees could go at half-speed and hide in the bureaucracy" (Bauder, 2002). Employees were provided with substantial autonomy and those who delivered their profit targets were handsomely rewarded with bonuses that sometimes reached one or two million dollars. Moreover, Enron had cachet for employees since it became the pride-and-joy of Houston, which had always been seen as the poor cousin of Dallas.

By the late 1990s, Skilling created his shopping list for job candidate characteristics: a very smart, sharply dressed extravert who could become a ruthless trader (Fowler, 2002). Skilling hired only the "best and the brightest" traders, investment bankers, information and computer experts, programmers, and financial engineers, most of whom were graduates of prestigious universities. As one insider commented:

> Skilling didn't really want eggheads. He wanted people like himself – ambitious, driven, self-made, with something of an edge. You had to be glib, you had to be aggressive and most of all, you had to be able to sell. You also had to have a healthy disrespect for the

> established order – how else could you keep innovating? ... In turn, Skilling engendered a kind of loyalty that, even in the early years, was almost cultlike. (Swartz & Watkins, 2003, p. 57)

As part of his Analyst and Associates' Program, Skilling would hire from 250 to 500 newly minted MBAs annually from the top business schools in the country (Zellner & Anderson, 2001). Fastow was also involved. In order to get help with his partnership deals, Fastow also hired from the elite business schools, enticing recruits with $20,000 signing bonuses and incentive plans that could double their salary. These bright and aggressive persons would be given great authority and the ability to make $5 million deals on their own (Byrne, France, & Zellner, 2002). Promotions and transfers came quickly, without providing time to learn industry details.

Employees were continually exposed to exaggerated claims about the organization. In 2000, Enron draped a huge banner at its entrance, enjoining employees to engage in the process of transforming Enron "FROM THE WORLD'S LEADING ENERGY COMPANY-TO THE WORLD'S LEADING COMPANY". Craig and Amernic (2004) have highlighted the persistence of hyperbole and hype in Enron's internal and external discourse. This extended to metaphors drawn from war, sport and extremism. Extreme wealth and ostentatious consumption was visibly available to those employees who achieved targets. On bonus day, upscale car dealers set up shop around the Enron headquarters building showing the latest most expensive Mercedes, BMWs, Aston Martins, Alpha Romeos and so on.

The nature of this indoctrination has been compared by two insiders to "a religious tract from a New Age megachurch" (Swartz & Watkins 2003, p. 103). Cruver (2003, p. 37) notes that the prevailing climate of visible conformity extended to visual artefacts such as dress and appearance.

> The first thing I noticed about Enron traders is that they all looked very similar: A goatee was fairly common; otherwise they maintained a clean-cut yet outdoorsy look; and if they didn't wear some version of a blue shirt every day, then it was like they weren't on the team ...

Swartz and Watkins (2003, p. 193) note that this also extended to language: "No one at Enron would ever 'build consensus,' they would 'come to shore,' as in 'We have to come to shore on this' ... Everyone [used] the term 'metrics' and anyone who used the term 'numbers' or 'calculations' was a 'loser', the most popular Enron label of all". Enron's socialization process was referred to "Enronizing" with people who didn't fit in called "losers", "damaged goods" or "shipwrecks" (Roberts & Thomas, 2002).

Performance Targets, Stock Options and Mark-to-Market Accounting

When Skilling came on board, he negotiated a lucrative employment contract that included salary and substantial bonuses based on the performance of Enron Finance Corporation (later the Enron Gas Services Group). Between October 1998 and November 2001, he sold 1,307,678 Enron shares with a gross proceeds value of $70,687,199. Another vital plank in Enron's strategy was the move, at Skilling's request, to switch its corporate accounting from traditional *historical cost* to *mark-to-market* accounting, a method that was already in widespread use throughout the banking and finance industries. In the second half of the 1990s, Enron would rely heavily on mark-to-market accounting, along with the use of SPEs, to "massage" its reported quarterly and annual earnings up or down as need be in order to meet analysts' earnings expectations.

While mark-to-market accounting provided better asset values for its contracts, more importantly, it permitted recording profits from long-term deals immediately rather than, as for traditional accounting, at the culmination of the contract. This meant that profits got recorded in the quarter in which the deals were signed, even for 20-year contracts. This had the effect of emphasizing short-term results since Enron's financial traders now had to start each quarter with a blank trading book and a new profit target. For Enron to continue to increase reported earnings at its current rate, an ever greater volume of deals was necessary. This put even more pressure on the traders for short-term output. In the words of one Enron executive, "you put yourself in a position where you had to kill to eat" (Fox, 2003, p. 42).

