

**AN EXAMINATION OF ORGANIZATIONAL PERFORMANCE  
MEASUREMENT SYSTEM UTILIZATION**

Gerald Kenneth DeBusk

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and  
State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy  
In  
General Business  
With a Major in Accounting

Robert M. Brown, Chair

John A. Brozovsky

Larry N. Killough

John J. Maher

Robert S. Stephens

April 2, 2004

Blacksburg, Virginia

Keywords: Performance Measurement, Balanced Scorecard,  
Tolerance for Ambiguity

Copyright 2004, Gerald K. DeBusk

# **AN EXAMINATION OF ORGANIZATIONAL PERFORMANCE MEASUREMENT SYSTEM UTILIZATION**

Gerald Kenneth DeBusk

## **(ABSTRACT)**

This dissertation provides results of three studies, which examine the utilization of organizational performance measurement systems. Evidence gathered in the first study provides insight into the number of perspectives or components found in the evaluation of an organization's performance and the relative weight placed on those components. The evidence suggests that the number of performance measurement components and their relative composition is situational. Components depend heavily on the strategies selected by the organization. Bottom-line financial measures like return on invested capital and net profit, while perceived as more important than their nonfinancial counterparts, were not part of the extracted components suggesting that they were viewed as outcomes to be achieved by controlling key nonfinancial measures.

The second study examines potential cognitive difficulties inherent in the use of performance measurement systems. Results suggest that whether performance was better than target, worse than target, or equal to target does not affect the perceived importance of the measures. Results also suggest an emphasis on historical financial measures and a lack of emphasis on more forward-looking nonfinancial measures. In addition, there is evidence of a halo effect in that an organization's performance on financial measures appears to influence an individual's perception of the organization's performance on nonfinancial measures.

The third study uses structural equation modeling and other related procedures to examine the relationships surrounding an executive's use of

performance measurement information. Results suggest that a personality characteristic of executives, specifically their intolerance of ambiguity, affects the amount of information perceived to be important in a performance measurement system. The results further suggest that the amount of information perceived to be important affects the evaluation of organizational performance with perceived risk serving as a mediating variable.

Overall, these three studies add to our knowledge of organizational performance measurement system utilization by examining the relative weightings of performance measures, the judgmental effects from utilization of performance measurement systems, and the impact of intolerance of ambiguity on the importance of performance measurement data. In addition, this dissertation examines the link between performance measurement data and the perception of risk in the evaluation of organizational performance.

## **DEDICATION**

This dissertation is dedicated to my parents Margaret V. and Charles M. DeBusk and to my brother Charles R. DeBusk. Thank you for your love and support. A special dedication is reserved for my Mother, who has for my entire life been a source of inspiration and wisdom.

## **ACKNOWLEDGEMENTS**

I would first like to thank God for all the blessings bestowed upon me. With His help, truly all things are possible.

I would like to acknowledge several faculty members. First, I would like to thank the members of my dissertation committee, Robert M. Brown (Chair), John A. Brozovsky, Larry N. Killough, John J. Maher, and Robert S. Stephens. I am especially indebted to Professor Brown and Professor Killough for their support and guidance throughout my studies. I would also like to thank Professor Wayne Leininger for his encouragement throughout the first three years of my doctoral program.

I appreciate the support given by several of my fellow Ph.D. students, Chris Allport, Duane Brandon, Aaron Crabtree, Scott Fleming, and Bill Kerler. They have also become great friends. I would like to thank several friends from my church for their encouragement over the years, Laura Berry, I-Ping Fu, Reed Kennedy, Pat Rodgers, and Emmett Wright. In addition, I appreciate the on-going support provided by some former colleagues, Judy Griffiths, Carol Cissel, and Ray Wurzberger.

Lastly, I would like to thank Professor Floyd A. Beams, Retired in whose name a scholarship was established that provided much needed financial support during my doctoral studies.

## TABLE OF CONTENTS

<b>Dedication .....</b>	<b>iv</b>
<b>Acknowledgements .....</b>	<b>v</b>
<b>List of Figures.....</b>	<b>x</b>
<b>List of Tables .....</b>	<b>xi</b>
<b>1 Introduction.....</b>	<b>1</b>
1.1 References.....	5
<b>2 Components and Relative Weights in Utilization of Dashboard Measurement Systems like the Balanced Scorecard .....</b>	<b>7</b>
2.1 Introduction.....	7
2.2 Dashboard Measurement Systems .....	8
2.2.1 <i>Tableau de Bord</i> .....	8
2.2.2 <i>Financial vs. Nonfinancial Measures</i> .....	8
2.2.3 <i>The Balanced Scorecard</i> .....	10
2.3 Research Questions.....	11
2.3.1 <i>Number of Components (Perspectives)</i> .....	11
2.3.2 <i>Relative Weights</i> .....	12
2.4 Method .....	13
2.4.1 <i>Case Development</i> .....	13
2.4.2 <i>Sample</i> .....	14
2.4.3 <i>Data Analysis</i> .....	15
2.5 Results.....	15
2.5.1 <i>Principal Components Analysis</i> .....	15
2.5.2 <i>T-Test</i> .....	17
2.6 Discussion.....	18
2.7 Research and Limitations.....	19
2.8 References.....	21
2.9 Figures and Tables .....	23
<b>3 Judgmental Effects in the Use of Performance Measurement Systems.....</b>	<b>26</b>

3.1	Introduction.....	26
3.2	Literature Review and Research Question Development.....	29
3.3	Method.....	34
3.3.1	<i>The Experiment</i> .....	34
3.4	Results.....	36
3.4.1	<i>Data Analysis – Research Question</i> .....	36
3.4.2	<i>Results – Research Question</i> .....	36
3.4.3	<i>Robustness Tests – Research Question</i> .....	37
3.4.4	<i>Data Analysis – Research Hypothesis</i> .....	38
3.4.5	<i>Results – Research Hypothesis</i> .....	40
3.4.6	<i>Robustness Tests – Research Hypothesis</i> .....	40
3.4.7	<i>Discussion of Results – Research Hypothesis and Halo Effects</i> .....	41
3.5	Discussion and Conclusion.....	42
3.6	References.....	44
3.7	Figures and Tables.....	47
<b>4</b>	<b>Intolerance of Ambiguity and Organizational Performance Measurement .....</b>	<b>54</b>
4.1	Introduction.....	54
4.2	Literature Review.....	56
4.2.1	<i>Measures of Intolerance of Ambiguity</i> .....	56
4.2.2	<i>Studies with Intolerance of Ambiguity as an Independent Variable</i> .....	58
4.2.3	<i>Information Usage and Perceived Risk</i> .....	62
4.2.4	<i>Perceived Risk versus Perceived Benefit</i> .....	63
4.2.5	<i>Balanced Scorecard</i> .....	64
4.3	Development of Needed Intolerance of Ambiguity Measures .....	65
4.3.1	<i>Pilot Study 1</i> .....	65
4.3.2	<i>Pilot Study 2</i> .....	66
4.4	Research Question and Hypothesis Development.....	68
4.4.1	<i>Importance of Information</i> .....	68
4.4.2	<i>Information Use, Perceived Risk and Assessment of Performance</i> .....	69

4.4.3	<i>Presentation Format</i> .....	70
4.5	Methodology .....	71
4.5.1	<i>Instrument and Measures</i> .....	71
4.5.2	<i>Sample</i> .....	72
4.5.3	<i>Structural Equation Modeling</i> .....	73
4.6	Results .....	74
4.6.1	<i>Sample 1</i> .....	74
4.6.2	<i>Sample 2</i> .....	75
4.6.3	<i>Sample 2 - Supplemental Analysis</i> .....	76
4.7	Discussion .....	79
4.7.1	<i>Research Question 1</i> .....	79
4.7.2	<i>Hypothesis 1</i> .....	80
4.7.3	<i>Hypothesis 2</i> .....	80
4.7.4	<i>Research Question 2</i> .....	81
4.8	Implications .....	81
4.9	References .....	83
4.10	Figures and Tables .....	87
<b>5</b>	<b>Conclusion</b> .....	<b>108</b>
5.1	Summary and Conclusions .....	108
5.2	Implications for Future Research .....	110
5.3	References .....	112
	<b>Appendices</b> .....	<b>113</b>
Appendix A	Case Materials for Chapter 2 and Chapter 3 .....	113
(a)	<i>Company Background</i> .....	113
(b)	<i>Descriptions of Performance Measures</i> .....	116
Appendix B	Performance Measures for Chapter 2 Survey .....	117
Appendix C	Performance Measures for Chapter 3 Survey .....	118
Appendix D	Survey Materials for Chapter 4, Pilot Study 1 .....	119
Appendix E	Survey Materials for Chapter 4, Pilot Study 2 .....	120



Appendix F	Survey Materials for Chapter 4, Main Study (BSC Format) .....	122
<b>Vita</b>	.....	<b>127</b>

## **LIST OF FIGURES**

Figure 2-1 Suggested No. of Measures per Balanced Scorecard Perspective .....	23
Figure 3-1 Mean Perceived Performance on Nonfinancial Measures .....	47
Figure 4-1 IOA and Organizational Performance – Proposed Model .....	87
Figure 4-2 IOA and Organizational Performance – Final Model .....	88
Figure 4-3 IOA and Organizational Performance – Final Model SEM Results...	89
Figure 4-4 IOA and Organizational Performance – Supplemental Model .....	90
Figure 4-5 IOA and Org. Performance – Supplemental Model SEM Results.....	91

## LIST OF TABLES

Table 2-1 Descriptive Statistics and Principal Components Analysis of the Perceived Importance of the Performance Measures.....	24
Table 2-2 Paired Samples T-Test of the Difference in the Perceived Mean Importance of Financial versus Nonfinancial Measures.....	25
Table 3-1 Organizational Performance Measures for Each Version of the Case .	48
Table 3-2 Descriptive Statistics – Perceived Importance .....	49
Table 3-3 Multivariate Analysis of Variance on Perceived Importance.....	50
Table 3-4 Descriptive Statistics – Perceived Performance .....	51
Table 3-5 Pearson Correlation Coefficients – Performance versus Importance...	52
Table 3-6 Version 2 versus 3 – Overall Performance Assessment T-Test .....	53
Table 4-1 Correlations between IOA Scales.....	92
Table 4-2 Prior Studies Using IOA as an Independent Variable .....	93
Table 4-3 Budner Scale Factor Analysis (Varimax Rotation), Pilot Study 1 .....	94
Table 4-4 Cross Reference of Items Used in Variable Development.....	95
Table 4-5 IOA Factor Analysis (Varimax Rotation), Pilot Study 2 .....	96
Table 4-6 IOA Factor Analysis, Main Study – Sample 1 .....	97
Table 4-7 Risk Factor Analysis, Main Study – Sample 1 .....	98
Table 4-8 Descriptive Statistics, Main Study – Sample 2 .....	99
Table 4-9 Pearson Correlations, Main Study – Sample 2 .....	100
Table 4-10 SEM Standardized Regression Weights, Main Study – Sample 2 ...	101
Table 4-11 Pct. Important Frequencies, Main Study – Sample 2 .....	102
Table 4-12 Supplemental SEM Regression Weights, Main Study – Sample 2 ..	103
Table 4-13 Additional Descriptive Statistics, Main Study – Sample 2 .....	104
Table 4-14 Pearson / Spearman Correlations, Main Study – Sample 2.....	105
Table 4-15 Correlations after Median Split, Main Study – Sample 2 .....	106
Table 4-16 Perceived Risk T-Test .....	107

## 1 INTRODUCTION

*“The choice of performance measurement systems is one of the most critical challenges facing organizations.”*

*(Ittner and Larcker 1998, 205)*

Organizational performance measurement systems are used to provide information to both external third parties (e.g. shareholders, lenders, and potential investors) and to those internal to the organization (e.g. directors, officers, managers, and employees). They are a critical aspect of strategic planning, business-unit performance evaluation, and executive compensation (Ittner and Larcker 1998).

Organizational performance measurement systems have traditionally focused on financial measures of performance; however, criticism has been levied in recent years at this approach. Financial measures are criticized as being “outcome” measures instead of measures that can be used to “drive” improved performance. Financial measures are also characterized as being too historical or backward-looking (Ittner and Larcker 1998). Birchard (1995, 43) comments,

Yesterday's accounting results say nothing about the factors that actually help grow market share and profits—things like customer service innovation, R&D effectiveness, the percent of first-time quality, and employee development.

In response to these criticisms, the American Institute of Certified Public Accountants (1994) and the American Accounting Association Financial Accounting Standards Committee (2002) have called for organizations to report more forward-looking information, including nonfinancial measures, for key business processes.

The Balanced Scorecard (BSC) has become popular in recent years as a system that combines the use of nonfinancial measures with financial measures. It emphasizes the linking of strategy to organizational performance measurement (Kaplan and Norton 2001a, 2001b, 2001c). A survey conducted by Renaissance Worldwide, Inc. estimated 60% of Fortune 1000 companies have implemented or are experimenting with the BSC (Silk 1998). Potential cognitive difficulties in the use of performance measurement systems like the BSC have been explored in recent literature (e.g. Lipe and Salterio 2000, 2002; Ittner et al. 2003; Banker et al. 2004). This dissertation builds on the existing literature by examining the utilization of performance measurement systems in three separate research studies.

The first study provides insight into the influence of strategy on the number of components or perspectives utilized in evaluating an organization's performance and the relative weight placed on each of those components. The evidence suggests that the number of performance measurement perspectives and their composition is situational. They depend heavily on strategies chosen by the organization. Bottom-line financial measures like return on invested capital and net profit, while perceived as more important than their nonfinancial counterparts, were not part of the perspectives extracted through the use of principal

components analysis (PCA). The PCA results suggest that financial measures may be viewed as outcomes to be achieved by controlling key nonfinancial measures.

The second study examines two potential cognitive difficulties in the utilization of systems containing financial and nonfinancial performance measures. First, this study examines the effect of relative performance to target on perceived importance of different performance measures. Whether performance was better than target, worse than target, or equal to target did not affect the perceived importance of the measure. These results suggest users of performance measurement systems will not alter the relative weight placed on a particular measure regardless of its position vis-à-vis the targeted or expected results. Second, this study examines the potential for overweighting financial measures as compared to nonfinancial measures in the evaluation of an organization's overall performance. The results suggest an emphasis on historical financial measures and a lack of emphasis on more forward-looking nonfinancial measures. In addition, the study provides evidence of a halo effect in that an organization's performance on financial measures appears to influence an individual's perception of the organization's performance on nonfinancial measures.

The third study examines the effect of personality on the usage of information by top executives in evaluating organizational performance. Results suggest that a cognitive characteristic of executives, specifically their intolerance of ambiguity, affects the amount of information perceived to be important in a performance measurement system. This was compelling in light of additional evidence suggesting the amount of information perceived to be important and the evaluation of organizational performance were positively correlated, with perceived risk as a mediating variable.

The remainder of this dissertation is organized as follows. Chapters 2 through 4 each contain separate research studies. Each includes an introduction, a review of the relevant prior literature, a description of the research methodology,

a summary of the results, a discussion of the findings, and a section containing references. Figures and tables are grouped together at the end of each chapter. A conclusion to the dissertation is provided in chapter 5. The appendices contain survey and case materials relating to chapters 2 through 4.

## 1.1 References

- American Accounting Association. 2002. Recommendations on disclosure of nonfinancial performance measures. *Accounting Horizons* 16 (4): 353-362.
- American Institute of Certified Public Accountants (AICPA). 1994. *Improving business reporting – a customer focus*. New York, NY: AICPA.
- Banker, R. D., H. Chang, and M. J. Pizzini. 2004. The balanced scorecard: judgmental effects of performance measures linked to strategy. *The Accounting Review* 79 (1): 1-23.
- Birchard, B. Making it count. *CFO* (October): 42-51.
- Ittner, C. D., and D. F. Larcker. 1998. Innovations in performance measurement: trends and research implications. *Journal of Management Accounting Research* 10: 205-238.
- Ittner, C. D., D. F. Larcker, and M. W. Meyer. 2003. Subjectivity and the weighting of performance measures: evidence from a balanced scorecard. *The Accounting Review* 78 (3): 725-758.
- Kaplan, R. S. and D. P. Norton. 2001a. Transforming the balanced scorecard from performance measurement to strategic management: part I, *Accounting Horizons* 15 (1): 87-104.
- and -----, 2001b. Transforming the balanced scorecard from performance measurement to strategic management: part II. *Accounting Horizons* 15 (2): 147-160.
- and -----, 2001c. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, MA: Harvard Business School Publishing Corporation.
- Lipe, M. S., and S. E. Salterio. 2000. The balanced scorecard: the judgmental effects of common and unique performance measures. *The Accounting Review* 75 (3): 283-298.
- and -----, 2002. A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society* 27: 531-540.



Silk, S. 1998. Automating the balanced scorecard. *Management Accounting* (May): 38-44.

## **2 COMPONENTS AND RELATIVE WEIGHTS IN UTILIZATION OF DASHBOARD MEASUREMENT SYSTEMS LIKE THE BALANCED SCORECARD**

### **2.1 Introduction**

“Dashboard measures” are intended to give management a quick view of organizational performance, i.e. organizational performance “at-a-glance”. The term “dashboard measures” is derived from examples like the set of instruments and gauges found on the dashboard of your car or on the instrument panel of an airplane. The gauges and indicators reflect the on-going performance of the various functions in the airplane or automobile. The dashboard of your automobile has indicators of current speed, distance traveled, engine temperature, and fuel reserves. A review of these indicators allows the driver to quickly assess the performance of the automobile and make necessary corrections, e.g. slow down or stop for fuel. Dashboard measures for an organization are intended to function in a similar fashion. After a quick review of the organizational dashboard measures, management should be able to assess the performance of the organization and make decisions related to that performance. This study examines aspects of how dashboard measures are utilized in evaluating organizational

performance. This study provides empirical evidence for use in answering two sets of research questions. The questions relate to (1) the number of components or perspectives in a dashboard measurement system and (2) the relative weight placed on those components.