Enron's accounting manoeuvres, much publicized later, were part and parcel of the egregious violations of its much-vaunted Code of Ethics. The majority of the ever-increasing flow of deals demanded by Skilling involved over-the-counter trades of sophisticated, long-term derivative-like contracts woven into a confusing web of subsidiaries, SPEs, investment bankers' loans and complex accounting entries. This meant that the mark-to-market accounting for them was up to Enron's traders to determine the market prices, which they "manufactured" according to complex Black and Scholes-type valuation models which involved a host of subjective assumptions about their long-term risk and volatility parameters. And given that the traders "... were continually pressured to meet targets and show 'as much profit' as possible ... Enron had the freedom (and the talent) to book trades at the *extreme edge* of what they could get away with" (Cruver, 2003, p. 274). The new culture encouraged dubious accounting practices for these deals.

Manipulation and Avariciousness

In 1999 and 2000, Enron electricity traders, along with those in other companies like Dynergy, Reliant Resources and CMS Energy, actively engaged in a variety of highly dubious schemes to drive up electricity prices in California and reap huge profits for themselves. One such tactic, ‘‘round-tripping’’, involves one trader selling electricity contracts to another trader (either internally or externally) at a ‘‘set’’ price, while the latter simultaneously sells the same electricity back to the former at the same price. The effect is to give the impression that demand has soared thus driving prices up since these trades establish the price of the last trade in the market, setting an artificially high benchmark price for the next regular trade.[8] Enron traders, using their own proprietary computerized trading networks, on one occasion round-tripped more than 11 million megawatts of electricity, making nearly 98 per cent of its trades to other Enron groups at prices spiking to $2,500 a megawatt hour.

Enron also engaged in ‘‘congesting’’ by deliberately overloading specific power lines. For example, its traders targeted the California Power Exchange (CPE). CPE had managed the State's electricity market and posed a large threat to Enron's trading operations in California. In May 1999, Enron traders submitted a bid to CPE for 2,900 megawatts on the transmission line running from the central California valley to San Diego that had a capacity to handle only 15 megawatts. This overload shut the line down, thus artificially boosting demand and driving prices up. In another version of congesting, one that Enron traders referred to as the ‘‘Death Star’’ project, Enron scheduled power movements over lines they knew were overloaded and collected ‘‘congestion payments’’ for calling them off with the result that Enron got paid for moving energy to relieve congestion without actually moving any energy or relieving congestion. Finally, the ‘‘Get Shorty’’ project involved falsifying information regarding standby inventory of energy, and then collecting payments for standby power generation capacity they did not have, but rather would buy later at lower prices to cover their obligations if and when needed.

Such dubious tactics, carried out by other energy firms as well, eventually drove CPE into bankruptcy leaving Enron as the key player in California. With electricity prices soaring, and in the wake of the infamous California ‘‘black-out’’, the California State Government, led by Governor Davis, legislated a cap on prices. Seizing the moment, Enron and other traders exploited this opportunity by buying California electricity at the capped price and selling it to other states that had no price caps at a large margin of profit and at the same time exacerbating the California shortage. This, along with round-tripping and sham congestion, at one point caused the price of electricity for public utilities to rise from $30 per kilowatt-hour to a spike of $1,500 per hour. In the face of severe criticism from California politicians and government officials who were accusing Enron of price gouging and unfair business practices that contributed to, if not caused, the black-outs, Enron accountants shifted large amounts of profits ($1.5 billion according to Fox, 2003, p. 220) to an account for ‘‘Future Contingencies’’ in order to deflate reported quarterly earnings. While such a practice is considered to be dubious accounting in many quarters, both Lay and Skilling admitted they were aware of the reserves but considered them proper in that they did not directly violate GAAP.

The increasingly instrumental and aggressive nature of Enron's corporate culture was also manifested in Enron traders' ‘‘trash talk'' during the California electricity crisis and black-outs of summer 2001 (recorded in taped transcripts and email files published by the Federal Energy Regulatory Commission in 2004). In a recorded conversation, Enron traders gloated about how much money they took from ‘‘Grandma Millie'' in California. Another reveals plots to deliberately drive prices up by shutting down power plants, ‘‘If you took the steamer down, how long would it take to get it back up?'' The reply was, ‘‘Oh it's not something you want to just be turning on and off every hour. Let's put it that way''. The reply, ‘‘Well, why don't you just go ahead and shut her down''.

Thus, Enron proved to be a key player in creating artificial energy shortages and manipulating prices in California during 1999 and 2000. For Enron this meant profits of nearly $46 billion, according to some estimates. Yet the means of accomplishing this concerned some Enron employees. One Enron manager later put it this way:

> The contrast between Enron's moral mantra (as stated in its Code of Ethics manual) and the behaviour of some of Enron's executives is bone chilling. Indeed, the Enron saga teaches us the limitations of corporate codes of ethics: how empty and ineffectual they can be. Long touted as crucial accoutrements to moral rectitude, codes are useless when words are hollow – when executives lack either the dedication to espoused values or the ability to make defensible ethical decisions. (Cruver 2003, p. xii)

Former Enron employees state that by 2001, ‘‘Enron had become less a company than a collection of mercenaries'' (Streitfeld & Romney, 2002, p. 2) and that ‘‘there wasn't anything they wouldn't try to make money'' (Streitfeld & Romney 2002, p. 7). This mercenary posture extended beyond national borders. In India, Enron executives paid local law enforcement officers to suppress legitimate and peaceful opposition to its power plant near Mumbai (Tourish & Vatcha, 2005). Even in the company's final hours, remaining executives furiously grasped at what remained. Watkins (2003c, p. 436) recalls

> In January, we all found out that a handful of executives paid themselves gargantuan retention bonuses the week before bankruptcy. There was an $8 million amount, a $5 million amount, a couple $1.5 million, lots of $900,000 and $700,000 ... And then they had the gall to stand up at floor meetings and tell handfuls of people, you know, "last Friday was your last paycheck, I'm so sorry, this is so heart wrenching for me," you know, with all that money sitting in their pockets. So something went horribly, horribly wrong with the culture.

Obscene Compensation

Compensation was another powerful shaper and emblem of Enron culture. Compensation plans were designed with one purpose in mind: to enrich the executives, not to enhance profits or increase shareholder value (McLean, 2001). "Those who closed major deals were paid up to 3% of the value of the entire deal, payable when it was struck, not when the project actually began earning money" (Fowler, 2002). Traders could earn as much as $1 million annually (Coy & Anderson, 2002). In the Energy Services (ES) Division, for example, executive bonuses were tied to the values of the deals struck. But under the circumstances, the value of the deal had to be estimated; the incentive plan thus used a market valuation formula for the estimate that was provided by the person making the deal. Eventually, the inflation in deal value spawned by the bonus program at ES was dropped (Fowler, 2002). For stock option incentives, instead of the usual fixed waiting or vesting period, Enron added the option that if profits and stock prices rose enough, the vesting schedules would be rapidly increased, meaning the executives could get their hands on the stock more quickly (Barnes, Barnett, & Schmitt, 2002). One employee recalled recommending trying to win a lucrative energy management contract for a public school district, noting that it might take a year or more to bring the project to fruition. The manager rejected the idea because it would take too long; he needed projects that could be done in three months or less for bonus purposes (Fowler, 2002).

Perks and rewards were lavish and flowed with champagne. Skilling handed out large pay cheques, bonuses and stock options to traders who successfully met their earnings targets; in 1999, Enron granted 93.5 million stock options compared with 25.4 in 1996. John Arnold, a gas trader, booked $700 million in 2001, took his $15 million bonus and left Enron. John Pai cashed $250 million in Enron stock over three years. As one observer put it, "the excess was obscene. We were just pissing money away" (Bryce, 2002, p. 134).[9] Six analysts were flown to Colorado for a skiing weekend after closing one deal (Barnes et al., 2002). Lay and Skilling's expansion of Enron's fleet of jets also took its toll on the company's cash position.[10]

Moreover, these executives manipulated quarterly earnings announcements. At the 2006 trial of Lay and Skilling, prosecutors charged them with making statements during 2000 intended to mislead Wall Street players about the true conditions of Enron's financial earnings and its financial position. Evidence emerged during the trial that Skilling had requested last minute changes to quarterly earnings per share (EPS) releases so that Enron's figures would "meet or beat" the analysts' consensus figure. According to testimony at the trial, Enron's investor relations executives including Koenig, under instructions from Richard Causey, Enron's chief accounting officer, carried out such orders. Paula Rieker, manager for investor relations, who helped write the "scripts" for such releases, testified that in January 2000, Enron officers were prepared to report quarterly EPS of 30 cents to match the analysts' consensus number (Barrionuevo, 2006). Just before Enron's conference call to analysts, she said, the analysts' consensus had risen by one cent to 31 cents and Mark Koenig, executive Vice President for investor relations, informed her that Skilling and Causey had decided to change the numbers to meet the new consensus. Accordingly, Wesley and Colwell, chief accountant of Enron's wholesale energy trading unit, transferred 7 million dollars to a profit account from a reserve contingency account set up in a prior period as a reserve for possible future contract settlements. Consequently, Rieker "modified" the news release to report 31 cents.[11]

The Subversion of Formal Management Control Systems under Skilling

Thus, as enumerated, Enron featured many of the formal accoutrements of management control that have been identified in prior research including the elaborate performance review regime, bonus regime, RAC as well as the conventional powers held by the Enron Board. Perhaps the centrepiece of the formal control regime was the company's emphatic Code of Conduct. Yet the reality of Enron's business practices flew in the face of the Code. Moreover, Enron's board of directors waived compliance with the Code on a number of occasions to permit conflict of interest transactions with SPEs.