The rest of this paper is organized as follows. The next section discusses dashboard measurement systems. The third section introduces the research questions and the fourth section lays out the method used in conducting the experiment and analyzing the data. The fifth section contains the results of the study, the sixth section is a discussion of the results and the last section deals with potential limitations.

## **2.2 Dashboard Measurement Systems**

### *2.2.1 Tableau de Bord*

A long-standing organizational dashboard measurement system is the French Tableau de Bord (TdB). The TdB dates back to the beginning of the 20<sup>th</sup> century with detailed writings on it going back at least 40 years. The TdB is a set of dashboard measures including both financial and nonfinancial measures. It is intended to be a method for translating the organization's mission and vision into objectives from which critical success factors can be derived. Critical performance indicators are then developed to track the organization's performance on the critical success factors. Unfortunately, the deployment of the TdB in practice is quite different than what is called for in the writings. In practice, there is often much more emphasis on the traditional financial measures than the nonfinancial measures (Epstein and Manzoni 1997).

### *2.2.2 Financial vs. Nonfinancial Measures*

Organizations have struggled for many years with the inadequacies of accounting performance measures. Using both financial and nonfinancial measures together is an attempt to solve the problems that arise from emphasizing

traditional financial measures. There are three main reasons for the use of nonfinancial performance measures: (1) perceived limitations in the use of traditional financial measures, (2) increased competitive pressure, and (3) implementation of other programs like Total Quality Management which call for the use of nonfinancial measures (Ittner and Larcker 1998). Ittner and Larcker (1998) cite various limitations of traditional financial measures. Specifically, financial measures...

- are too historical and backward looking.
- lack predictive ability.
- reward short-term or incorrect behavior.
- are not actionable.
- do not capture key business changes until too late.
- are too aggregated and summarized to guide management action.
- are too departmentalized instead of cross-functional.
- do not effectively consider intangibles.

Because of these problems, some firms, in recent years, have made changes to their financial measures. The new measures focus more on cash flow and value creation. Measures like EVA<sup>®</sup> (Economic Value Added) and CFROI (Cash Flow Return on Investment) have become more popular. Other firms have emphasized “forward-looking” nonfinancial measures to counter the problems mentioned above. Examples of these “forward-looking” nonfinancial measures include measures of customer satisfaction, employee satisfaction, product and process innovation, community involvement, and defect rates. Many firms have utilized the Balanced Scorecard (BSC), another type of dashboard measurement system similar to the TdB, to incorporate nonfinancial measures into an overall set of measures. A survey conducted by Renaissance Worldwide, Inc. estimated 60% of Fortune 1000 companies have implemented or are experimenting with the BSC (Silk 1998).

### 2.2.3 *The Balanced Scorecard*

Kaplan and Norton introduced the BSC in 1992. Like the TdB, the BSC is a mechanism for translating the organization's strategy into operational terms (Kaplan and Norton 1992, 1996, 2001a, 2001b, 2001c). Implementing the BSC requires management to (1) develop coherent strategies to achieve their objectives and (2) develop a set of measures for gauging the organization's performance in implementing those strategic objectives. According to Kaplan and Norton (2001c), a well-developed BSC will allow employees to discern the organization's strategy simply by examining the scorecard.

A BSC is a one-page document with 18 – 25 key measures comparing the organization's performance to planned targets. The BSC document typically organizes measures into four categories and is the framework for organizing the firm's strategic objectives into these four perspectives or views (Kaplan and Norton 2001a, 90):

1. *Financial – the strategy for growth, profitability, and risk viewed from the perspective of the shareholder.*
2. *Customer – the strategy for creating value and differentiation from the perspective of the customer.*
3. *Internal Business Processes – the strategic priorities for various business processes that create customer and shareholder satisfaction.*
4. *Learning and Growth – the priorities to create a climate that supports organizational change, innovation, and growth.*

The BSC helps to define a strategy's cause-and-effect relationships in relation to the four perspectives. To execute a strategy, the strategic intent, formulated at the top levels of the organization must cascade downward and have an impact on the lower levels of the organization, particularly in their development of objectives (Porth and Maheshkumar 1998). A BSC implementation tool known as "strategy maps" is used for describing and implementing strategies throughout all levels of the organization. The strategy map specifies elements and their linkages critical to the strategy (Kaplan and Norton 2001a).

## **2.3 Research Questions**

### *2.3.1 Number of Components (Perspectives)*

The first research question deals with the number of perspectives or components being employed by dashboard measurement systems. Kaplan and Norton (1992) determined after a one-year research project with 12 companies that four perspectives were ideal. The four perspectives represent the views of the three major stakeholders in an organization: owners, customers, and employees plus a fourth view representing internal processes. Kaplan and Norton (1996, 25) state, “The four perspectives of the scorecard permit a balance between short-term and long-term objectives, between outcomes desired and the performance drivers of those outcomes, and between hard objectives measures and softer, more subjective measures.” However, Kaplan and Norton (1996, 34) express a cautionary note, “... the four perspectives should be considered a template, not a straitjacket. No mathematical theorem exists [stating] that four perspectives are both necessary and sufficient.”

There has been very little empirical research done concerning the number of perspectives in a performance measurement system. Hoque and James (2000) studied Australian manufacturing companies. As part of their study they asked chief financial officers to assess the extent of usage within their firms of 20 separate performance measures. The 20 performance measures were typical measures picked a priori without regard to the specific needs of the organizations. Extent of usage was measured on a five-point Likert scale. Hoque and James (2000) performed principal components analysis with a varimax rotation on the extent of usage questions. The results suggest that Kaplan and Norton’s four perspectives were appropriate. However, the authors noted a limitation (Hoque and James, 2000, 12), “This instrument has failed to pick up the strategic linkages of a real BSC.” Does strategy play a role in determining the number of perspectives in a performance measurement system? The number of perspectives

could be linked to the specific strategies of an organization and thus not considered in the Hoque and James (2000) study. This brings us to the first set of research questions.

Q1a: How many components (perspectives) are there to dashboard performance measurements—the four components (perspectives) proposed by Kaplan and Norton or a different number?

Q1b: Does strategy play a role in determining the number of perspectives in a performance measurement system?

### 2.3.2 *Relative Weights*

To utilize the BSC, relative weights for each of the scorecard measures must be established. These weights are employed in evaluating performance. Kaplan and Norton (2001c) propose a balanced weighting scheme where financial measures receive only 22% of the relative weight. This suggests that nonfinancial measures are at least as important, if not more important, than financial measures. Figure 2-1 contains the breakdown for number of measures in each category and their relative weights as suggested by Kaplan and Norton (2001b, 375).

In addition, Kaplan and Norton (1992) believe the BSC minimizes information overload by limiting the number of measures used and forcing managers to focus on the key measures used in the scorecard. However, a 1996 Towers Perrin survey found that BSC adopters placed 56% of the relative weight on financial measures and studies on information overload suggest that a large number of performance measures can hinder a manager's ability to evaluate the organization (Ittner and Larcker 1998). An example of placing a preponderance of weight on financial measures is Cigna Insurance's Property and Casualty Division. They tie bonuses to their BSC; but the financial measures account for 50% of the weight in the bonus calculation (McWilliams 1996). Kaplan and Norton (1992) believe the BSC guards against suboptimal decision-making because it forces managers to consider all the important measures together. An

alternate theory suggests that managers rely on what they know best—financial measures. Anthony and Govindarajan (2001, 451) state,

... not only are most senior managers well trained and adept with financial measures, they also keenly feel pressure regarding the financial performance of their companies. Shareholders are vocal, and boards of directors frequently apply pressure on the shareholders' behalf. This pressure may overwhelm the long-term, uncertain payback of the nonfinancial measures.

Considering Kaplan and Norton's proposed emphasis (relative weight) on nonfinancial measures, the following research question needs to be answered.

Q2: Are financial measures weighted more (i.e. more important) than nonfinancial measures?

## **2.4 Method**

### *2.4.1 Case Development*

A case was provided to participants giving background data, information on vision and strategies, and measures of performance against targets for a major brewery. The case was adapted from an Institute of Management Accountants case designed for classroom use (Grove, Cook and Richter 2000). See Appendix A and Appendix B. Specific information was given on 24 individual performance measures including a definition of the measure, the target for the year, and the actual results for the year. The measures were organized onto one page and listed alphabetically. See Appendix B. An explanation of the measures was also provided. See Appendix A – Descriptions of Performance Measures. Participants in the experiment were asked to evaluate the organization's performance on individual measures and on an overall basis on a scale of one to six, where one equaled "very poor" and six equaled "very good". Participants were also asked to assess the importance of each measure in evaluating the organization's performance on a scale of one to six, where one equaled "not important" and six equaled "critical".



#### 2.4.2 *Sample*

The case and survey instruments were given to graduate students at a large state university in the eastern United States. Students were contacted and material was distributed via email. Participants completed a web-based survey instrument. A total of 88 surveys were received. Three surveys were discarded due to missing information; another was discarded for failure to understand the instrument. Thus, 84 usable surveys were obtained. The overall usable response rate was 62% (84 out of 136 possible students). Most of the usable surveys (80 out of 84 or 95%) were from students participating in a course offered in the MBA program. Four additional volunteers were obtained from a graduate accounting course.

The demographics of the sample suggest that the student subjects used display characteristics similar to those found in the business environment. Many of the students were full-time working professionals attending MBA courses on a part-time basis. The median age of the participants was 29 years (Mean 29.7, S.D. 6.8 years) and the median amount of professional work experience was 5.5 years (Mean 7.3, S.D. 7.0 years).<sup>1</sup> Ninety-nine percent (83 out of 84) of the participants were graduate students, 38% were female and 62% were male. The median age of the participants was 29 years (Mean 29.7, S.D. 6.8 years).

---

<sup>1</sup> It is sometimes suggested that students are often not the best subject pool. In this situation, however, the authors believe that experienced MBA students give us an appropriate subject pool for addressing our particular research questions. The subject pool contained experienced professionals (mean work experience was 7.3 years) with 39% of the participants' current or last position being in some form of managerial or supervisory capacity. As an alternative to the use of MBA students, we could have sampled key managers in a cross-section of firms. However, firms within different industries tend to use different measures and firms tailor their measures to fit their strategic objectives. As Hoque & James (2000) have stated, sampling a cross-section of firms has the limitation of failing to detect linkages between strategy and measurement. Another alternative would have been to draw a sample from the employees of a single firm. However, the use of a single firm's employees would severely hamper generalizability.

### *2.4.3 Data Analysis*

To answer the first set of research questions (Q1a and Q1b), principal components analysis was utilized. Principal components analysis is a data reduction technique, which can be utilized to extract a reduced number of components out of the 24 measures included in the case. Principal components analysis was run on the questions indicating the importance of each measure in evaluating the organization. In order to interpret the results more easily, an oblique rotation was employed (Hair et al. 1998). Oblique rotations are utilized when correlation among the components is anticipated. They are also utilized when the goal of the principal components analysis is to obtain theoretically meaningful components or constructs (Hair et al. 1998). The Direct Oblimin oblique rotation method was utilized for this analysis. Instead of determining the number of components a priori and forcing the analysis to extract four components as suggested by Kaplan and Norton (1992, 1996, 2001a, 2001c), the latent root criterion (eigenvalue greater than 1) was utilized in determining the number of components.

To examine Q2, a paired sample t-test was conducted comparing the mean importance of selected nonfinancial variables to the mean importance of the two main financial variables, net profit per barrel and return on invested capital.

## **2.5 Results**

### *2.5.1 Principal Components Analysis*

Principal components analysis was run twice in an attempt to obtain the cleanest possible interpretation of the dimensions. The first run included all 24 measures. Variables where the component loadings (correlation of the individual variable with the factor or component) were less than plus or minus 0.4 in the first run were deleted in the second run. The deleted variables were net profit per barrel and barrels produced per labor hour. Table 2-1 includes the results of the

second principal components analysis.<sup>2</sup> In Table 2-1 loadings less than plus or minus 0.5 are deemed not to be of practical importance and are omitted. Bartlett's Test of Sphericity indicated a significant level of correlation among the variables ( $p = 0.000$ ), a requirement for principal components analysis. Another measure of the degree of correlation among the variables and the appropriateness of principal components analysis is the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (MSA). The MSA score was 0.660. A MSA of 0.660 is within the acceptable range (Hair et al. 1998).

In this case, eight separate components were identified containing 18 of the 24 measures. The Performance to Schedule component includes measures of load item accuracy and load schedule performance. The Growth component contains the measures for sales growth in both baseline markets and incremental markets as well as the annual market share increase measure. The Volume component contains the revenue per barrel measure and the throughput per month measure. The Face to the Outside World component represents the organization's image. Things like advertising and community volunteer work shape image. The component includes selling, general & administrative cost per barrel and a measure of employee hours doing community volunteer work. The Safety component contains both safety measures. The Productivity component contains measures of plant productivity, production stability, and beer waste & package scrap. The Employee Skills component contains the training hours per employee measure and the skills inventory measure. The Quality component includes the new brands introduced each year and the quality index measurement. The six variables with insufficient loadings in Table 2-1 are net profit per barrel, barrels produced per labor hour, warehouse moves (actual shipments as a percent of

---

<sup>2</sup> The differences between the two analyses were minimal, except for the deletion of net profit per barrel and barrels produced per labor hour. The second principal components analysis contained the same number of components with exactly the same variables loading on each component with the exception of the two variables that were excluded.

plan), customer complaints, return on invested capital, and manufacturing cost per barrel.

The loadings of the 18 variables contained within the eight components were relatively strong. Variables with a loading greater than plus or minus 0.5 are considered to have practical significance (Hair et al. 1998). Two of the components contained both positive and negative loadings. The new products measure (new brands introduced each year) loaded negatively on the Quality component suggesting a perceived adverse effect on quality when new products are introduced. Also, selling, general & administrative cost per barrel loaded negatively on the Face to the Outside World component. To interpret this negative loading one should keep in mind that selling, general & administrative cost contains advertising, a significant expense for beer companies. A possible explanation might suggest beer advertising reflects negatively on the organization while community involvement represents a more positive image.

### 2.5.2 *T-Test*

Table 2-1 also contains the mean importance of each variable. To determine if the participants regarded financial measures to be more important than nonfinancial measures, a paired sample t-test was conducted. See Table 2-2.<sup>3</sup> The mean importance of the 18 variables remaining after the principal components analysis was used as a proxy for the mean importance of nonfinancial measures. The mean importance of the net profit per barrel and return on invested capital was used as a proxy for mean importance of financial measures. Participants rated the importance of the two financial measures with a mean of 5.01 (S.D. 0.843) and the nonfinancial measures with a mean of 4.36 (S.D. 0.472). The means were significantly different ( $p = 0.000$ ) indicating participants believe

---

<sup>3</sup> The six-point scale and sample size of 84 creates problems assuming normality of the data. Therefore, the p-values associated with the paired sample t-test should be interpreted with caution.

financial measures are more important. Further evidence can be found in that 67 of the 84 participants (80%) rated the financial measures more important.

## **2.6 Discussion**

The evidence obtained by this study fails to support the four perspectives defined by Kaplan and Norton (1992, 1996, 2001a, 2001c) by suggesting the eight components identified in Table 2-1. The results support an idea proposed by Olve, Roy and Wetter (1999) and Simons (2000), that the number of perspectives is situational. The number of perspectives in a performance measurement system should depend on strategies, competitive threats, and economic conditions facing the organization. According to Olve et al. (1999, 120), “If as we have indicated the scorecard [or performance measurement system] should help us to guide the development of the business, it is natural to consider changing the number of perspectives or area of focus.” The goal of a performance measurement system is to communicate and implement strategy (Anthony and Govindarajan 2001; Olve et al. 1999; Kaplan and Norton, 1996, 2001a, 2001b, 2001c). As Hoque and James (2000) had contemplated, strategy appears to be the driving force behind some of the extracted components. At least four of the components can be traced to organizational strategies outlined in the case. Quality and Employee Skills are mentioned in the case as two of the fundamentals on which to focus. Growth and Productivity are listed as individual strategies in the case. Participants appear to have linked organizational strategies provided in the case with available performance measures. These strategy-related components (perspectives) are evidence of the movement away from traditional financial control and management control systems to systems of strategic control.

Anthony and Govindarajan (2001) discuss the difference between outcome and driver measures. Outcome measures are the result of a strategy and are lagging indicators, while driver measures indicate success in implementing a strategy and are leading indicators. Six variables were omitted from Table 2-1 due

to poor component loadings (net profit per barrel, barrels produced per labor hour, warehouse moves, customer complaints, return on invested capital, and manufacturing cost per barrel). Most could be considered outcome measures. Manufacturing cost per barrel, barrels produced per labor hour, and warehouse moves (expressed as actual shipments as a percentage of plan) may be viewed as the result or outcome of the Productivity component. Customer complaints are an outcome of poor quality. Return on invested capital and net profit per barrel are the result of a combination of all the components. The variables remaining in Table 2-1 are all driver measures or at least more closely associated with performance drivers. Participants do not dispute the importance of outcome measures, particularly the omitted financial measures. The results of the paired t-test clearly indicate the relative importance of financial measures. However, it is interesting that these outcome measures fail to load significantly on any component. This suggests that success in bottom-line financial measures is perceived to be a result of controlling nonfinancial driver-type measures.

## **2.7 Research and Limitations**

This study used a single, specific case with 24 predefined measures. The use of a case exercise with MBA students as subjects might cause problems with the generalizability of the findings. However, the authors feel the risk is minimal in this particular case, because of the experienced nature of our research participants. The authors also feel this methodology was the best available, in light of our research objectives. A realistic case study exercise was presented to graduate students, making every effort to obtain MBA students with “real-world” work experience.<sup>4</sup> Nevertheless, we recognize that the use of students raises the question of validity. This seemingly endless debate over subject types and validity

---

<sup>4</sup> Subjects in this study had a mean of 7.3 years of work experience. Additionally, 39% of the participants’ current or last position was in some form of managerial or supervisory capacity.

is likely to be resolved only by subsequent replications where both types of subjects, students and working professionals, are used.