In practice, the PRC system worked to encourage "entourages" or "fiefdoms" (Dallas, 2003) of loyal employees who gravitated towards powerful players for protection. The PRC was a powerful mechanism for preventing the emergence of subcultures running counter to the organizational tone set by Enron's hierarchy. Members of the Risk Management and Assessment Group who reviewed the terms and conditions of deals (and who were largely inexperienced recent MBA graduates) as well as internal auditors, were fearful of retaliation in the PRC from persons whose deals they were reviewing (Chaffin & Fidler, 2002; Dallas, 2003). At best, control was compliance-based, seldom encouraging employees to follow either the letter or the intent of laws (Dallas, 2003).

This punitive environment brought the consequences of dissent sharply into focus. Enron's culture has been characterized as "ruthless and reckless ... lavish[ing] rewards on those who played the game, while persecuting those who raised objections" (Chaffin & Fidler, 2002, pp. 4–5). Led by Skilling's cavalier attitude to rules, top management conveyed the impression that all that mattered was for employees to book profits. In sum, this led to an erosion of employees' confidence in their own perceptions and, most crucially, to further compliance with the organization's leaders in a way that strengthened conformist behaviour. Former employees have noted how "loyalty required a sort of group think" (Chaffin & Fidler, 2002, p. 2) and "that you had to 'keep drinking the Enron water' " (Stephens & Behr, 2002, p. 2). A myth of smooth, flawless operations was perpetuated with problems "papered over" (McLean, 2001, p. 58). The net effect of the rank-and-yank system was to decrease the likelihood that employees would raise objections to any illegal or unethical behaviour of powerful players. The competitiveness the PRC created was exacerbated by Enron's bonus regime. As one insider put it, "sure, the culture at Enron was treacherous, but that was the point" (Swartz & Watkins, 2003, p. 56). Ultimately, the overestimation of profits and underestimation of costs was endemic to the organization.

DISCUSSION

When Kinder left in 1996, a very different corporate climate came to pervade Enron, one that effectively neutralized and subverted his extensive and well-designed diagnostic control and governance structures that were in place. In short, the transformation of Enron's corporate culture and shift in control style were at the heart of Enron's demise.

As Enron rapidly grew into market areas where it did not enjoy a comparative advantage (such as fibre optics and broadband markets), its mercenary corporate culture combined with subverted controls meant that Enron lost its ability to keep track of relevant risks, often taking large positions and encountering unforeseen risks. Skilling was able to bring together a constellation of structural factors that enabled the Enron expansion and re-branding: deregulation, the high-tech investment bubble, enhancements in technological capabilities and a hungry and increasingly expectant investment community. That is, although he was abetted by favourable developments in Enron's institutional environment, it was the agency of Skilling that was able to bring these elements together in a coherent package and craft a culture celebrating creative deal making, innovation, and entrepreneurial and mercenary practices. The emerging cultural climate at Enron in turn had an effect on Skilling. A former executive officer observed:

> Over the years, Jeff changed. He became more of a creature of his own creation. His hubris came to outweigh some of the more attractive parts of his personality. He became more intolerant, more opinionated, more bombastic. Jeff was always right, and that got worse. He had a little bit of a God syndrome. (Stephen & Behr, 2002, p. 4)

Table 2 summarizes the key differences in Enron's control systems, strategy and operating environment under Kinder and Skilling.

To date, the role of culture in the operation of performance management systems has only been touched upon in research (Chenhall, 2003). This has generally been at the level of national culture (see Harrison & McKinnon, 1999 for a review) with very little work assessing the impact of organizational culture on the operation of performance management systems. Individuals are affected by a range of cultural differences beyond those of the nation in which they were brought up. Simons' notion of belief systems represents an important exception, emphasizing the role of an organization's published vision, mission statements, credos and statements of purpose (Simons 1990, 1995). Our approach, however, extends beyond this notion of formal belief systems. Rather than focus on physical artefacts and espoused values, we stress the role of basic assumptions and values in management control.

Numerous methods were used by Skilling to reshape organizational culture in a way that celebrated attempts to exploit and "bend the rules" in order to maximize reported financial returns. Table 3 identifies some of the primary and secondary ways through which Skilling and a dominant coalition within Enron was able to embed elements of this culture.

Table 2. Differences between the Kinder Era and Skilling Era at Enron.

	Kinder Era (1990–1996)	Skilling (1996–2000)
Strategy	Predominantly a physical pipeline and gas exploration business steadily expanding into the shiny new world of energy trading	Expand Enron's energy trading expertise into a range of new commodities to sustain earnings growth. Skilling envisioned taking on markets ranging from paper goods to metals to broadband capacity
Management control systems	Personal, interactive	Impersonal, formal, diagnostic
Operational priority	Cash flow and meeting earnings targets	Margin and volume
Operating environment	Regulated	Deregulated
Technology	Focused on Enron's asset-rich liquid-gas pipeline operations	Focused on online trading of a range of commodities
Work force	Range of employees primarily focused in traditional gas energy trades and vocations	Recruitment of hundreds of MBA students with limited practical experience from top programs throughout the United States
Accounting	Consistent with GAAP conventions for companies in the gas sector	Mark-to-market valuation The complexity of Enron's business model made their balance sheet opaque to the markets Off-balance sheet entities, self-created partnerships run by several members of Enron's own top management, most particularly by CFO Andy Fastow Irregular treatment of derivatives

Table 3 features many of the mechanisms for cultural manipulation discussed by Schein: elimination of dissent (and therefore the promotion of a homogenous and insular group mentality), the accumulation of power at the centre, exaggerated claims about corporate missions and the restriction of negative information (and embellishment of positive information) were all present at Enron under Skilling. An extreme performance-oriented culture that both institutionalized and tolerated deviant behaviour came into being. The enculturation process effectively acted to reduce the

Table 3. Culture-Embedding Mechanisms at Enron under Skilling.

Schein's Primary Culture-Embedding Mechanisms	Primary Embedding Mechanisms at Enron	Secondary Articulation and Reinforcement Mechanisms at Enron
What leaders pay attention to, measure, and control on a regular basis	Mercenary, profit-centred style of management reflected in strategic agendas, key performance metrics and corporate memos	Tight organizational hierarchy and structure
	Almost exclusive focus on key metrics relating to stock price and earnings	
How leaders react to critical incidents and organizational crises	Developing consensus and reaching goals by means of social removal of members who deviate from the culture	Organizational systems and procedures that restricted feedback
	Mercenary public remarks, casual comments and ridiculing	
	Restricting negative information and maximizing positive information	
Criteria by which leaders allocated scarce resources	Performance appraisals and incentive system based almost exclusively on transaction volume and size	Narcissism, highly visible consumption, braggadocio and excess
	Profligate spending and disregard for mounting debt	Exaggerated claims for the Enron vision
Deliberate role modelling, teaching and coaching	Charismatic role modelling, teaching, and coaching predicated on a pragmatic, "doing" orientation to risk	Design of physical space, facades, air planes and buildings
	Dramaturgical events, such as annual conferences and academic forums, where Skilling self-promoted in an exaggerated, sometimes theatrical manner (e.g. dressing up as Darth Vader at a corporate retreat and fostered this moniker)	Skilling led an opulent lifestyle characterized by conspicuous consumption
Criteria by which leaders allocate rewards and status	PRC system which fostered strong competition between organizational members	Myths, stories, legends of outstanding (and well rewarded) performance
	Nepotism and favouritism for successful dealers	
	Distributive negotiation focus of the bonus regime	

Table 3. (*Continued*)

Schein's Primary Culture-Embedding Mechanisms	Primary Embedding Mechanisms at Enron	Secondary Articulation and Reinforcement Mechanisms at Enron
Criteria by which leaders recruit, select, promote, retire and excommunicate organizational members	"Rank and yank" performance evaluations that quickly excommunicated members deemed not to fit in Intense recruitment rituals, designed to engage employees in a process of affiliation Top down communication and very limited upward communication Cronyism and group think reflected in physical artefacts and cultivation of obscure jargon	Formal statements of organizational philosophy, values and creed

range of decisions available to Enron employees to alternatives assessed to be compatible with Enron's mission. The lauding of "creative risk-taking" and "revolution" led to legal and ethical boundaries being stretched, circumvented and even broken. Resistance to bad news created an important pressure point on information sharing internally and externally. Fierce internal competition coupled with huge incentives led to private information, deceit and extensive efforts to bolster short-term performance. The result was social contagion[12] and the normalization of deviancy: the social pathologies promoted by Skilling and Lay became widespread and ultimately toxic in Enron.

Underplaying the role of culture in management control systems results in under-specified models of corporate management. The Enron collapse suggests that control frameworks that focus primarily on formal systems such as Simons' levers of control model, fail to take account of the potentially critical role of shared meanings, accepted norms and rules that may in practice act to influence or even undermine formal systems in practice. The widespread attachment to shared, albeit unwritten, values certainly makes an organization more cohesive. The case expands on Simon's conceptualisation of belief systems, with its emphasis on espoused values and visual artefacts rather than basic assumptions.

While culture should not be seen as placing totalizing, unmediated constraints upon human subjects (Alvesson & Willmott, 2002), in the absence of counter-discourses that interpret enculturation processes as intrusive or offensive, we can anticipate not only instrumental compliance but also increased identification with the cultural values. Accordingly, dominant management control system research should incorporate organizational culture as a powerful lever for guiding organizational behaviour, by informally approving (or prohibiting) patterns of behaviour. The Enron demise indicates that organizational culture provides shared patterns of cognitive interpretations or perceptions, so organization members know how they are expected to act and think. From a managerial viewpoint, cultural control presents a less obtrusive, and potentially more effective, means of organizational control than methods that rely upon external stimuli.