## 2.8 References

- Anthony, R. N. and V. Govindarajan. 2001. *Management Control Systems*, 10<sup>th</sup> ed. New York, NY: The McGraw-Hill Companies, Inc.
- Epstein, M. J. and J. Manzoni, 1997. The balanced scorecard and tableau de Bord: translating strategy into action. *Management Accounting* 79 (2): 28-36.
- Grove, H., T. Cook, and K. Richter. 2000. Coors case: balanced scorecard in W. Bremser and J. Mackey (eds.) *Cases from Management Accounting Practice*, Volume 15. Montvale, NJ: Institute of Management Accountants.
- Hair, J. F. Jr., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. *Multivariate Data Analysis*, 5<sup>th</sup> ed. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Hoque, Z. and W. James. 2000. Linking balanced scorecard measures to size and market factors: impact on organizational performance. *Journal of Management Accounting Research* 12: 1-17.
- Ittner, C. D., and D. F. Larcker. 1998. Innovations in performance measurement: trends and research implications. *Journal of Management Accounting Research* 10: 205-238.
- Kaplan, R. S., and D. P. Norton. 1992. The balanced scorecard – Measures that drive performance. *Harvard Business Review* (January-February): 71-79.
- and -----, 1996. *Translating Strategy into Action: The Balanced Scorecard*. Boston, MA: Harvard Business School Publishing Corporation.
- and -----, 2001a. Transforming the balanced scorecard from performance measurement to strategic management: Part I. *Accounting Horizons* 15 (1): 87-104.
- and -----, 2001b. Transforming the balanced scorecard from performance measurement to strategic management: Part II. *Accounting Horizons* 15 (2): 147-160.
- and -----, 2001c. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, MA: Harvard Business School Publishing Corporation.
- McWilliams, B. 1996. The measure of success. *Across the Board* (February): 16-20.



- Olve, N., J. Roy, and M. Wetter. 1999. *Performance Drivers: A Practical Guide to Using the Balanced Scorecard*. West Sussex, England: John Wiley & Sons, Ltd.
- Porth, S.J., R. Kathuria, and P. J. Maheshkumar. 1998. Performance impact of the fit between manufacturing priorities of general managers and manufacturing managers. *The Journal of Business and Economic Studies* 4 (1): 13-35.
- Silk, S. 1998. Automating the balanced scorecard. *Management Accounting* 79 (11): 38-44.
- Simons, R. 2000. *Performance Measurement and Control Systems for Implementing Strategy*. Upper Saddle River, NJ: Prentice-Hall, Inc.

## 2.9 Figures and Tables

**Figure 2-1 Suggested No. of Measures per Balanced Scorecard Perspective**

Perspective	# of Measures	Weight
Financial	5	22%
Customer	5	22%
Internal Business Processes	8 to 10	34%
Learning & Growth	5	22%

(Source: Kaplan and Norton 2001b, 375)

**Table 2-1 Descriptive Statistics and Principal Components Analysis of the Perceived Importance of the Performance Measures  
(n = 84)**

			Component Loadings Greater than plus or minus 0.5 after Oblique Rotation							
Item	Mean	Std. Dev.	1 Performance to Schedule	2 Growth	3 Volume	4 Face to the Outside World	5 Safety	6 Productivity	7 Employee Skills	8 Quality
<b>Variables with loadings greater than plus or minus 0.5</b>										
Load Item Accuracy	4.49	0.843	0.876							
Load Schedule Performance	4.23	0.998	0.824							
Baseline Growth	4.96	0.963		0.866						
Incremental Growth	4.81	0.885		0.798						
Annual Market Share Increase	4.82	0.946		0.599						
Revenue per Barrel	4.56	0.998			0.828					
Throughput per Month	4.24	0.965			0.516					
Selling, General & Admin. Cost per Barrel	4.38	0.993				-0.794				
Community Involvement	2.81	1.197				0.734				
Safety (lost work incident rate)	4.00	1.232					0.947			
Safety (total case incident rate)	3.95	1.231					0.935			
Plant Productivity	4.80	0.941						-0.752		
Production Stability	4.37	1.128						-0.638		
Beer Waste & Package Scrap	3.95	1.289						-0.575		
Training	3.95	1.040							-0.907	
Skills	4.08	1.143							-0.820	
New Products	3.63	1.180								-0.743
Quality Index	4.98	1.041								0.550
<b>Variables with loadings less than plus or minus 0.5</b>										
Net Profit per Barrel	5.18	1.066								
Barrels Produced per Labor Hour	4.25	1.160								
Warehouse Moves	4.08	1.164								
Customer Complaints	4.82	1.055								
Return on Invested Capital	4.85	1.000								
Manufacturing Cost per Barrel	4.98	1.029								

**Table 2-2 Paired Samples T-Test of the Difference in the Perceived Mean Importance of Financial versus Nonfinancial Measures  
(n = 84)**

Item	Mean	Standard Deviation	t stat.	df	Significance (2-tailed)
Nonfinancial Measures <sup>(a)</sup>	4.36	0.472			
Financial Measures <sup>(b)</sup>	5.01	0.843			
Difference	-0.65	0.770	-7.79	83	0.000

<sup>(a)</sup> The following measures were used as a proxy for the mean importance of nonfinancial measures: load item accuracy; load schedule performance; baseline growth; incremental growth; annual market share; revenue per barrel; throughput per month; selling, general & administrative cost per barrel; community involvement; safety (lost work incident rate); safety (total case incident rate); plant productivity; production stability; beer waste & package scrap; training; skills; new products; and quality index.

<sup>(b)</sup> Net profit per barrel and return on invested capital were used as a proxy for the mean importance of financial measures.

### **3 JUDGMENTAL EFFECTS IN THE USE OF PERFORMANCE MEASUREMENT SYSTEMS**

#### **3.1 Introduction**

Users of management accounting information have been critical of organizations whose performance measurement system is dominated by traditional financial measures. Traditional financial measures are often characterized as being too historical or backward-looking (Ittner and Larcker 1998b). In 1994, the American Institute of Certified Public Accountants called for organizations to report more forward-looking information and nonfinancial measures for key business processes. More recently, the American Accounting Association Financial Accounting Standards Committee (2002) has urged the Financial Accounting Standards Board to encourage the reporting of nonfinancial measures in order to increase the relevance of external reporting.

Several studies have examined the value-relevance and predictive ability of nonfinancial performance information. Amir and Lev (1996) examined the value-relevance of nonfinancial information in the cellular phone industry. The nonfinancial information was positively associated with stock prices. Amir and Lev (1996, 5) state, “In the cellular industry, the value-relevance of nonfinancial

information overwhelms that of traditional, financial indicators. ... we expect this to be the case in other science-based, high-growth sectors.” Ittner and Larcker (1998a) provide evidence that customer satisfaction is positively related to market value and Dempsey et al. (1997) provides evidence financial analysts use or want to use a wide range of nonfinancial information. These sentiments are echoed by Birchard (1995, 43) in his comments concerning a recent survey. He states, “According to a recent survey, 80 percent of large American companies want to change their performance measurement systems. No wonder. Yesterday’s accounting results say nothing about the factors that actually help grow market share and profits—things like customer service innovation, R&D effectiveness, the percent of first-time quality, and employee development.”

In order to overcome perceived limitations of managing solely with traditional financial measures, many firms have adopted new performance measurement systems that utilize a combination of financial and nonfinancial measures. Probably the most widely used of these new systems is the Balanced Scorecard (BSC), introduced by Kaplan and Norton in 1992. It combines financial and nonfinancial measures in a more “balanced” approach to performance measurement. A survey conducted by Renaissance Worldwide, Inc. estimated 60% of Fortune 1000 companies have implemented or are experimenting with the BSC (Silk 1998).

Management accountants have a role to play in the development, implementation, and use of these new performance systems. Barsky and Bremser (1999, 12) state, “Financial managers will be called upon to integrate diverse sets of data and provide sophisticated analysis and support for critical business decisions.” Kaplan (1995) believes it is possible for management accountants to become part of the management team, participating in the formulation and implementation of strategy. Management accountants can be involved in translating strategies into operational measures and the design of new information

systems, according to Kaplan (1995). Additionally, Barsky and Bremser (1999) suggest management accountants should take the lead in the measurement and management of business risk. The Institute of Management Accountants' 1999 Practice Analysis, *Counting More, Counting Less*, provides evidence management accountants are spending more time than ever before as internal consultants, business analysts, and even partners in decision-making processes. According to their survey, nearly 80 per cent of management accountants spend more time actively involved in business decisions than five years earlier. Also, 82 per cent of those surveyed expect to spend greater time in the next three years actively involved in business decisions (Siegel and Sorenson 1999).

This paper presents two analyses examining accountants' cognitive difficulties and judgmental effects in the use of performance measurement systems containing both financial and nonfinancial measures. It uses accounting undergraduate and graduate students as a proxy for accounting professionals. Accountants are the chief custodian, user, and disseminator of performance measurement information and they, to a large extent, influence other manager's attitudes towards organizational performance. In this sense, accounting students are also a proxy for other professionals in the organization.

The results of the first analysis suggest users of performance measurement systems with multiple measures will not alter the relative weight placed on a particular measure, regardless of its position vis-à-vis the targeted or expected results for that particular measure. In the second analysis, participants appear to have over-weighted the financial measures and underweighted the nonfinancial measures in evaluating overall organizational performance. Evidence is also provided of a halo effect in that an organization's performance on financial measures appears to influence an individual's perception of the organization's performance on nonfinancial measures.

The remainder of this paper is organized as follows. The following section contains a literature review and presents the research questions, the next section discusses the methodology, and the next section presents the results. Following the results is a section discussing implications of the study.

### **3.2 Literature Review and Research Question Development**

Use of a BSC or similar performance measurement system where both financial and nonfinancial measures are employed is considered a central element in a modern strategic management system (Kaplan and Norton 2001a, 2001c). Implementing the BSC requires management to (1) develop coherent strategies to achieve their objectives and (2) develop a set of financial and nonfinancial measures for gauging the organization's performance in implementing those strategic objectives (Kaplan and Norton 1992, 1996, 2001a, 2001b, 2001c). The BSC usually contains 20 – 25 key measures that compare the organization's performance to planned targets. It typically is organized into four categories of measures (Kaplan and Norton 2001a, 90):

1. *Financial – the strategy for growth, profitability, and risk viewed from the perspective of the shareholder.*
2. *Customer – the strategy for creating value and differentiation from the perspective of the customer.*
3. *Internal Business Processes – the strategic priorities for various business processes that create customer and shareholder satisfaction.*
4. *Learning and Growth – the priorities to create a climate that supports organizational change, innovation, and growth.*

The BSC gauges progress in the achievement of strategic objectives and helps to define a strategy's cause-and-effect relationships in relation to the four perspectives.

Managing performance with a set of several measures could be considered inherently more complex than managing with a single bottom-line financial measure and could therefore introduce potential cognitive difficulties. Recent



literature on the BSC has explored some of the potential cognitive difficulties in the use of these more balanced performance measurement systems (e.g. Lipe and Salterio 2000, 2002; Tully et al. 2002).

Lipe and Salterio (2000) was the first study to examine cognition in the use of the BSC. MBA students were asked to evaluate the performance of two division managers of a clothing firm based on a review of each division's BSC. The case materials documented that the two divisions (RadWear and WorkWear) were pursuing different strategies, thus allowing for the scorecards to contain some measures that were unique to each division. Each scorecard contained 16 measures—eight were common to both divisions and eight unique to each division. Each participant evaluated both divisions (the within-subjects factor). The two between-subjects factors were the division's performance (relative to the other division) on the common and unique measures. Lipe and Salterio (2000) found measures unique to each division were not utilized in the evaluations. A potential ramification of these findings is that managers may not utilize strategically important measures unique to a particular division. A second ramification deals with the use of financial versus nonfinancial measures. Because the unique measures are often nonfinancial, managers may rely too heavily on the common financial measures when evaluating performance.

Tully et al. (2002) utilized MBA and accounting undergraduate students to test the robustness of the Lipe and Salterio (2000) findings. In Lipe and Salterio (2000), all of the performance measures for both divisions exceeded target and the differences between the divisions as compared to the target was approximately 4.1% points (Tully et al. 2002). In other words, the superior division's measures averaged approximately 10.6% above target and the lesser division averaged approximately 6.5% above target. Tully et al. (2002) proposed the differences between the divisions may not have been sufficiently large enough to allow detection. They also proposed a bias may have resulted because all performance

measures exceeded their targets. Tully et al. (2002) discussed a potential goal-centered, make-the-numbers orientation that may confound Lipe and Salterio's results. Performance that exceeds targets or goals may be judged as satisfactory and performance that falls short of the target may be judged as unsatisfactory with very little discrimination within each condition. Lock and Latham (1990, 87) state,

...goals *define* for the individual what an acceptable level of performance or direction of action is. Actions that fall short of desired ends...lead to negative performance evaluations and/or self-evaluations. ... Actions that attain or exceed desired ends lead to positive appraisals.

If a goal centered make-the-numbers confound were in effect, subjects would ignore differences in the unique divisional measures where both exceeded the target but by slightly differing degrees.

Tully et al. (2002) employed a design where all measures were either 20% above or 20% below target in order to test the robustness of the earlier findings. They used basically the same case materials as Lipe and Salterio (2000) except for the above differences in performance to target. Tully et al. (2002) did not find evidence of a goal-centered make-the-numbers orientation and they found the unique measures were utilized and not underweighted in the evaluation decision. They further found the unique measures from each division were perceived differently even when performance to target is the same.

The Tully et al. (2002) and Lipe and Salterio (2000) studies provide evidence of a contingency effect by demonstrating a difference in the utilization of unique measures in evaluation of two divisions within an organization. When unique measures were slightly to moderately above target and the differences between divisions were not large, unique measures were ignored (Lipe and Salterio 2000); but, when unique measures were well above or below target and

varied significantly between divisions, the subjects utilized the unique measures (Tully et al. 2002).

The combination of these two studies suggests a possible relationship between performance relative to target and the perceived importance of a measure. The existence of this relationship could have serious consequences for developers and users of performance measurement systems. If performance relative to target affects an assessment of the importance of a particular measure within the performance measurement system, then a particular measure will have a different relative weight that depends on its status relative to targeted performance. If a measure is over (under) target for a particular period, then it may receive more (less) weight (or visa versa) than in another period when the performance to target is different. Therefore, this study will attempt to answer the following research question.

*Question: Does performance against target for a particular measure affect the perception of the importance of that measure?*

To evaluate overall organizational performance, one must establish relative weights for each of the performance measures in a performance measurement system. The weighting of financial versus nonfinancial measures is especially critical. Financial measures have been criticized as too historical and backward-looking. Nonfinancial measures are perceived to be predictive and forward-looking (Ittner and Larcker 1998a, 1998b) although sometimes perceived to suffer from poor measurement quality (Lingle and Schiemann 1996). Nonfinancial measures are value relevant (Amir and Lev 1996, Ittner and Larcker 1998a). Recognizing these facts, the AICPA (1994) has called for more extensive reporting of nonfinancial performance measures. Kaplan and Norton (2001c, 376) proposed “nearly 80% of the measures on a Balanced Scorecard should be nonfinancial,” suggesting that nonfinancial measures be emphasized in management decision-making and evaluations of performance.

While there is a trend towards a greater reporting of nonfinancial measures, there is some doubt as to whether nonfinancial measures receive the proper weight in evaluating performance. “At Volvo, the principal challenge is considered to be that of giving nonfinancial measures as much weight as the more established financial ones,” report Olve et al. (1999, 119). A 1996 Towers Perrin survey found BSC adopters placed an average of 56% of the relative weight on financial measures (Ittner and Larcker 1998b). DeBusk et al. (2003) found support for the finding that performance measurement system users view bottom-line financial measures as more important than nonfinancial measures.

Does the use of a set of measures overwhelm accountants and managers to such an extent that they must concentrate on traditional financial measures? Studies on information overload suggest a large number of performance measures can hinder a manager’s ability to evaluate the organization (Ittner and Larcker 1998b). Managers may rely on what they know best—financial measures. Anthony and Govindarajan (2001, 451) state,

... not only are most senior managers well trained and adept with financial measures, they also keenly feel pressure regarding the financial performance of their companies. Shareholders are vocal, and boards of directors frequently apply pressure on the shareholders’ behalf. This pressure may overwhelm the long-term, uncertain payback of the nonfinancial measures.

The anecdotal evidence (Olve et al. 1999), survey evidence (Ittner and Larcker 1998b), and data from experiments (DeBusk et al. 2003) suggest that Govindarajan’s analysis may be correct. Users of a performance measurement system may discount the importance of nonfinancial measures and rely more on financial measures. Therefore, the following research hypothesis is examined in this study.

*H: Financial measures will receive a disproportionate share of the weight, as compared to nonfinancial measures, in tasks involving the assessment of overall organizational performance.*

### **3.3 Method**

#### *3.3.1 The Experiment*

A case was developed giving sufficient background data, information on vision and strategies, and measures of performance against targets for a major brewery. The case was modified from an Institute of Management Accountants' case titled *Coors Case: Balanced Scorecard* by Hugh Grove, Tom Cook and Ken Richter. Specific information was given on 24 individual performance measures including a definition of the measure, the target for the year, and the actual results for the year. The measures were organized onto one page and listed alphabetically.<sup>5</sup> Participants were given two packages of information. The first contained the background information (see Appendix A). The second contained the actual survey instrument and included the actual measures (see Appendix B). Participants in the experiment were asked to evaluate the organization's performance on individual measures (variable names are PERF1 ... PERF24) and on an overall basis (ORGPREF) on a scale of one to six, where one equals very poor and six equals very good. Participants, in addition, were asked to assess the importance of each measure in evaluating the organization's performance on a scale of one to six, where one equals not important and six equals critical (variable names are IMPORT1 ... IMPORT24).