The prescriptive literature on cultural change must also be careful to avoid adopting an overly sanitized, normative perspective that classes "strong" organizational cultures as "good", by pacifying uncertainty and creating stability (e.g. Aaltio-Marjosola, 1994). An organization's shared history and stability can contribute to the internalization and institutionalization of specific attitudes in individuals. Once employees over-align themselves with a company – and invest heavy commitment in organizational routines and the wisdom of leaders – they are liable to lose their original sense of identity, tolerate and rationalize ethical lapses that they would have previously deplored, find a new and possibly corrosive value system taking root, and leave themselves vulnerable to manipulation by organizational leaders to whom they have mistakenly entrusted many of their vital interests (Tourish & Vatcha, 2005). The Enron demise points to numerous risks associated with strong cultures: the risk that a culture motivating and rewarding creative entrepreneurial deal making may provide strong incentives to take additional risks, thereby pushing legal and ethical boundaries; resistance to bad news creates an important pressure point of culture; and internal competition for bonuses and promotion can lead to private information and gambles to bolster short-term performance. At Enron, these risks ultimately disabled the company's elaborate web of controls.

CONCLUSION

The motivation for this article stemmed from the puzzle that while at the time of its demise Enron had in place all the trappings of an extensive, state-of-the-art set of management control and governance systems. Yet, these systems failed to keep the company on a sound trajectory. The article, following Simon's urging to include culture in such research, adopted Schein's (2004) cultural framework for investigating this issue. Our investigation, relying on the vast archival database available regarding the Enron saga, indicated that these controls were systematically ignored, thwarted, and even corrupted by many of the company's managers and executives during the its final years. This led us to conclude that the core values, norms and dispositions of the new culture that emerged during the Lay-Skilling era ran counter to the principles and tenets of sound management control. The article thus contributes to the literature by, first, introducing and using the Schein framework for management control and governance research and, second, bringing into the light the neglected story and attendant lessons that the Enron debacle has for understanding of management control and governance systems.

NOTES

1. Numerous management accounting articles published in refereed journals have relied on similar databases. Toms (2002), for example, relied extensively on sources such as newspaper articles in the *Oldham Standard* and the *Oldham Chronicle* in the nineteenth century, Directors' reports, the trade magazine the *Textile Manufacturer*. Similarly, Walker (2004) relied heavily on nineteenth-century newspaper articles in the *Financial Times*, the *Spectator*, the *Law Times*, the *Daily Courier*, the *Liverpool Mercury* and the *Economist*.

2. The following quotation is typical:

> Controls form the cauldron in which Enron's innovative energies circulate. The heat comes from Enron's ambition to transform global energy markets and from the chance individual deal-makers have for personal wealth accumulation. (Hamel, 2000, p. 214)

3. Lay's political influence ran so high that there was much speculation in Washington that he would be appointed to the cabinet and it was rumored that he harbored the ambition to be Secretary of the Treasury. Lay's lobbying efforts eventually bore fruit. In 1994, the federal government opened the door to deregulation (although it shifted much of the decision-making to the individual state legislatures). After a six-year debate in the Texas Legislature, deregulation was approved in 1999.

4. For example, in a public address in California during the State's 2000 rolling electricity blackouts, when Enron trades was making vast profits on energy trading in the area, he callously joked, "at least when the Titanic went down the lights were on". In another example, during a telephone conference call with investment analysts across the country, Skilling called a major Wall Street banker an "asshole" after being asked why Enron was not reporting balance sheet items of price risk management assets and liabilities as is customary for investment firms.

5. For example, the World Bank cancelled a $100 million Enron water project in Ghana because of "corruption concerns". And, in the case of Enron's UK Azurix water project, strong allegations were made by a civilian watchdog committee that the company had paid a $5 million bribe to senior officials. Allegations of bribery surrounding Enron's widely publicized Indian Dabhol Power project were also widespread (see Bryce, 2002).

6. Most of those receiving "1" chose to accept a severance package rather than stick it out. Furthermore, those in categories "2" and "3" were effectively put on notice that they were liable to be yanked within the next year (Swartz & Watkins, 2003).

7. For example, in June 1999, Skilling transferred risk-management specialist Vince Kaminski out of Risk Management because "he was acting like a cop, trying to kill deals" after Kaminski had expressed concerns about a set of proposed hedges (Behr & Witt, 2002).