To evaluate if performing better or worse than target in varying degrees affects the participant's assessment of the importance of the measure, three

---

<sup>5</sup> This study does not intend to test the effects of organizing measures into the BSC format. Interested parties should consult Lipe and Salterio (2002). In their study, BSC format had no effect in evaluations of two division managers when differences were scattered across the four BSC categories. However, Lipe and Salterio (2002) found that BSC format moderated the evaluations when the differences were concentrated into one BSC category.

different measurement scenarios or versions (VER1, VER2, and VER3) were utilized to alter performance against target. Each version was developed so the individual measures would support an overall scenario where Return on Invested Capital (ROI or measure no.16) was below target (VER1), on target (VER2), or above target (VER3). See Table 3-1. In addition to ROI, four other measures followed the same pattern where in VER1 actual was below target, in VER2 actual was on target, and in VER3 actual was better than target. The other measures were varied in an attempt to support the performance scenario in each version (see Table 3-1). Version 1 performance was generally below target while version 2 performance was generally equal to target. Version 3 performance was better than target for net profit and return on invested capital due to strength in selling, general and administrative costs but other measures were worse than target signaling future profitability problems.

The study was conducted with undergraduate and graduate students at a large state university in the eastern United States. Undergraduate students were obtained from two sections of a Cost Accounting course and graduate students from an Advanced Managerial Accounting course. The case and survey instruments were introduced after a 45 – 60 minute lecture on performance measurement systems focusing on the Balanced Scorecard (BSC) system. The purpose of the lecture was to insure students were familiar with performance measurement systems like the BSC that emphasize nonfinancial measures. Students read the case and completed the surveys as part of a take home exercise.

A total of 67 usable surveys were obtained. Ten surveys were dropped because of missing information. Ninety-one percent of the participants were accounting majors or had accounting listed as part of a double major. Of the 67 participants, 35 (52%) were female and 32 (48%) were male; 44 (66%) were undergraduates and 23 (34%) were graduate students. Graduate students comprise 43% of the version 1 sample, 29% of the version 2, and 30% of the version 3. No

significant differences in responses were noted between the two groups: undergraduates and graduate students.

### 3.4 Results

#### 3.4.1 Data Analysis – Research Question

The first analysis (research question) is a between-subjects comparison of the measurement scenarios in order to ascertain if there are any effects on the participants' evaluations of a measure's importance caused by the organization's performance relative to its target. Multivariate Analysis of Variance (MANOVA) was used to address the research question by testing the equality of the three vectors of participants' responses with respect to the importance of each measure (IMPORT1 ... IMPORT24) for each of the three versions (VER1, VER2, VER3) of the case. Thus the hypothesis being tested in the research question can be stated as follows (in null form):

$$H_0: (\mu_{11} \dots \mu_{241})' = (\mu_{12} \dots \mu_{242})' = (\mu_{13} \dots \mu_{243})'$$

#### 3.4.2 Results – Research Question

Table 3-2 provides descriptive statistics for each of the 24 dependent variables (IMPORT1 ... IMPORT24). Table 3-3 provides the MANOVA results. The significance levels for the test statistics range from 0.339 to 0.586 indicating a failure to reject the null hypothesis in all cases. In other words, no significant differences were detected among the three versions of the case. Observed power levels using  $\alpha = 0.05$  were above the recommended 0.80 threshold (Hair et al. 1998) for all test statistics but Roy's Largest Root, which had an observed power of 0.71. With sufficient power and no significant effects in the MANOVA, we can be confident the effect size was too small to be of statistical significance or practical importance (Hair et al. 1998).

As a validation that the different versions did portray differences in performance, another MANOVA was run on the participant's evaluation of the organization's performance on the 24 individual measures. The results were significant ( $p = 0.000$ ) indicating perceived differences in performance among the three version's 24 measures. The participants could differentiate performance among the versions but there was no differentiation in importance of the various 24 measures among the versions. Table 3-4 shows the mean evaluation scores on the individual performance measures<sup>6</sup> and on the organization's overall performance.<sup>7</sup>

### 3.4.3 Robustness Tests – Research Question

The results from the MANOVA on the importance variables suggest the amount the measure is above or below target does not influence the subject's view of the importance of the measure. Additional evidence can be obtained through an analysis of the correlations of the importance variables (IMPORT1 ... IMPORT24) with their corresponding performance evaluation variables (PERF1 ... PERF24). If performance against target affects the assessed importance of the measure, a substantial correlation should exist. Some correlation will naturally exist since both items are measured on six-point scales. See Table 3-5. All of the absolute values of the correlation coefficients are less than 0.5 and the average absolute value of the correlations is only 0.248 (explaining only 6% of the variation). Thus, the correlations between importance and assessed performance lack practical significance.

---

<sup>6</sup> Two of the 67 cases had missing values on the performance variables. Thus Table 3-4 reflects the means of a total of 65 cases (22 VER1, 21 VER2, and 22 VER3).

<sup>7</sup> Utilizing ANOVA to examine the overall performance variable (ORGPREF), further evidence is provided of significant differences in the versions ( $p = 0.050$ ). The overall performance scores and the average of the 24 individual measures demonstrate that the participants view performance in VER1, VER2 and VER3 differently.



Another test of the robustness of the findings is to run 24 separate ANOVAs comparing the versions on the importance variables (IMPORT1 ... IMPORT24). While this fails to control the total experimentwise Type-1 error, it is appropriate for a robustness test. Only one of the 24 measures is significant at the 0.05 level (results not shown), providing further evidence performance against target does not affect the evaluation of the importance of that particular measure.

#### *3.4.4 Data Analysis – Research Hypothesis*

The second analysis (concerning the research hypothesis) compares version 2 and version 3 in an evaluation of the weighting of financial versus nonfinancial measures. Evidence was obtained by examining performance evaluations between versions 2 and 3. Versions 2 and 3 were selected for analysis because of the tension provided by comparing the financial and nonfinancial measures between the two versions. Versions 2 and 3 each outperform version 1 on both financial and nonfinancial measures and thus provide no tension in studying version 1. However, a comparison of versions 2 and 3 is a study where each version outperforms the other on one category (financial or nonfinancial) of measures. The average performance for the nonfinancial measures is 5% below target in version 2 but 10% below target in version 3; however, financial measures were on target in version 2 (0%) but 13% better than target in version 3.<sup>8</sup>

In version 2, all of the financial measures and most of the nonfinancial measures indicate performance exactly equal to target. In version 3 the favorable financial measures of net profit and ROI are driven by selling, general & administrative costs (SG&A) being \$3 (11%) better than target. Offsetting some of the SG&A savings is worse than targeted results in manufacturing cost. Barrels

---

<sup>8</sup> See Table 3-1. For measures 4, 18, and 19, the target is 0 which presents a problem in the computation of the percentage difference (division by 0). However, actual performance is the same in all three versions. Averages ignore these three variables. Financial variables are 10, 11, 16, 17, and 20.

produced per direct labor hour, throughput per month, load schedule performance, production stability, and plant productivity are all worse in version 3 than in version 2, indicating poor manufacturing and overall productivity in version 3 compared to target and version 2. Worse than targeted and version 2 performances in training and skills measures might help explain some of the productivity problems in version 3 and possibly indicate yet additional future declines in productivity. Cuts in training might have accounted for some of the version 3 savings on the SG&A line but could lead to more manufacturing problems in future years. Cuts in advertising might also account for some of the lower SG&A costs in version 3 and some of the growth problems. Version 3 is also behind version 2 and below target in annual market share increase, incremental growth for high potential brands and markets, and new products introduced each year. These results suggest problems for the future in maintaining revenue growth. Version 3 outperforms version 2 on only three nonfinancial measures, customer complaints (0.04% VER3; 0.05% VER2), quality index (95 VER3; 90 VER2), and load item accuracy (96% VER3; 95% VER2).

The nonfinancial measures indicate version 2 is superior to version 3 (version 2 averages 5% worse than target while version 3 averages 10% worse than target – see Table 3-1). The poor performance in version 3 within the nonfinancial measures signals a future decline in financial performance even though the version 3 financial measures are better than those in version 2 in the current period (average of 13% favorable to target in VER3 versus on target in VER2 – see Table 3-1).

Average performance of all measures is 4% worse than target in version 2 and 5% worse than target in version 3. If each measure is weighted equally, version 2 should be perceived as outperforming version 3. Considering the recommendations for a “balanced approach” and that nonfinancial measures are said to predict future financial performance; overall, the two versions should be

perceived as roughly equivalent, in a best case scenario, or version 2 as slightly superior.

It is difficult to determine the relative weights of individual measures subjects use in making an evaluation of organizational performance; however, we may find evidence of suboptimal weighting by examining their assessment of overall organizational performance. If the study participants assess version 3 as superior to version 2, it would suggest a less than optimal weighting of nonfinancial measures. We can analyze the subjects' perception of the two versions' performance by examining responses on the overall organizational performance variable (ORGPFRF). A t-test is performed on ORGPFRF (version 2 compared to version 3) to test for a difference in perceived overall performance.

#### *3.4.5 Results – Research Hypothesis*

The results of the t-tests to determine if subjects perceive a difference in the overall organizational performance between versions 2 and 3 are given in Table 3-6. The research participants receiving version 3 evaluate the organization better than those receiving version 2 ( $p = 0.037$ ).<sup>9</sup> These results support the hypothesis and suggest a possible over-weighting of the financial measures and a less than optimal weighting of nonfinancial measures.

#### *3.4.6 Robustness Tests – Research Hypothesis*

A paired t-test was performed comparing the mean importance of the five financial measures (4.94) to the mean importance of the 19 nonfinancial variables (4.37). The mean importance of the financial measures was significantly higher ( $p = 0.000$ ).

---

<sup>9</sup> Because 42 of the 44 participants responded with either a 4 or 5 for the rating of organizational performance, splitting the data into two groups ( $\leq 4$  and  $\geq 5$ ) and performing the t-test is appropriate. Nonparametric Mann-Whitney tests yielded similar results.

### 3.4.7 Discussion of Results – Research Hypothesis and Halo Effects

An emphasis on financial results is not the only item driving the view that version 3 outperforms version 2. The results in Table 3-4 reveal participants view version 3 as outperforming version 2 on the nonfinancial measures as well (mean of 4.44 versus 4.31). This is intriguing considering the average performance to target, shown in Table 3-1, is worse in version 3 than version 2 (VER3 -10%; VER2 -5%). See Figure 3-1 for an illustration of the findings. The perceived performance on the nonfinancial measures seems more related to the financial measures than the nonfinancial measures. A MANOVA performed on the assessed performance on the 19 nonfinancial measures indicates a significant difference ( $p = 0.001$ ) between participants' assessment of performance on the two versions.<sup>10</sup>

Pedhazur and Schmelkin (1991, 121) state, "One of the most common sources of bias is the halo effect, a constant error that occurs when raters' general impressions bias their ratings of distinct aspects of the ratees." Figure 3-1 can be explained by version 3 subjects forming a general impression of the organization based on the financial results. This general impression biases their ratings on the version 3 nonfinancial results, a halo effect.

Evidence of a halo effect has been found in other studies involving performance measurement. In Nelson et al. (1992), it was noted that a halo effect caused hospital patients to rate the hospital as generally high or low on all measures. Herman and Renz (1997) expected to find multiple dimensions in measuring the effectiveness of non-profit organizations. All of the measures,

---

<sup>10</sup> Version 3 did outperform version 2 on three nonfinancial measures. To determine if this was causing the effect, another analysis was conducted excluding those nonfinancial measures where version 3 outperformed version 2. On the 16 nonfinancial measures where version 3 performed the same as or worse than version 2, the mean performance was 4.29 for version 2 and 4.35 for version 3. MANOVA results indicated a significant difference ( $p = 0.005$ ).

however, loaded on a single factor with a high degree of intercorrelation among the items (Cronbach alpha was .85). This unidimensionality and high intercorrelation is indicative of a halo effect.

Additional ad hoc analysis was performed on the subjects' assessment of performance on the 24 individual performance measures. Principal components analysis did reveal the presence of multiple dimensions among the measures. However, there was a high amount of intercorrelation. The Cronbach alpha was .88 for all three versions combined. Alphas for each version ranged from .85 to .90. In version 3, one should not expect perception of financial performance to be correlated with perceived nonfinancial performance since financial performance was above target on average and nonfinancial performance was below target on average. The subject's overall perception of financial performance in version 3 was highly correlated with their perception of performance on the nonfinancial measures (Pearson Correlation equals 0.42;  $p = 0.052$  2-tailed) and thus provides more evidence of a halo effect.

### **3.5 Discussion and Conclusion**

This study extends the findings of Tully et al. (2002). No evidence was found of a goal centered, making-the-numbers orientation or any other budget (target) related effect. An individual's perception of a measure's importance was not affected by the measure being below target, on target, or in excess of target or by the magnitude of the difference from target. These results are important because they suggest an accountant or manager can count on a measure being perceived to have the same relative weight in a performance measurement system regardless of its position relative to its target (i.e. the measures will not be over or under-weighted because of relative performance to target).

The bad news is that users of performance measurement systems may be expected to emphasize historical financial measures. The predictive ability of

nonfinancial measures has been a major factor in the development of performance measurement systems that contain both financial and nonfinancial measurements. In this study, subjects failed to place the suggested emphasis on the leading nonfinancial measures which signaled problems in productivity and growth for the organization in version 3. The participants evaluating version 3 seem to be fixated on the financial measures and allowed the financial results to bias their view of the nonfinancial results (i.e. halo effect).

The results suggest that when confronted with multiple measures of organizational performance, a person will form a general impression based on the financial results. This general impression will distort their perception of performance on the remaining nonfinancial measures. This bias was present even in the presence of a potential debiaser—participants were given training that promoted the predictive ability of nonfinancial measures.

The generalizability of this study may be hampered because of two limitations. First, students were used as subjects. Students lack the experience of managers and may use different criteria in evaluating an organization. Also, they may be more susceptible to the effects of information overload due to their lack of experience. Future research could extend these findings using experienced managerial accountants. Second, the study used a single, specific case with 24 predefined measures. It could be argued that a different case with different facts and different measures may have yielded different results.

### 3.6 References

- American Accounting Association. 2002. Recommendations on disclosure of nonfinancial performance measures. *Accounting Horizons* 16 (4): 353-362.
- American Institute of Certified Public Accountants (AICPA). 1994. *Improving business reporting – a customer focus*. New York, NY: AICPA.
- Amir, E., and B. Lev. 1996. Value-relevance of nonfinancial information: the wireless communications industry. *Journal of Accounting & Economics* 22: 3-30.
- Anthony, R. N., and V. Govindarajan. 2001. *Management Control Systems, 10<sup>th</sup> ed.*, New York, NY: The McGraw-Hill Companies, Inc.
- Barsky, N. P., and W. G. Bremser. 1999. Performance measurement, budgeting and strategic implementation in the multinational enterprise. *Managerial Finance* 25 (2): 3-15.
- Birchard, B. Making it count. *CFO* (October): 42-51.
- DeBusk, G. K., R. M. Brown, and L. N. Killough. 2003. Components and relative weights in utilization of performance measurement systems like the balanced scorecard. *British Accounting Review* 35 (3): 215-231.
- Dempsey, S. J., J. F. Gatti, D. J. Grinnell, and W. L. Cats-Baril. 1997. The use of strategic performance variables as leading indicators in financial analysts' forecasts. *The Journal of Financial Statement Analysis* (Summer): 61-79.
- Grove, H., T. Cook, and K. Richter. 2000. Coors case: balanced scorecard in W. Bremser and J. Mackey (eds.) *Cases from Management Accounting Practice*, Volume 15. Montvale, NJ: Institute of Management Accountants.
- Hair, J. F. Jr., R. E. Anderson, R. L. Tatham, and W.C. Black. 1998. *Multivariate Data Analysis, 5<sup>th</sup> ed.*, Upper Saddle River, NJ: Prentice-Hall, Inc.
- Herman, R. D. and D. O. Renz. 1999. Theses on nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly* 28 (2): 107-126.
- Ittner, C. D., and D. F. Larcker. 1998a. Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research* 36 (Supplement): 1-35.