8. Round-tripping is technically not illegal in unregulated markets.

9. In spite of this deteriorating cash situation, stock options remained a mainstay of Enron's remuneration packages. For example, from January 1999 through July 2001, 18 senior executives and 9 board members cashed stock options in the amount of $1.2 billion. In spite of Enron's chronic and serious debt situation and its inadequate cash flow stream, during Skilling's reign as CEO and president, Enron also paid out multi millions on office buildings in Houston and London.

10. During Kinder's era, Enron's fleet was limited to five small economical jets with a cost of $1,500 per flight hour. As soon as he left these were sold and replaced with expensive jets with a per flight hour cost of $4,200. Moreover in 2001, Lay got board approval for a $41.6 million new Gulf Stream V, allegedly for non-stop flights to Europe, Asia and South America. The plane, however, was essentially reserved for Lay, his wife and other Enron board members for personal trips.

11. Similarly, Investor Relations officer Paula Rieker testified that for the June 2000 quarterly release, Enron officers had planned to meet analysts' consensus figure of 32 cents, but at the last minute Skilling told him that "he wanted to beat earnings by 2 or 3 cents. Four days later, Enron reported 34 cents. Analysts were never told about the sudden change" (Barrionuevo, 2006).

12. Demski (2003, p. 67) has also introduced the possibility of contagion or "herding" in the context of corporate malfeasance at Enron.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support of the Certified Management Accountants/Canadian Academic Accounting Association research program. We also recognize comments received at presentations at the Canadian Academic Accounting Association Annual Meeting (May 2006) and the European Accounting Association Annual Meeting (March 2006). Norman Macintosh also acknowledges the financial support of the Social Science and Humanities Research Council (Canada).

REFERENCES

Aaltio-Marjosola, I. (1994). From a "grand story" to multiple narratives: Studying an organisational change project. *Journal of Organisational Change Management*, *7*(5), 56–67.

Abdel-khalik, A. R. (2002). Reforming corporate governance post Enron: Shareholders' board of trustees and the auditor. *Journal of Accounting and Public Policy*, *21*(2), 97–103.

Alvesson, M., & Willmott, H. (2002). Identity regulation as organizational control: Producing the appropriate individual. *Journal of Management Studies*, *39*(5), 619–644.

Amigoni, F. (1978). Planning management control systems. *Journal of Business Finance and Accounting*, *5*(3), 279–291.

Barnes, J. E., Barnett, M., & Schmitt, C. H. (2002). How a titan came undone. *US News and World Report* (March 18), 27–36.

Barrionuevo, A. (2006). Enron data modified, jury is told. *New York Times*, February 22, Section C, p. 1.

Bauder, D. (2002). Enron case will label many as fools and knaves. *The San Diego Union-Tribute* (January 18), H2.

Behr, P., & Witt, A. (2002). Concerns grow amid conflicts. *Washington Post* (30 July), 54.

Benston, G. J., & Hartgraves, A. L. (2002). Enron: what happened and what we can learn from it. *Journal of Accounting and Public Policy*, *21*(2), 105–127.

Bernard, H. (2002). *Research methods in anthropology* (3rd ed.). Walnut Creek, CA: AltaMira Press.

Bisbe, J., & Otley, D. T. (2004). The effects of interactive use of management control systems on product innovation. *Accounting, Organizations and Society*, *29*(8), 709–737.

Bryce, R. (2002). *Pipe dreams: Greed, ego, and the death of Enron*. New York: Public Affairs.

Byrne, J., France, M., & Zellner, W. (2002). At Enron, 'the environment was ripe for abuse'. *Business Week Online*, February 25. Available at http://www.businessweek.com/magazine/content/02_08/b3771092.htm

Chaffin, J., & Fidler, S. (2002). Enron Revealed to be rotten to the core. *Financial Times* (April 9), 3–6.

Chenhall, R. H. (2003). Management control system design within its organizational context: Findings from contingency-based research and directions for the future. *Accounting, Organizations and Society*, *28*(2–3), 127–168.

Conrad, C. (2003). Stemming the tide: Corporate discourse and agenda denial in the 2002 'corporate meltdown'. *Organization*, *10*, 549–560.

Coy, P., & Anderson, S. (2002). Enron: How good an energy trader? *Business Week Online*, (February 11).

Craig, R., & Amernic, J. (2004). Enron discourse: The rhetoric of a resilient capitalism. *Critical Perspectives on Accounting*, *15*(6/7), 813–851.

Cruver, B. (2003). *Anatomy of greed: The unshredded truth from an Enron insider*. London: Arrow.

Dallas, L. (2003). A preliminary inquiry into the responsibility of corporations and their directors and officers for corporate climate: The psychology of Enron's demise. *Rutgers Law Journal*, *35*, 1–68.

Demski, J. S. (2003). Corporate conflicts of interest. *Journal of Economic Perspectives*, *17*(2), 51–72.

Fowler, T. (2002). The pride and the fall of Enron. *The Houston Chronicle* (October 20), 14.