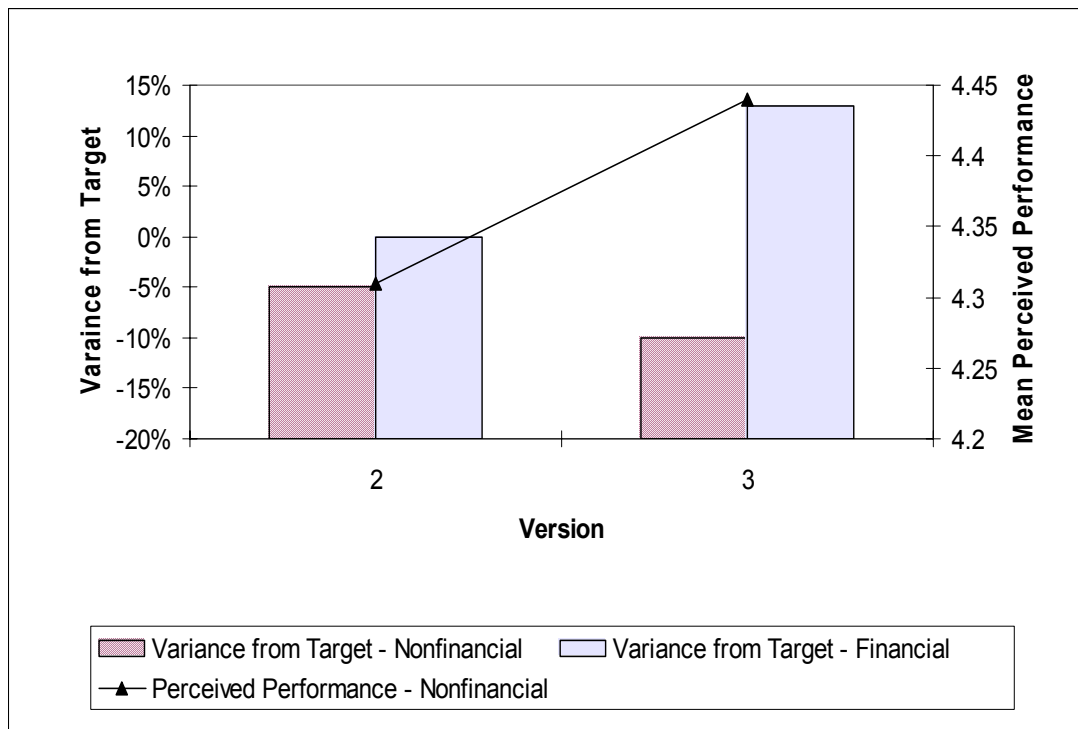
- Ittner, C. D., and D. F. Larcker. 1998b. Innovations in performance measurement: trends and research implications. *Journal of Management Accounting Research* 10: 205-238.
- Kaplan, R. S. (1995). New roles for management accountants. *Journal of Cost Management* (Fall), 6-13.
- Kaplan, R. S., and D. P. Norton. 1992. The balanced scorecard – Measures that drive performance. *Harvard Business Review* (January-February): 71-79.
- and -----, 1996. *Translating Strategy into Action: The Balanced Scorecard*. Boston, MA: Harvard Business School Publishing Corporation.
- and -----, 2001a. Transforming the balanced scorecard from performance measurement to strategic management: Part I. *Accounting Horizons* 15 (1): 87-104.
- and -----, 2001b. Transforming the balanced scorecard from performance measurement to strategic management: Part II. *Accounting Horizons* 15 (2): 147-160.
- and -----, 2001c. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, MA: Harvard Business School Publishing Corporation.
- Lingle, J. H., and W. A. Schiemann. 1996. From balanced scorecard to strategic gauges: is measurement worth it? *Management Review* (March): 56-61.
- Lipe, M. S., and S. E. Salterio. 2000. The balanced scorecard: the judgmental effects of common and unique performance measures. *The Accounting Review* 75 (3): 283-298.
- and -----, 2002. A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society* 27: 531-540.
- Locke, E. A., and G. P. Latham. 1990. *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice Hall.
- Nelson, E. C., R. T. Rust, A. Zahorik, R. L. Rose, P. Batalden, and B. A. Siemanski. 1992. Do patient perceptions of quality relate to hospital performance? *Marketing Health Services* 12 (4): 6-13.



- Olve, N., J. Roy, and M. Wetter. 1999. *Performance Drivers: A Practical Guide to Implementing the Balanced Scorecard*. West Sussex, England: John Wiley & Sons Ltd.
- Pedhazur, E. J. and L. P. Schmelkin. 1991. *Measurement, Design, and Analysis: An Integrated Approach*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Siegel, G., and J. E. Sorenson. 1999. *Counting More, Counting Less: Transformations in the Management Accounting Profession*, Montvale, NJ: Institute of Management Accountants.
- Silk, S. 1998. Automating the balanced scorecard. *Management Accounting* 79 (11): 38-44.
- Tully, G. T., M. L. Gagne, and J. Hollister. 2002. The judgmental effects of balanced scorecard performance measures: evidence of contingencies. Working paper, Marist College.

### 3.7 Figures and Tables

**Figure 3-1 Mean Perceived Performance on Nonfinancial Measures  
(Compared to Variance from Target for both Financial and Nonfinancial Measures)**



**Table 3-1 Organizational Performance Measures for Each Version of the Case**

No.	Performance Measure	Target	Actual 2000			Actual vs. Target <sup>(a)</sup>		
			Version			Version		
			1	2	3	1	2	3
1	Annual Mkt. Share Increase	0.05%	0.04%	0.05%	0.04%	-20%	0%	-20%
2	Barrels Produced per DLH	6.0	5.0	6.0	5.8	-17%	0%	-3%
3	Baseline Growth (for key brands and markets)	2.0%	2.0%	2.0%	2.0%	0%	0%	0%
4	Beer Waste & Package Scrap	0%	0.5%	0.5%	0.5%			
5	Community Involvement (volunteer hours per employee annually)	30	25	30	25	-17%	0%	-17%
6	Customer Complaints (per 100,000 barrels sold)	0.05%	0.06%	0.05%	0.04%	-20%	0%	20%
7	Incremental Growth (for high potential brands and markets)	3.0%	2.5%	3.0%	2.5%	-17%	0%	-17%
8	Load Item Accuracy	100%	93%	95%	96%	-7%	-5%	-4%
9	Load Schedule Performance	100%	55%	65%	60%	-45%	-35%	-40%
10	Mfg. Cost per Barrel	\$53	\$55	\$53	\$54	-4%	0%	-2%
11	Net Profit per Barrel	\$6	\$3	\$6	\$8	-50%	0%	33%
12	New Products (new brands introduced each year)	6	5	6	5	-17%	0%	-17%
13	Plant Productivity	80%	76%	80%	78%	-5%	0%	-3%
14	Production Stability	100%	50%	65%	60%	-50%	-35%	-40%
15	Quality Index (out of 100)	90	85	90	95	-6%	0%	6%
16	Return on Invested Capital	12.0%	7.1%	12.0%	14.5%	-41%	0%	21%
17	Revenue per Barrel	\$100	\$100	\$100	\$100	0%	0%	0%
18	Safety (lost work incident rate)	0	0.0001	0.0001	0.0001			
19	Safety (total case incident rate)	0	0.0002	0.0002	0.0002			
20	S, G & A Cost per Barrel	\$27	\$28	\$27	\$24	-4%	0%	11%
21	Skills (inventory of cross-functional employee skills)	7	6	7	6	-14%	0%	-14%
22	Throughput per Month (millions of barrels)	1.75	1.72	1.75	1.74	-2%	0%	-1%
23	Training (hours per employee annually)	40	42	40	38	5%	0%	-5%
24	Warehouse Moves (actual shipments as a % of plan)	100%	95%	95%	95%	-5%	-5%	-5%
	Average					-16%	-4%	-5%
	Avg. of nonfinancial measures					-15%	-5%	-10%
	Avg. of financial measures					-20%	0%	13%

<sup>(a)</sup> For measures 4, 18, and 19, the target is 0 which presents a problem in the computation of the percentage difference (division by 0). Averages ignore these three variables.

**Table 3-2 Descriptive Statistics – Perceived Importance  
of Individual Measures  
(n = 23, 21, and 23 for Versions 1, 2, and 3 Respectively)**

No.	Performance Measure	Mean Importance				Standard Deviation			
		Version				Version			
		1	2	3	All	1	2	3	All
1	Annual Mkt. Share Increase	4.83	4.52	5.17	4.85	0.887	1.327	0.717	1.019
2	Barrels Produced per DLH	4.57	4.29	4.43	4.43	0.728	1.231	1.121	1.033
3	Baseline Growth (for key brands and markets)	4.57	4.81	4.70	4.69	1.080	0.981	0.926	0.988
4	Beer Waste & Package Scrap	3.78	3.76	4.04	3.87	1.126	1.136	1.261	1.166
5	Community Involvement (volunteer hours per employee annually)	2.74	2.95	2.96	2.88	1.176	1.396	1.065	1.200
6	Customer Complaints (per 100,000 barrels sold)	5.00	4.81	5.13	4.99	0.853	1.078	0.757	0.896
7	Incremental Growth (for high potential brands and markets)	4.57	4.71	4.96	4.75	0.945	1.007	0.767	0.910
8	Load Item Accuracy	4.48	4.10	4.39	4.33	0.898	0.889	0.891	0.894
9	Load Schedule Performance	4.26	4.10	4.09	4.15	0.964	0.889	0.900	0.909
10	Mfg. Cost per Barrel	5.17	4.95	5.04	5.06	0.778	0.805	0.706	0.756
11	Net Profit per Barrel	5.22	5.05	5.43	5.24	0.671	0.805	0.662	0.720
12	New Products (new brands introduced each year)	4.13	4.14	4.22	4.16	1.180	1.195	0.998	1.109
13	Plant Productivity	4.65	4.86	4.70	4.73	0.775	0.964	0.822	0.845
14	Production Stability	4.57	4.48	4.43	4.49	0.843	0.873	0.728	0.805
15	Quality Index (out of 100)	4.96	4.71	4.52	4.73	0.928	0.902	1.238	1.038
16	Return on Invested Capital	4.57	5.00	5.39	4.99	1.273	1.049	0.722	1.080
17	Revenue per Barrel	4.87	4.90	4.96	4.91	0.968	0.889	1.107	0.981
18	Safety (lost work incident rate)	4.04	4.48	4.09	4.19	1.065	1.123	0.949	1.048
19	Safety (total case incident rate)	4.22	4.10	3.91	4.07	1.043	1.338	1.083	1.146
20	S, G & A Cost per Barrel	4.61	4.33	4.52	4.49	0.891	1.017	1.039	0.975
21	Skills (inventory of cross-functional employee skills)	4.57	4.67	4.48	4.57	0.945	1.017	0.947	0.957
22	Throughput per Month (millions of barrels)	4.22	4.76	4.48	4.48	0.850	0.889	0.994	0.927
23	Training (hours per employee annually)	4.65	4.81	4.35	4.60	1.112	0.981	0.982	1.031
24	Warehouse Moves (actual shipments as a % of plan)	4.13	4.24	3.91	4.09	1.254	0.831	1.379	1.177
	Average	4.47	4.48	4.51	4.49				
	Avg. of nonfinancial measures	4.37	4.38	4.37	4.37				
	Avg. of financial measures	4.89	4.85	5.07	4.94				

**Table 3-3 Multivariate Analysis of Variance on Perceived Importance  
of Individual Performance Measures  
(n = 67)**

Effect		Value	F	Hypothesis df	Error df	Sig.	Observed Power <sup>(a)</sup>
VER	Pillai's Trace	0.717	0.977	48.000	84.000	0.527	0.870
	Wilks' Lambda	0.410	0.958	48.000	82.000	0.556	0.858
	Hotelling's Trace	1.128	0.940	48.000	80.000	0.586	0.845
	Roy's Largest Root	0.656	1.148	24.000	42.000	0.339	0.712

<sup>(a)</sup> Computed using alpha = 0.05

**Table 3-4 Descriptive Statistics – Perceived Performance  
on Individual Measures and Overall  
(n = 22, 21, and 22 for Versions 1, 2, and 3 Respectively)**

No.	Performance Measure	Mean Evaluation				Standard Deviation			
		Version				Version			
		1	2	3	All	1	2	3	All
1	Annual Mkt. Share Increase	4.27	4.76	4.73	4.58	1.120	0.995	0.767	0.983
2	Barrels Produced per DLH	3.91	4.48	4.68	4.35	0.684	1.030	0.839	0.909
3	Baseline Growth (for key brands and markets)	4.86	4.76	5.36	5.00	0.990	0.889	0.727	0.901
4	Beer Waste & Package Scrap	3.86	3.76	3.68	3.77	1.037	1.338	1.287	1.209
5	Community Involvement (volunteer hours per employee annually)	4.05	4.57	3.95	4.18	0.844	1.165	0.899	0.998
6	Customer Complaints (per 100,000 barrels sold)	4.27	4.62	5.27	4.72	1.202	1.161	0.767	1.125
7	Incremental Growth (for high potential brands and markets)	3.73	4.67	4.32	4.23	0.935	1.017	0.894	1.012
8	Load Item Accuracy	3.82	3.95	4.27	4.02	1.220	0.921	0.767	0.992
9	Load Schedule Performance	2.23	2.52	2.77	2.51	1.572	1.365	1.152	1.371
10	Mfg. Cost per Barrel	4.05	4.67	4.59	4.43	0.844	0.913	1.008	0.951
11	Net Profit per Barrel	2.86	4.90	5.50	4.42	1.167	1.044	0.859	1.530
12	New Products (new brands introduced each year)	4.50	4.48	4.64	4.54	0.740	1.327	0.658	0.937
13	Plant Productivity	4.14	4.52	4.68	4.45	1.037	0.928	0.716	0.919
14	Production Stability	2.18	2.76	2.95	2.63	1.181	1.375	1.290	1.306
15	Quality Index (out of 100)	3.95	4.62	5.23	4.60	0.722	0.865	0.922	0.981
16	Return on Invested Capital	3.18	4.90	5.50	4.52	0.795	0.995	0.673	1.288
17	Revenue per Barrel	5.23	5.05	5.05	5.11	0.869	0.921	0.950	0.904
18	Safety (lost work incident rate)	5.09	4.38	4.64	4.71	0.750	0.805	1.049	0.914
19	Safety (total case incident rate)	5.00	4.29	4.73	4.68	0.816	0.902	0.935	0.920
20	S, G & A Cost per Barrel	4.36	4.57	5.09	4.68	1.002	1.207	1.065	1.120
21	Skills (inventory of cross-functional employee skills)	4.09	4.67	4.45	4.40	0.811	0.913	0.739	0.844
22	Throughput per Month (millions of barrels)	4.32	4.86	4.86	4.68	0.894	0.910	0.710	0.868
23	Training (hours per employee annually)	5.27	4.95	4.68	4.97	0.767	0.740	0.780	0.790
24	Warehouse Moves (actual shipments as a % of plan)	4.23	4.19	4.41	4.28	0.813	0.750	0.854	0.801
Overall Org. Performance		4.35	4.52	4.78	4.55	0.647	0.680	0.422	0.610
Average		4.07	4.42	4.59	4.36				
Avg. of nonfinancial measures		4.09	4.31	4.44	4.28				
Avg. of financial measures		3.94	4.82	5.15	4.63				

**Table 3-5 Pearson Correlation Coefficients – Performance versus Importance  
(n = 67)**

No.	1	2	3	4	5	6	7	8
1	0.176							
2		0.322**						
3			0.382**					
4				-0.074				
5					0.284*			
6						0.041		
7							0.339**	
8								0.264*

No.	9	10	11	12	13	14	15	16
9	-0.022							
10		0.208						
11			0.252*					
12				0.294*				
13					0.370**			
14						-0.105		
15							0.226	
16								0.389**

No.	17	18	19	20	21	22	23	24
17	0.409**							
18		0.008						
19			0.294*					
20				0.207				
21					0.317**			
22						0.287*		
23							0.428**	
24								0.245*

**Note:** Average absolute value of correlations equals 0.248.

\* Significant at the 0.05 level.

\*\* Significant at the 0.01 level.

**Table 3-6 Version 2 versus 3 – Overall Performance Assessment T-Test**

	Mean	Std. Dev.	n	Sig. (1-tailed)
VER2	0.52	0.512	21	
VER3	0.78	0.422	23	
Difference	-0.26			0.037 <sup>(a)</sup>

Note, a binary split was done on overall organizational performance (0 if ORGPREF <= 4; 1 if ORGPREF >=5).

<sup>(a)</sup> Nonparametric tests yielded similar results.



## **4 INTOLERANCE OF AMBIGUITY AND ORGANIZATIONAL PERFORMANCE MEASUREMENT**

### **4.1 Introduction**

Many firms are replacing traditional performance measurement systems, those that focus solely on financial performance, with systems that combine the use of financial and nonfinancial measures. One such system is the Balanced Scorecard (BSC), introduced by Kaplan and Norton in 1992. The BSC promotes a more “balanced” approach in measuring progress towards achievement of strategic objectives (Kaplan and Norton 2001a, 2001b, 2001c). A survey conducted by Renaissance Worldwide, Inc. estimated 60% of Fortune 1000 companies have implemented or are experimenting with the BSC (Silk 1998).

Managing performance with a group of measures like the BSC could be considered inherently more complex than managing with a single bottom-line financial measure and could, therefore, introduce potential cognitive difficulties. Recent literature on the BSC has explored some of the potential cognitive difficulties in the use of these types of performance measurement systems (e.g. Lipe and Salterio 2000, 2002; Ittner et al. 2003; Banker et al. 2004). Ittner et al. (2003, 725) states, “The evidence suggests that psychology-based explanations

may be equally or more relevant than economics-based explanations in explaining a firm's management practices.”

We are interested in individual differences among users of performance measurement systems and how these differences affect how much performance measurement information is perceived to be important. Cognitive psychology and the study of individual differences have long been important motivations for accounting studies. An aspect of individual differences that has been examined in accounting studies dating back to the 1970s is intolerance-tolerance of ambiguity. Early accounting studies attempted to examine, with limited amounts of success, the relationship between intolerance of ambiguity (IOA)<sup>11</sup> and the use of information (e.g. Dermer 1973; Oliver and Flamholtz 1978; McGhee et al. 1978). More recent studies in the 1980s and 1990s have continued to produce inconsistent results, in part due to poor measurement reliability and the use of different measures of IOA (Ho and Rodgers 1993). Unlike many prior studies which use multidimensional measures for IOA, we use unidimensional measures. We examine the relationship between IOA and the perceived amount of information contained in a set of organizational performance measures (financial and nonfinancial). No study to-date has examined this relationship.

The amount of important (i.e. useful) information available to evaluate organizational performance can affect both the perception of risk in the evaluation and the outcome of the evaluation. The negative correlation between information available and perceived risk has been established in the marketing arena. Studies have shown that less information leads to more perceived risk in a purchasing situation and visa versa (Crocker 1986; Kim and Lennon 2000). We hypothesize that the perceived amount of information contained in a set of organizational

---

<sup>11</sup> Intolerance-tolerance of ambiguity can be measured on a continuum with intolerance on one extreme and tolerance on the other extreme. This paper uses the term IOA to refer to this measurement continuum.

performance measures is negatively correlated with the perceived risk in the evaluation of organizational performance. The theory of cognitive consistency suggests that evaluations will be more positive when perceived risk is low and visa versa (Alhakami and Slovic 1994). Therefore, we also hypothesize that the perceived risk in the evaluation of organizational performance is inversely related to the actual evaluation of organizational performance.

Differences in how data is formatted can have an impact on a decision maker. Lipe and Salterio (2002) found that the presentation format of performance measures (the BSC format versus an unformatted listing) could have an effect on the evaluation of organizational performance. In addition to the research questions and hypotheses previously mentioned, we attempt to determine if the presentation format of performance measures has any impact on the perceived importance of those measures.

The remainder of this paper is organized as follows. The next section reviews pertinent literature on IOA, information usage, perceived risk, perceived benefit, and the BSC. The following section develops measures for IOA. The next section presents the research questions and hypotheses. Next the methodology is discussed. The results are next, followed by a discussion of the results. The final section discusses ramifications of this paper.

## **4.2 Literature Review**

### *4.2.1 Measures of Intolerance of Ambiguity*

An interest in IOA has been present in the psychology literature for more than 50 years.<sup>12</sup> IOA is considered to be the tendency to see ambiguous situations

---

<sup>12</sup> The first person to examine IOA in a comprehensive fashion was Else Frenkel-Brunswick (1949).

as threatening (Budner 1962). Budner (1962, 30) identifies three types of ambiguous situations:

a completely new situation in which there are no familiar cues; a complex situation in which there are a great number of cues to be taken into account; and a contradictory situation in which different elements or cues suggest different structures—in short, situations characterized by novelty, complexity, or insolubility.

Budner's self-report questionnaire contains 16 items, scored with a Likert-type scale with half being scored positively and half scored negatively. The Budner scale has been used more frequently than others in IOA research (Furnham 1994; Furnham and Ribchester 1995). Budner validated his scale on 17 different populations, many of which were student populations. It was shown to be free of acquiescence and social desirability response bias and to have content, concurrent, and construct validity. However, it has been criticized for poor internal reliability. Budner (1962) reported his scale had a mean Cronbach alpha of 0.49 over all samples. Furnham's (1994) analysis of four IOA scales reported an alpha of 0.59 for the Budner scale.

Another scale for IOA is the MAT-50 (Norton 1975). The MAT-50 questionnaire effectively deals with the reliability problems associated with the Budner scale but contains enough items to hinder its use in behavioral research. Furnham (1994) reported an alpha of 0.89 for the MAT-50 scale. Some other IOA scales are Walk's A Scale (O'Conner 1952) and the AT-20 Scale (Rydell and Rosen 1966; MacDonald 1970). The Walk scale contains 8 items but suffers from poor internal reliability (Ehrlich 1965; Furnham 1994). The AT-20 scale contained 16 true-false questions when first introduced by Rydell and Rosen (1966). MacDonald (1970) added another four questions. Furnham (1994) reported an alpha of 0.78 for the AT-20 Scale.

Furnham (1994) conducted factor analysis on the Walk, Budner, and AT-20 scales. Each of the scales was shown to be multidimensional. The Walk scale was shown to contain three independent dimensions, the Budner scale contained four dimensions, and the AT-20 scale contained six dimensions. Furnham (1994) also provided correlation coefficients between the different instruments. See Table 4-1 for these correlations. Table 4-1 illustrates that the instruments measure different constructs. In particular, the Budner scale is not highly correlated with the others.

#### *4.2.2 Studies with Intolerance of Ambiguity as an Independent Variable*

Including cognitive characteristics is a trend in behavioral accounting research (Ho and Rodgers 1993). Dermer (1973, 511) states, "Of particular interest to accountants is the possibility that the cognitive characteristics of an information user may affect his perception of what information is important and, hence, may affect how information influences his ultimate behavior." One early stream of research examined the relationship of decision style with information usage (e.g. Driver and Mock 1975; San Miguel 1976; and Savich 1977). The Budner (1962) studies support a possibility of a relationship between IOA and the use of information. Theory suggests individuals intolerant of ambiguity will be more likely to seek additional information in order to reduce the threat from existing ambiguous data or an overall lack of data. Several studies have been performed over the past 30 years that examined IOA. The Budner scale, the AT-20 scale, and the MAT-50 scale have been used extensively in the accounting and organizational psychology literature. Results of many of these studies are presented in an abbreviated fashion in Table 4-2 and summarized below.

Dermer (1973) examined IOA using the Budner scale. He found IOA was significantly correlated with perceived importance of information. Dermer (1973) used sales supervisors, district sales managers, and regional sales managers to sort informational aspects of their jobs in order of importance, using a Q-sort

technique. He then asked participants to quantify the number of pieces of information that were of little importance. Dermer (1973) reported the correlation between IOA score and number of unimportant items was negative 0.33 ( $p < 0.05$ ).

McGhee et al. (1978) examined the actions of MBA students in a quasi-experiment. The students were asked to play the role of investment advisors and recommend (or not) various companies to be included in an investment portfolio. The subjects received eight cues, all containing financial information. Budner's scale was used to classify subjects into three groups based upon IOA. The ANOVA results failed to detect differences among the IOA groups in additional information wanted and decision confidence.

Oliver and Flamholtz (1978) studied the effect of human resource replacement cost information on confidence in a layoff decision. Subjects intolerant of ambiguity, as determined using Budner's scale, gained confidence in their layoff decision from the replacement cost information more often than their more tolerant counterparts.

Faircloth and Ricchiute (1981) used the MAT-50 scale developed by Norton (1975) to study desire for financial reporting alternatives and desire for disclosure among practicing CPAs. ANOVA tests failed to detect significant differences among the IOA categories (low, low-medium, high-medium, and high).

Like Oliver and Flamholtz (1978), Gul (1984) studied the effect of human resource replacement cost information on confidence in a layoff decision. Unlike Oliver and Flamholtz (1978), Gul (1984) used the AT-20 scale to measure IOA. The study used male managers from Malaysian electronics companies. Gul (1984) failed to find significant main effects for IOA but reported a significant

interaction between IOA and field dependence; as well as, significant main effects for field dependence.

Ashford and Cummings (1985) examined feedback-seeking behavior in a job setting using 7 items from the MAT-50 scale, three items for job-related ambiguity and four items related to problem-solving. Factor analysis verified that these items contained the two expected dimensions. In their examination of feedback-seeking behavior that was a follow-up to Ashford and Cummings (1985) using the same dataset, Bennett et al. (1990) found that a general intolerance for ambiguity did not motivate the seeking of job-specific feedback but a job-related intolerance was significant.

Gul (1986) examined confidence in loan decisions in an experimental setting using male Australian bankers. The study included a between-subjects manipulation of the audit opinion (qualified versus unqualified). The AT-20 scale was used to measure IOA. A significant interaction was found between audit opinion and IOA. Only for qualified opinions was there a significant difference in mean confidence scores between managers low on tolerance (more intolerant) versus high on tolerance (less intolerant) of ambiguity. Managers low on tolerance were less confident when faced with a qualified opinion presumably because there was no way to seek additional information to raise their confidence level.

Tsui (1993) used a sample of New Zealand bankers to examine IOA (AT-20 scale) and interest rates for a hypothetical loan decision. Interest rates represent a loan officer's assessment of risk. The case involved a "subject-to" audit opinion for a material uncertainty. Subjects low in tolerance for ambiguity (highly intolerant) established a significantly higher interest rate for the loan.

Ghosh and Ray (1997) examined attitude towards risk and IOA (using the Budner scale) and their effects on decision confidence in an experimental setting using MBA students. The case involved establishing the number of quality control

inspections for a series of plants, a type of risk-reducing activity. Subjects that were intolerant of ambiguity required more inspections than those who were more tolerant, but they had less confidence in their decisions.

Lal and Hassel (1998) used the Budner IOA questionnaire, a measure for perceived environmental uncertainty (PEU), and a term for the interaction of the two in a regression equation that predicted the perceived usefulness of non-conventional management accounting system (MAS) information.<sup>13</sup> A random sample of managing directors of New Zealand manufacturing companies was used to conduct the study. The correlation coefficient between the perceived usefulness and IOA was 0.20 but not statistically significant. However, the interaction term in the regression equation was both positive and significant.

Wright and Davidson (2000) found an association between IOA using the AT-20 scale and assessed loan risk in commercial lending decisions. They used Canadian bank loan officers in their experiment. Wright and Davidson's (2000) results contradict those of Tsui (1993) where subjects more intolerant of ambiguity established higher interest rates for loans than their more tolerant counterparts. In Wright and Davidson (2000), subjects less tolerant (more intolerant) of ambiguity tended to view the loan as less risky and tended to charge a lower interest rate.<sup>14</sup> The authors attribute these contradictory findings to the proposed uncertainty in which persons highly intolerant of ambiguity react to ambiguous situations.

---

<sup>13</sup> Non-conventional MAS information was defined by a questionnaire developed by Chenhall and Morris (1986). The regression equation used by Lal and Hassel (1998) was  $Y = \alpha_0 + \beta_1 PEU + \beta_2 IOA + \beta_3 PEU * IOA + \varepsilon$ , where, Y = perceived usefulness of non-conventional MAS information.

<sup>14</sup> Wright and Davidson (2000) also examined levels of auditor attestation. In the one cell of no attestation, subjects less tolerant of ambiguity charged a higher interest but saw the loan as less risky.



#### 4.2.3 *Information Usage and Perceived Risk*

In the marketing context, “perceived risk refers to the nature and amount of risk perceived by a consumer in contemplating a particular purchase decision” (Cox and Rich 1964, 33). A central theme in the previous marketing literature on perceived risk has been the theory that it is positively related to information search. When faced with a perceived risk, individuals theoretically will engage in risk-reducing behaviors like searching for additional information (Cox and Rich 1964; Gemünden 1985; Murray 1991). In a survey of telephone shoppers, “The most favored strategy of reducing uncertainty among the telephone shoppers surveyed was to seek information, usually newspaper advertising” (Cox and Rich 1964, 38). Gemünden’s (1985) meta-analysis identifies 100 studies in which the association of perceived risk and information search are examined. About half of the studies either confirm a positive association or show a tendency towards a confirmation. Although the prior research findings are not consistent, Gemünden (1985, 89) states, “Our findings offer some tentative support for the validity of the risk-information-search hypothesis.”<sup>15</sup>

The preceding paragraph discusses information search as a potential risk reducing activity. Alternatively, when the amount of information presented to the consumer is manipulated, it is the perceived risk that is affected. Crocker (1986) found that more information negatively affected consumers of legal services perception of risk. Kim and Lennon (2000) found that even the perceived amount of information (actual amount of information was constant among the subjects) negatively affected television shoppers’ perceived risk.

---

<sup>15</sup> Some later studies provide additional evidence. As an example, Mitra et al. (1999) found evidence that high risk credence services are associated with greater information search.

#### 4.2.4 *Perceived Risk versus Perceived Benefit*

Alhakami and Slovic (1994) examined the correlations between perceived risk and perceived benefit of 40 items and found evidence supporting the theory of cognitive consistency. The theory of cognitive consistency suggests individuals who view an item as beneficial also view that item as being low in risk (Alhakami and Slovic 1994 citing Heider 1946 and McGuire 1968). Alhakami and Slovic (1994) asked subjects to assess the risk and benefit of 40 items like water fluoridation, herbicides, smoking, handguns, etc. Negative correlations were observed for 38 of the 40 items and 26 of the negative correlations were significant ( $p < .01$ ). “The results confirmed the existence of strong inverse interdependence between risk and benefit judgments” (Alhakami and Slovic 1994, 1094).

Alhakami and Slovic (1994) studied perceived risk and perceived benefit in a general context. In a marketing context, purchase intent is a proxy for perceived benefit. In their study on telephone shopping, Cox and Rich (1964, 38) write, “...it would seem reasonable to conclude that high perceived risk is likely to be a strong deterrent to purchasing an item by telephone.” Marketing studies have often examined the link between information available and both perceived risk and purchase intent, but have failed to examine directly the link between perceived risk and purchase intent (e.g. Crocker 1986; Kim and Lennon 2000). Crocker (1986) found a significant negative relationship between amount of information and perceived risk and a significant positive relationship between amount of information and intent to retain legal services. Kim and Lennon (2000) examined information available, perceived risk, and purchase intent in a setting of television shopping for apparel. Participants in their study were exposed to the same television segment advocating an apparel purchase. Since participants were exposed to the same television segment, there were no differences in the amount of actual information presented; however, there were differences among the participants in their perceptions of information content. The more information, as

perceived by the participants in the television segment, the less risky the apparel purchase and the more likely the participant was to make the purchase. Crocker (1986) and Kim and Lennon (2000) provide evidence supporting cognitive consistency even if they do not directly examine the inverse relationship between risk and purchase intent (a proxy for perceived benefit).

#### 4.2.5 *Balanced Scorecard*

The Balanced Scorecard (BSC) is a mechanism for translating the organization's strategy into operational terms (Kaplan and Norton 2001a, 2001b, 2001c). Its implementation requires management to (1) develop coherent strategies to achieve their objectives and (2) develop a set of measures for gauging the organization's performance in implementing those strategic objectives. The BSC is typically 20 – 25 key measures comparing actual performance to a budget or target, organized into specific categories, on a one-page document. Kaplan & Norton (2001a, 90) outline the four recommended categories or perspectives:

1. *Financial – the strategy for growth, profitability, and risk viewed from the perspective of the shareholder.*
2. *Customer – the strategy for creating value and differentiation from the perspective of the customer.*
3. *Internal Business Processes – the strategic priorities for various business processes that create customer and shareholder satisfaction.*
4. *Learning and Growth – the priorities to create a climate that supports organizational change, innovation, and growth.*

Lipe and Salterio (2000) was the first study to examine cognition in the use of the BSC. In their experiment, MBA students were asked to evaluate the performance of two division managers of a clothing firm based on a review of each division's BSC. The case materials documented that the two divisions were pursuing different strategies, thus allowing for the scorecards to contain some measures that were unique to each division. Lipe and Salterio (2000) found measures unique to each division were not utilized in the evaluations of division

managers. Banker et al. (2004) confirmed the Lipe and Salterio (2000) findings and examined further the impact of strategy linkage on the use of performance measures.<sup>16</sup>

Lipe and Salterio (2002) used the same case in conducting an experiment to determine the effect of format (BSC format versus unformatted) on evaluations of two division managers. When differences between divisions were scattered across the four BSC categories, Lipe and Salterio (2002) did not find a significant difference between the formats. However, they found BSC format moderated the evaluations when differences between divisions were concentrated into a single BSC category.

### **4.3 Development of Needed Intolerance of Ambiguity Measures**

#### *4.3.1 Pilot Study 1*

Ho and Rodgers (1993) discuss the differing results found in the prior IOA literature. One possible reason for the differing results mentioned by Ho and Rodgers (1993) was usage of different measurement instruments. Three IOA scales have been used in the prior accounting literature: Budner's scale, the MAT-50 scale, and the AT-20 scale. The Budner scale is the most widely used in IOA research (Furnham 1994; Furnham and Ribchester 1995). Looking at Table 4-2, one can see that the Budner scale was used in many of the prior accounting studies that successfully found evidence of a link between IOA and the dependent variable of interest. Five of the studies in Table 4-2 used the Budner scale for IOA; three of the studies found IOA was significant. In addition, the Lal and Hassel (1998) study found the IOA's correlation with the perceived usefulness of information was 0.20 but they did not have sufficient sample size for the correlation to be significant. The prior success of a particular scale is a motivating

---

<sup>16</sup> Interested readers may also consult DeBusk et al. (2003). They examine the effects of strategy on the usage of performance measures and the number of perspectives.

factor in selecting that scale for a new experiment. We, therefore, selected the Budner scale for the first pilot study.

Another possible reason for the inconsistent results in the prior literature is poor measurement reliability (Ho and Rodgers 1993). The Budner scale, although widely used, suffers from poor measurement reliability (Budner 1962; Norton 1975; Furnham 1994). Factor analysis can be used to reduce a multidimensional scale into its unidimensional constructs. One or more of these individual constructs may have more reliability than the overall scale. We, therefore, planned to use factor analysis as a tool to develop a reliable instrument for IOA.

The Budner questionnaire (see Appendix D) was given to students in an introductory accounting class at a major university in the United States. Since all undergraduate business students were required to take this course, the sample reflects a broad spectrum of business majors. A total of 764 usable responses were obtained from the survey. Cronbach alpha for the entire instrument is 0.44, well below acceptable levels. The sample was subjected to factor analysis with a varimax rotation. Only the first factor emerged with more than two variables having loadings greater than 0.4. See Table 4-3. The first factor is very similar to the factor labeled “predictability” in Furnham’s (1994) examination of IOA scales. Question 2 loaded primarily on the first factor; however, its loading of 0.23 was weak enough that question 2 was eliminated from further consideration. Questions 9, 10, 8, 1, and 13 were combined into a single construct called predictability. Cronbach alpha for the predictability construct is 0.69.

#### *4.3.2 Pilot Study 2*

Pilot study 2 further refined the construct that emerged from pilot study 1 by utilizing information from the Furnham (1994) study and the Bennett et al. (1990) study. Furnham (1994) conducted a second-order factor analysis of the 21 factors determined from analyzing separate IOA scales. The predictability factor

from the Budner scale loaded on the same dimension as the completer/finisher factor from the AT-20 scale and the conservatism factor from the Walk scale. The individual items from these correlated factors were included in the second pilot study. Questions from the job-related segment of the MAT-50 scale were also included, consistent with Bennett et al. (1990). The questionnaire contained fourteen items, ten from the predictability and related factors and four from the job-related section of the MAT-50 scale. The ordering of the items was randomized. See Appendix E and Table 4-4. The objective of the pilot study 2 was to use factor analysis to reduce these fourteen items into one or two factors (predictability and job-related IOA).

Subjects were obtained for the second pilot study via an email request of individuals working in an industrial / research park located in the eastern United States. A total of 36 people participated in the web-based survey. Multiple iterations of factor analysis were conducted in order to distill the items down into two factors by deleting variables with insufficient loadings or significant loadings on multiple factors. Five variables were deleted (Appendix E – Questions 2, 4, 8, 9, and 10). Results of the final factor analysis with varimax rotation are presented in Table 4-5. Two factors emerged: predictability and job-related IOA. One variable, Appendix E – question 6, loaded heavily on both constructs. However, the heaviest loading for this variable was on the anticipated construct: job-related. One variable, Appendix E – question 1, loaded more heavily on the job-related factor than on the anticipated predictability factor. Cronbach alphas of 0.73 for the predictability IOA factor and 0.74 for the job-related IOA factor are within the acceptable range (Nunnally 1978). The nine questions that make up these two constructs were used in this paper's main study.