Fox, L. (2003). *Enron: The rise and fall*. Hoboken NJ: Wiley.

Fusaro, P., & Miller, R. (2003). *What went wrong at Enron: Everyone's guide to the largest bankruptcy in US history*. Hoboken, NJ: Wiley.

Gruley, B., & Smith, R. (2002). Anatomy of a fall, keys to success left Kenneth Lay open to disaster. *The Wall Street Journal*, (April 26), A1 and A5.

Hamel, G. (2000). *Leading the revolution*. Boston, MA: Harvard Business School Press.

Harrison, G., & McKinnon, J. (1999). Cross cultural research in management accounting systems design: A review of the current state. *Accounting, Organizations and Society*, *24*(5/6), 483–506.

Healy, P. M., & Palepu, K. G. (2003). The fall of Enron. *The Journal of Economic Perspectives*, *17*(2), 3–26.

Joint Committee on Taxation, Senate Finance Committee. (2003). *Report of investigation of Enron corporation and related entities regarding federal tax and compensation issues, and policy recommendation*. Volume I, February.

Kaplan, R. S. (1984). The evolution of management accounting. *The Accounting Review*, *59*(3), 390–417.

Kets de Vries, M., & Miller, D. (1987). Interpreting organizational texts. *Journal of Management Studies*, *24*(3), 233–247.

McLean, B. (2001). Why Enron went bust. *Fortune*, *144*(12), 52–58.

McLean, B., & Elkind, P. (2004). *The smartest guys in the room: The amazing rise and scandalous fall of Enron*. New York: Portfolio.

Merchant, K. A. (1998). *Modern management control systems: Text and cases*. Upper Saddle River, NJ: Prentice Hall.

Modell, S. (1995). A three-dimensional approach to management control systems. *Journal of Business and Management*, *2*(2), 35–73.

Ouchi, W. G. (1979). A conceptual framework for the design of organizational control mechanisms. *Management Science*, *25*(9), 833–848.

Permanent Subcommittee on Investigations of the Committee on Governmental Affairs, United States Senate. (2002). *The role of the Board of Directors in Enron's collapse*. Report 107-70, July 8.

Roberts, J., & Thomas, E. (2002). Enron's dirty laundry. *Newsweek* (March 11), 22–30.

Schein, E. (1996). Culture: The missing concept in organization studies. *Administrative Science Quarterly*, *41*(2), 229–240.

Schein, E. (2004). *Organizational culture and leadership* (3rd ed.). San Francisco, CA: Jossey-Bass.

Simons, R. (1990). The role of management control systems in creating competitive advantage: New perspectives. *Accounting, Organizations and Society*, *15*(1/2), 127–143.

Simons, R. (1995). *Levers of control: How managers use innovative control systems to drive strategic renewal*. Boston: Harvard Business School Press.

Simons, R. (2000). *Performance measurement and control systems for implementing strategy*. Upper Saddle River, NJ: Prentice Hall.

Sims, R., & Brinkmann, J. (2003). Enron ethics (Or: Culture matters more than codes). *Journal of Business Ethics*, *45*(3), 243–256.

Stephens, J., & Behr, P. (2002). Enron culture fed its demise. *Washington Post*, (June 27), Sections A1–A2.

Streitfeld, D., & Romney, L. (2002). One cozy bunch. *Los Angeles Times*, *2*(January 27), Part A1, 1–8.

Swartz, M., & Watkins, S. (2003). *Power failure: The inside story of the collapse of Enron*. New York: Doubleday.

Toms, S. (2002). The rise of modern accounting and the fall of the public company: The Lancashire cotton mills 1870–1914. *Accounting, Organizations and Society*, *27*(1–2), 61–84.

Tourish, D., & Vatcha, N. (2005). Charismatic leadership and corporate cultism at Enron: The elimination of dissent, the promotion of conformity and organizational collapse. *Leadership*, *1*(4), 455–480.

Walker, S. (2004). Conflict, collaboration, fuzzy jurisdictions and partial settlements: Accountants, lawyers and insolvency practice during the late 19th century. *Accounting and Business Research*, *34*(3), 247–265.

Watkins, S. (2003a). Ethical conflicts at Enron: Moral responsibility in corporate capitalism. *California Management Review*, *45*, 6–19.

Watkins, S. (2003b). Former Enron vice-president Sherron Watkins on the Enron collapse. *Academy of Management Executive*, *17*(4), 119–127.

Watkins, S. (2003c). Pristine ethics. *Vital Speeches of the Day*, *69*(14), 435–439.

Whitley, R. (1999). Firms, institutions and management control: The comparative analysis of coordination and control system. *Accounting, Organizations and Society*, *24*(5/6), 507–524.

Zellner, W., & Anderson, S. (2001). The fall of Enron. *Business Week Online*, (December 17).