## **4.4 Research Question and Hypothesis Development**

### *4.4.1 Importance of Information*

Theory suggests that persons intolerant of ambiguity, when faced with an ambiguous situation, will pursue risk-reducing activities like the seeking of additional information. If the seeking of additional information is not a viable option, these persons are likely to perceive the situation as more risky than their more ambiguity tolerant counterparts. Dermer (1973) found that the more a person was intolerant of ambiguity, the more job-related information he/she deemed to be of importance. Similarly, Oliver and Flamholtz (1978) found intolerant subjects gained confidence from additional human resource information. Some studies like McGhee et al. (1978) have failed to find support, however, for a link between IOA and the seeking of additional information.

We examine, in this paper, IOA and its relationship to the importance of information (i.e. the percentage of performance measures perceived to be important) in assessing overall organizational performance. This can be expressed in the form of the following research questions. Research questions are deemed more appropriate than hypotheses due to the inconclusive nature of prior research.

Q1: Is IOA correlated with the importance of information in the evaluation of organizational performance?

Q1a: Is IOA, measured using the predictability construct, correlated with the importance of information (i.e. the percentage of performance measures perceived to be important) in the evaluation of organizational performance?

Q1b: Is IOA, measured using the job-related construct, correlated with the importance of information (i.e. the percentage of performance measures perceived to be important) in the evaluation of organizational performance?

See Figure 4-1 for a view of the proposed model.

#### *4.4.2 Information Use, Perceived Risk and Assessment of Performance*

Several prior studies in marketing have examined and documented a positive relationship between perceived risk and information search (Cox and Rich 1964; Gemünden 1985; Murray 1991). Other marketing studies provide evidence of a negative relationship between both the perceived and actual amounts of information available and perceived risk (Crocker 1986; Kim and Lennon 2000). We build on the results found in marketing by examining the relationship between the percentage of performance measures perceived to be important and an individual's perceived risk in assessing organizational performance. The available research in the marketing arena suggests the following research hypothesis.

- H1: The amount of available information perceived to be important (i.e. the percentage of performance measures perceived to be important) is inversely associated with perceived risk.

Alhakami and Slovic (1994) found evidence of a negative correlation between perceived risk and perceived benefit, supporting the theory of cognitive consistency. Cox and Rich (1964) found a negative relationship between perceived risk and purchase behavior in telephone shopping, further supporting cognitive consistency theory. Other studies have documented the inverse relationship between information available and perceived risk and the positive relationship between information available and purchase intent without examining directly the relationship between perceived risk and purchase intent (Crocker 1986; Kim and Lennon 2000). We build on the psychology and marketing literature by examining the relationship between perceived risk and the actual assessment of organizational performance. The following hypothesis is based on the above information.



- H2: The perceived risk in an evaluation of organizational performance is inversely associated with the assessment of organizational performance.

See Figure 4-1 for a view of the proposed model.

#### 4.4.3 *Presentation Format*

Lipe and Salterio (2002) found evidence that presentation format of performance measures may impact the perception of performance. Subjects in their study assessed the performance of two division managers. When the differences between the two divisions were placed solely into one BSC category, the BSC format moderated the evaluations. There was no difference found when differences between the two divisions were scattered among the four BSC categories. Lipe and Salterio (2002) only examined one organization and the actual performance is close to target on all measurements making it unlikely that differences in overall organizational performance would be detected, where some subjects receive BSC formatted data and others receive unformatted data. However, the Lipe and Salterio (2002) study did not report data on the perceived importance of performance measures and it remains unresolved if the presentation format of performance measures may have an impact on the perceived importance of those measures. We examine two formats: (1) the BSC format and (2) a proxy for unformatted data where measures are arranged in alphabetical order. See the research question below.

- Q2: Does format (BSC versus unformatted) affect the importance of information (i.e. the percentage of performance measures perceived to be important) in the evaluation of organizational performance?

## **4.5 Methodology**

### *4.5.1 Instrument and Measures*

Two instruments were developed in order to conduct a survey of executives. The only difference between the two instruments was the format of the performance measures (BSC format or unformatted). The BSC formatted version is provided in Appendix F. The instrument provides performance measures for a fictional business unit.

In the first part of the instrument, subjects were asked to assess the business-unit performance and to assess the importance of each measure in assessing business-unit performance. The overall business-unit performance was reflected in the subject's response to question 1 which asked the subject to evaluate performance on a Likert-type scale from 1 to 10. The importance on each of the 20 performance measures was also assessed on a Likert-type scale from 1 to 10. See Appendix F, Part I: questions 2 – 21. The variable utilized to represent the percentage of performance measures perceived to be important was calculated by summing, for each subject, the number of measures with an importance score of 6 through 10 and dividing by the total number of performance measures.

After answering questions on performance measures, the subjects were asked five questions designed to measure the perceived risk in the assessment of organizational performance. See Appendix F – Part II. The risk variables were assessed on a Likert-type scale containing six possible responses. The scoring of these questions was done on a 7 point scale where a score of 4 was assigned when the question was omitted.<sup>17</sup> A high score means the subject perceives a high amount of risk.

---

<sup>17</sup> Reported results are qualitatively the same as results excluding the eight cases where subjects failed to answer either a Risk question or an IOA question.

Part III of the instrument contains the nine questions for measuring IOA and one demographic question. See Appendix F, questions 28 – 36. Question 27 provides data on the demographic variable of experience. The IOA questions were scored the same way as the risk questions, which was the same way Budner (1962) scored his scale. A high score represents a high amount of intolerance of ambiguity.

#### *4.5.2 Sample*

The sample consisted of members of the American Management Association (AMA) likely to assess business-unit performance as part of their job. These individuals were in the AMA classification of top executives / vice-president. The file obtained from the AMA contained SIC (standard industrial classification) codes for a portion of the firms. Where possible, government agencies, financial services and professional service firms were excluded from the mailing (4-digit SIC codes 6000 – 9999). A total of 2,714 surveys were mailed, 49 were returned by the postal service as non-deliverable, leaving a net total of 2,665 surveys mailed. After two weeks, a follow-up post card was sent to approximately 83% of the potential respondents.<sup>18</sup> A total of 430 responses were received but seventeen were excluded because of multiple instances of missing data or multiple answers to questions requiring a single response. A net total of 413 (430 – 17) usable surveys were received for a 15.5% response rate.

Of the 413 usable responses, 166 failed to answer question 1, “On a scale of 1 – 10 where 1 is very poor and 10 is very good, please circle the number that indicates your opinion of Division A’s overall 2003 performance.” The omission probably occurred because question 1 was a single question placed between two sets of instructions and thus easily overlooked. See Appendix F, Part I. A

---

<sup>18</sup> The post cards were sent using the United States Postal Service’s online mailing service. Due to a problem at the mailing post office, approximately 17% of the follow-up post cards could not be mailed.

MANOVA (multivariate analysis of variance) analysis was conducted to compare those that answered question 1 with those that failed to answer question 1 on the remaining questions. The results are not significant ( $p = .583$ ) indicating a failure to reject the hypothesis that the two groups were not equal. Observed power levels of 0.877 using alpha equal to 0.05 are above the recommended 0.80 threshold for concluding that there were no meaningful differences between the groups (Hair et al. 1998). Since MANOVA detected no significant differences, the 413 usable responses were split into two separate samples on the basis of whether or not they answered question 1. The group of 166 subjects who failed to answer question 1 (sample 1) were used for additional factor analysis on the two IOA constructs and the perceived risk construct and to provide preliminary evidence relating to regression weights.

The remaining 247 subjects (sample 2) were used to test the final model. Of the 247 subjects in sample 2, 95.5% had at least 10 years of experience, at least 20 years – 57.1%, and at least 30 years – 17.4%. Sample 2 was examined for a potential non-response bias. Early-responders were compared to late-responders on all variables using MANOVA. There are no significant differences between early and late-responders ( $p = .521$ ) with observed power of 0.878 (alpha = .05).

#### *4.5.3 Structural Equation Modeling*

Structural equation modeling (SEM) was utilized to examine the final model in a two-step process. It was first utilized in a fashion that is similar to confirmatory factor analysis, in order to evaluate the factor loadings and to assess the overall fit of the measurement model (a model containing only the constructs and their indicators). The second step was to use SEM to evaluate the strength, direction (positive or negative) and significance of the relationships proposed in the final structural model (the measurement model plus the regression paths) while simultaneously assessing the overall model fit. It is recommended that several fit indices be examined when evaluating model fit in SEM (Hair et al.

1998; Dilalla 2000). This paper reports four indices: the chi-square statistic ( $\chi^2$ ), the adjusted gross fit index (AGFI), the root mean square error of approximation (RMSEA), and the incremental fit index (IFI). It is recommended that the p-value for the  $\chi^2$  statistic be greater than 0.05, that the AGFI and IFI indices be greater than 0.90, and that the RMSEA be less than 0.08 (Dilalla 2000).

## **4.6 Results**

### *4.6.1 Sample 1*

Two IOA factors emerged from pilot study 2, although with a small sample ( $n = 36$ ), further analysis would be beneficial. Factor analysis was run on the nine variables that made up the two IOA constructs (predictability and job-related). Two of the weakest variables from pilot study 2 were deleted due to insufficient loadings, Appendix F – questions 28 and 30. The remaining seven variables loaded on a single IOA construct. See Table 4-6. Cronbach alpha for the seven variables was 0.71. The loading of the variables on a single factor was less surprising when two items were considered: (1) the cross-loadings in pilot study 2 and (2) the variable in pilot study 2 that loaded more heavily on the factor other than what was anticipated a priori. The single IOA construct was used with SEM for testing with sample 2. The collapsing of IOA into a single unidimensional construct makes it impossible to examine research questions 1a and 1b, dealing separately with the predictability and job-related constructs. However, the main question was examined with sample 2: (Q1) Is IOA correlated with the importance of information in the evaluation of organizational performance?

Factor analysis was also run on the five variables that were anticipated to make up the perceived risk construct. One variable was deleted due to an insufficient loading, Appendix F – question 26. The remaining four variables loaded on a single construct with a Cronbach alpha of 0.77. See Table 4-7. These four variables were used for testing with sample 2.

SEM was used to test some of the regression weights in the proposed model in Figure 4-1.<sup>19</sup> The path representing Q2, between format and the percentage of performance measures perceived to be important, was not significant ( $p = .372$ ).<sup>20</sup> In this study, format did not affect the importance of information. Format was, therefore, deleted from the final model.

The path between IOA and the percentage of performance measures perceived to be important (Pct. Important), Q1, was also not significant ( $p = .780$ ) but the path between importance and risk, H1, was significant ( $p = .021$ ). The achievement of a good fitting SEM model with the Pct. Important variable included would be problematic. However, if the Pct. Important variable is eliminated and the path is drawn between IOA and risk, the regression weight was negative and significant ( $p = .010$ ). The Pct. Important variable was deleted from the final model. Q1 and H1 are examined in a supplemental analysis using sample 2 data.

The proposed model in Figure 4-1 was modified to reflect the results from sample 1. The revised model is presented in Figure 4-2.

#### 4.6.2 *Sample 2*

Descriptive statistics and correlation coefficients are provided in Table 4-8 and Table 4-9. The first step before proceeding to the testing of the structural model represented in Figure 4-2 is to evaluate the measurement model. All factor loadings in the measurement model are significant and the construct's reliability is acceptable. The model fit is good. The  $\chi^2$  is 65.25 ( $df = 52$ ;  $p = .10$ ), AGFI is

---

<sup>19</sup> Sample 1 contains no data on the evaluation of performance so the path between risk and the evaluation of performance was not examined. IOA and risk were measured using the single constructs distilled earlier from factor analysis.

<sup>20</sup> In addition, a t-test was used to compare perceived risk (calculated as a summated scale of the four risk variables) between the two formats, BSC and unformatted. Risk was not statistically different between formats.

0.937, IFI is 0.974, and RMSEA is 0.032. It is therefore appropriate to proceed to the examination of the structural model.

The model shown in Figure 4-2 fits the sample 2 data well. The  $\chi^2$  is 65.53 ( $df = 53$ ;  $p = .12$ ), AGFI is 0.938, IFI is 0.975, and RMSEA is 0.031.<sup>21</sup> Some of the research questions / hypotheses are represented by the regression paths in Figure 4-2 and can be tested by examining the regression weights. The regression weights are provided in Table 4-10 and Figure 4-3. Some evidence on Q1 and H1 can be obtained by examining the regression weight of the IOA → Risk path. The standardized regression coefficient is negative and significant ( $\beta = -0.34$ ;  $p = .000$ ) providing evidence that the more ambiguity intolerant the subject, the less risk the subject perceives. The IOA → Risk path is examined further in the supplemental analysis below.

H2 can be tested by examining the regression weight of the Risk → Performance path. The standardized regression coefficient is negative and significant ( $\beta = -0.40$ ;  $p = .000$ ) supporting the hypothesis that perceived risk is negatively correlated with the evaluation of firm performance.

#### 4.6.3 Sample 2 - Supplemental Analysis

The data in Table 4-10 provides strong evidence that the more a subject is intolerant of ambiguity, the less risk is perceived in evaluation of the organization's performance. Theory suggests that those intolerant of ambiguity perceive ambiguous situations, like evaluating organizational performance, as threatening and therefore containing more risk. The results in Table 4-10 suggest that intolerant subjects are engaging in some risk reducing activity; otherwise,

---

<sup>21</sup> The data is not normally distributed due in part to the use of categorical measures. The reported fit indices and regression weights were calculated using maximum likelihood estimation which assumes multivariate normality. The asymptotically distribution-free (ADF) estimation is insensitive to departures from normality (Hair et al. 1998). SEM was run using ADF estimation with similar results.

they would perceive more risk than those more tolerant of ambiguity. The presence of a mediating variable like importance of information is suggested by these counter-intuitive results; however, this variable was deleted in order to obtain a good fitting SEM model. The results of sample 1 are mixed in regards to the importance of information variable, Pct. Important. The significant and negative regression coefficient for the IOA → Risk path; however, suggests that additional examination is warranted into Q1 and H1.

One type of available supplemental analysis would be to discard the “middle” group on the Pct. Important variable and examine the high and low groups on the possibility that the subjects in the middle could be obscuring the expected relationships. A summarized frequency distribution is presented in Table 4-11. The group of 62 subjects between 85% and 100% was discarded. The data on the remaining 185 subjects was examined using SEM with maximum likelihood estimation and the model shown in Figure 4-4. The model fit the data well except for the  $\chi^2$  statistic's p-value. The  $\chi^2$  statistic is extremely sensitive with larger sample sizes and tends to reject both good and bad fitting models (Dilalla 2000). The  $\chi^2$  is 96.58 ( $df = 64$ ;  $p = .005$ ), AGFI is 0.900, IFI is 0.921, and RMSEA is 0.053. Table 4-12 and Figure 4-5 contain the standardized regression weights. Regression coefficients are significant in the hypothesized direction for both H1 ( $\beta = -0.16$ ;  $p = .049$ ) and H2 ( $\beta = -0.42$ ;  $p = .000$ ). The regression coefficient for Q1 is positive but not significant ( $\beta = 0.14$ ;  $p = .128$ ). The positive sign of the regression coefficient with the marginal two-tailed p-value provides some evidence that many subjects behaved as IOA theory would suggest: ambiguity intolerant subjects sought additional information.<sup>22</sup>

Descriptive statistics for IOA, Pct. Important, Risk, and Performance are presented in Table 4-13. Summated scales were created for IOA and Risk in order to perform the additional supplemental analyses outside of SEM. Table 4-14

---

<sup>22</sup> SEM run using ADF analysis yield significant regression weights for all relationships.



presents the Pearson and Spearman correlations for IOA, Pct. Important, Risk, and Performance. Q1 relates to the correlations between IOA and Pct. Important. H1 relates to the correlations between Pct. Important and Risk. While the Pearson correlations are less than impressive in Table 4-14, the Spearman correlations are significant for both correlations. The presence of significant Spearman correlations suggests a possible non-linear relationship.

The sample 2 data was split on the median score of 23 on the IOA summated scale. The correlations between IOA and Pct. Important after the median split are presented in Table 4-15. Panel A represents subjects below the median score on IOA. Neither the Pearson or Spearman correlations are significant ( $p < .05$ ) in Panel A but the sign is negative. The negative sign would tend to indicate that within the group of subjects tolerant of ambiguity (i.e. below the median score); the more intolerant they are the more they are comfortable making evaluations of organizational performance with less and less information. Panel B of Table 4-15 represents subjects intolerant of ambiguity (i.e. above the median score). Here the relationship between IOA and perceived importance of information is positive and significant. The Pearson and Spearman correlations are 0.235 and 0.320 respectively ( $p < .015$ ). These results suggest that within the group of subjects intolerant of ambiguity, the more they are intolerant the more information they consider important. These findings are consistent with a risk-reducing behavior.

Preliminary results from sample 1 support H1 which predicts that the relationship between the perceived importance of information and perceived risk is positive. Sample 2 data was split into two groups based on the median score of 95% on the Pct. Important variable. The only scores reported above the 95% median are those subjects who reported that all or 100% of the performance measures are important. A t-test is presented in Table 4-16 comparing the mean Risk scores between the two groups. The mean Risk score for those considering

fewer of the performance measures important is 15.56, while the mean Risk score for those considering more of the performance measures important is 14.21. Members of the high importance group perceive 8.7% less risk than the low importance group. The difference of 1.35 is significant ( $t = 2.238, p = .026$ ).<sup>23</sup> These results suggest that the more information one considers important in evaluating organizational performance, the less risk they will perceive exists in that evaluation.

## **4.7 Discussion**

### *4.7.1 Research Question 1*

Much of the success in previous research (see Table 4-2) was achieved using the Budner scale. Bennett et al. (1990) also found that job-related IOA was significantly related to feedback seeking behavior. This paper builds on this prior research by utilizing the primary factor from the Budner (1962) scale and the job-related factor from Bennett et al. (1990). An IOA scale was developed through two pilot studies and was refined through use of a split sample in the main study. The new IOA instrument's reliability was acceptable and validity was provided by sampling from a population of top executives.

SEM results fail to detect a significant relationship between IOA and the percentage of performance measures perceived to be important in the evaluation of organizational performance. Supplemental analysis, however, provides evidence that persons low on IOA behave differently than those high on IOA. Results presented in Table 4-15 show that those subjects below the median score on IOA have a negative correlation coefficient for IOA and Pct. Important; while

---

<sup>23</sup> The underlying data may violate the normality assumption. However, the t-test is generally considered robust when samples sizes are equivalent and variances are equal. Levene's Test for Equality of Variances failed to reject the null hypothesis ( $F = 1.372, p = .243$ ). Nonparametric tests also produced results similar to the t-test.

those above the median score on IOA have a positive correlation coefficient. The results suggest that individuals over a certain baseline score on IOA behave as IOA theory predicts. The more they are intolerant of ambiguity, the more information they seek in order to reduce the perceived threat and make a better decision. However, some evidence is provided that the more tolerant individuals seem to behave in a way counter to those more intolerant of ambiguity. These results may help explain why there are inconsistent results in the prior literature. Prior studies have utilized samples from different populations with perhaps differing mean levels of IOA. Most of the prior research has also assumed that the relationship between IOA and the dependent variable was linear. The data in Table 4-15 does not support that assumption.

#### *4.7.2 Hypothesis 1*

Kim and Lennon (2000) found that even the perceived amount of information provided to consumers was negatively related to their perception of risk in the purchase decision. The results presented in Table 4-16 are consistent with Kim and Lennon (2000) and support H1. Subjects above the median in percentage of performance measures perceived to be important perceive less risk in the evaluation of organizational performance than those below the median.

#### *4.7.3 Hypothesis 2*

The theory of cognitive consistency suggests a negative correlation between perceived risk and benefit (Alhakami and Slovic 1994). H2 proposes that there will be a negative correlation between perceived risk and the evaluation of organizational performance. SEM results presented in Table 4-10 support H2 and are consistent with the theory of cognitive consistency. Individuals that perceive less risk will judge firm performance more favorably.

#### 4.7.4 *Research Question 2*

Lipe and Salterio (2000) provided evidence that format may lead to differences in evaluating performance but our experiment was not structured in a way to replicate their findings. Not surprisingly, we find no difference between the two format groups (BSC and unformatted) in the evaluation of organizational performance ( $t = -.32, p = .749$ ). We were curious, however, if format would have any effect on the percentage of performance measures perceived to be important. We find no evidence that format has any such effect.

### 4.8 **Implications**

With many firms replacing traditional performance measurement systems with systems that combine the use of financial and nonfinancial measures, this paper provides relevant information on how managers utilize the information provided in these new performance measurement systems. Individuals appear to rate the performance of an organization differently even when they are provided the same information on which to make their evaluation. We provide further evidence on IOA and its impact on the use of information by examining the relationship between IOA and the number of measures perceived to be important in evaluating organizational performance. We find that individuals intolerant of ambiguity behave as expected and seek out more information the more intolerant they are. We further provide some support that more tolerant individuals behave in an opposite fashion. Future researchers can utilize this information in order to design experiments that do not rely on tools like regression which assumes linearity.

Many prior studies use students to examine the relationships that we have studied in this paper. Using students leads to potential problems of generalizability. This paper does not suffer from this limitation. We used working professionals who were top executives in their firms. Over 95% of these

professionals had at least 10 years of experience and over 57% had at least 20 years of experience.

This paper provides evidence in the performance measurement arena of the relationships between perceived amount of information, perceived risk, and evaluation. It is these relationships that answer the question: Why study factors like IOA that may influence the use of information in evaluations of organizational performance? It is important to study the factors that may influence the amount of information used in an evaluation of organizational performance because the amount of information used affects the evaluation itself. Our paper provides additional evidence that supports literature in psychology and marketing suggesting that information content is positively related to how an individual perceives the item of interest, in our case the performance of the organization. We explore this relationship in more detail and find that perceived risk is a mediating variable within this relationship.

## 4.9 References

- Alhakami, A. S. and P. Slovic. 1994. A psychological study of the inverse relationship between perceived risk and perceived benefit. *Risk Analysis* 14 (6): 1085-1096.
- Ashford, S. J. and L. L. Cummings. 1985. Proactive feedback seeking: the instrumental use of the information environment. *Journal of Occupational Psychology* 58: 67-79.
- Banker, R. D., H. Chang, and M. J. Pizzini. 2004. The balanced scorecard: judgmental effects of performance measures linked to strategy. *The Accounting Review* 79 (1): 1-23.
- Bennett, N., D. M. Herold, and S. J. Ashford. 1990. The effects of tolerance for ambiguity on feedback-seeking behavior. *Journal of Occupational Psychology* 63: 343-348.
- Budner, S. 1962. Intolerance of ambiguity as a personality variable. *Journal of Personality* 30: 29-50.
- Chenhall R. H. and D. Morris. 1986. The impact of structure, environment, and interdependence on the perceived usefulness of management accounting systems. *The Accounting Review* 61 (1): 1-35.
- Cox, D. F. and S. U. Rich. 1964. Perceived risk and consumer decision-making – the case of telephone shopping. *Journal of Marketing Research* 1 (4): 32-39.
- Crocker, K. E. 1986. The influence of the amount and type of information on individuals' perception of legal services. *Journal of the Academy of Marketing Science* 14 (4): 18-27.
- DeBusk, G. K., R. M. Brown and L. N. Killough. Components and relative weights in utilization of performance measurement systems like the balanced scorecard. *British Accounting Review* 35 (3): 215-231.
- Dermer, J. D. 1973. Cognitive characteristics and the perceived importance of information. *The Accounting Review* 48 (3): 511-519.
- Dilalla, L. F. 2000. Structural equation modeling: uses and issues in H.E.A. Tinsley and S.D. Brown (eds.) *Handbook of Multivariate Statistics and Mathematical Modeling*. San Diego, CA: Academic Press.

- Driver, M. J. and T. J. Mock. 1975. Human information processing, decision style theory, and accounting information systems. *The Accounting Review* 50 (3): 490-508.
- Ehrlich, D. 1965. "Intolerance of ambiguity", Walk's A scale: historical comment. *Psychological Reports* 17: 591-594.
- Faircloth, A. W. and D. N. Ricchiute. 1981. Ambiguity intolerance and financial reporting alternatives. *Accounting, Organizations and Society* 6 (1): 53-67.
- Furnham, A. 1994. A content, correlational and factor analytic study of four tolerance of ambiguity questionnaires. *Personality and Individual Differences* 16 (3): 403-410.
- and T. Ribchester. 1995. Tolerance of ambiguity: a review of the concept, its measurement and applications. *Current Psychology* 14 (3): 179-199.
- Frenkel-Brunswick, E. 1949. Intolerance of ambiguity as an emotional and perceptual personality variable. *Journal of Personality* 18:108-143.
- Ghosh, D. and M. R. Ray. 1997. Risk, ambiguity, and decision choice: some additional evidence. *Decision Sciences* 28 (1): 81-104.
- Gul, F. A. 1984. The joint and moderating role of personality and cognitive style on decision making. *The Accounting Review* 59 (2): 264-277.
- , 1986. Tolerance for ambiguity, auditors' opinions and their effects on decision making. *Accounting and Business Research* (Spring): 99-105.
- Gemünden, H. G. 1985. Perceived risk and information search. A systematic meta-analysis of the empirical evidence. *International Journal of Research in Marketing* 2 (2): 79-100.
- Heider, F. 1946. Attitudes and cognitive organization. *Journal of Psychology* 21: 107-112.
- Hair, J. F. Jr., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. *Multivariate Data Analysis, 5<sup>th</sup> ed.*, Upper Saddle River, NJ: Prentice-Hall, Inc.
- Ho, J. L. and W. Rodgers. 1993. A review of accounting research on cognitive characteristics. *Journal of Accounting Literature* (12): 101-130.
- Ittner, C. D., D. F. Larcker, and M. W. Meyer. 2003. Subjectivity and the weighting of performance measures: evidence from a balanced scorecard. *The Accounting Review* 78 (3): 725-758.

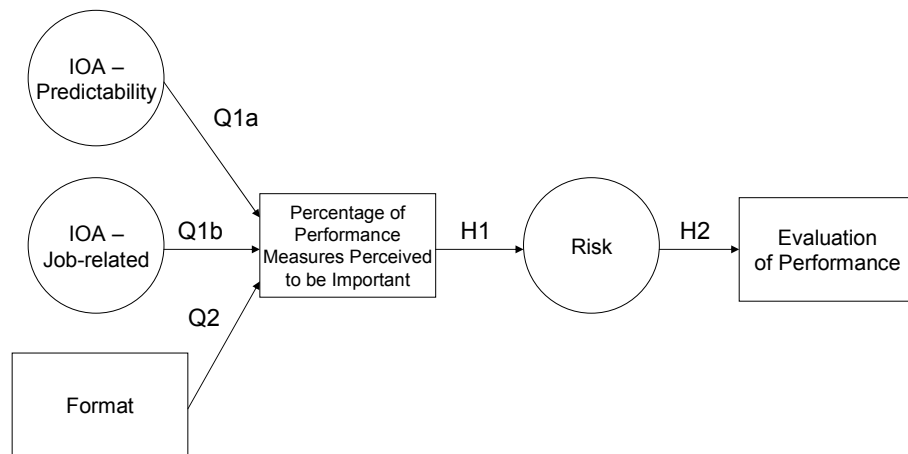
- Kaplan, R. S. and D. P. Norton. 2001a. Transforming the balanced scorecard from performance measurement to strategic management: part I, *Accounting Horizons* 15 (1): 87-104.
- and -----, 2001b. Transforming the balanced scorecard from performance measurement to strategic management: part II. *Accounting Horizons* 15 (2): 147-160.
- and -----, 2001c. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, MA: Harvard Business School Publishing Corporation.
- Kim, M. and S. J. Lennon. 2000. Television shopping for apparel in the United States: effects of perceived amount of information on perceived risks and purchase intentions. *Family and Consumer Sciences Research Journal* 28 (3): 301-330.
- Lal, M. and L. Hassel. 1998. The joint impact of environmental uncertainty and tolerance of ambiguity on top managers' perceptions of the usefulness of non-conventional management accounting information. *Scandinavian Journal of Management* 14 (3): 259-271.
- Lipe, M. S., and S. E. Salterio. 2000. The balanced scorecard: the judgmental effects of common and unique performance measures. *The Accounting Review* 75 (3): 283-298.
- and -----, 2002. A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society* 27: 531-540.
- MacDonald, Jr., A. P. 1970. Revised scale for ambiguity tolerance. *Psychological Reports* (26): 791-798.
- McGhee, W., M. D. Shields and J. G. Birnberg. 1978. *The Accounting Review* 53 (3): 681-697.
- McGuire, W. J. 1968. Theory of the structure of human thought in R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, and P. H. Tannenbaum (eds.) *Theories of Cognitive Consistency*. Chicago, IL: Rand McNally: 140-162.
- Mitra, K., M. C. Reiss, and L. M. Capella. 1999. An examination of perceived risk, information search and behavioral intentions in search, experience and credence services. *The Journal of Services Marketing* 13 (3): 208-228.



- Murray, K. B. 1991. A test of services marketing theory: consumer information acquisition activities. *Journal of Marketing* 55 (1): 10-25.
- Norton, R. W. 1975. Measurement of ambiguity tolerance. *Journal of Personality Assessment* 39 (6): 607-619.
- Nunnally, J. 1978. *Psychometric theory*, 2<sup>nd</sup> edition. New York: McGraw-Hill.
- O'Conner, P. 1952. Ethnocentrism, "intolerance of ambiguity" and abstract reasoning ability. *Journal of Abnormal and Social Psychology* (47): 526-530.
- Oliver, J. and E. Flamholtz. 1978. Human resource replacement cost numbers, cognitive information processing, and personnel decisions: a laboratory experiment. *Journal of Business Finance and Accounting* 5 (2): 137-157.
- Rydell, S. and E. Rosen. 1966. Measurement and some correlates of need cognition. *Psychological Reports* (19): 139-165.
- San Miguel, J. G. 1976. Human information processing and its relevance to accounting: a laboratory study. *Accounting, Organizations and Society* 1 (4): 357-373.
- Savich, R. 1977. The use of accounting information in decision making. *The Accounting Review* 52 (3):642-652.
- Silk, S. 1998. Automating the balanced scorecard. *Management Accounting* (May): 38-44.
- Tsui, J. 1993. Tolerance for ambiguity, uncertainty audit qualifications and bankers' perceptions. *Psychological Reports* 72: 915-919.
- Wright, M. E. and R. A. Davidson. 2000. The effect of auditor attestation and tolerance for ambiguity on commercial lending decisions. *Auditing: A Journal of Practice & Theory* 19 (2): 67-81.

#### 4.10 Figures and Tables

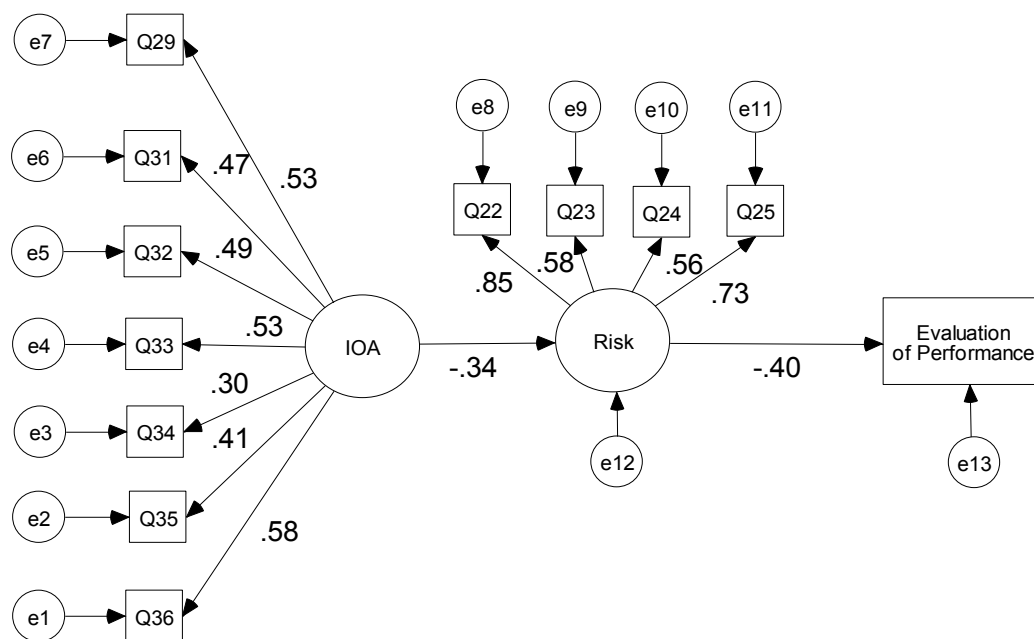
**Figure 4-1 IOA and Organizational Performance – Proposed Model**



**Figure 4-2 IOA and Organizational Performance – Final Model**

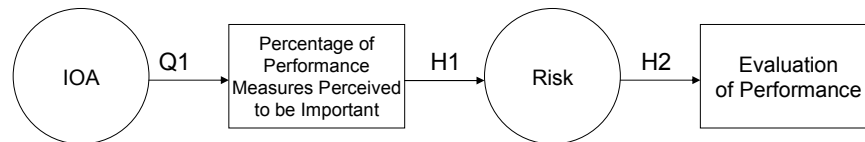


**Figure 4-3 IOA and Organizational Performance – Final Model SEM Results**

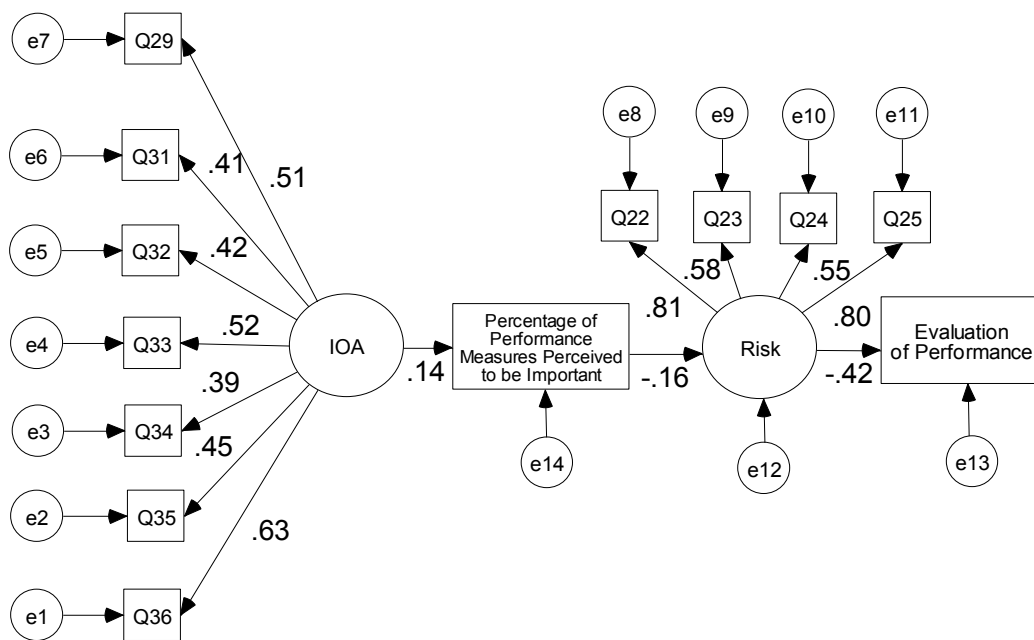


Indicator variables refer to question numbers from Appendix F.  
Numbers on paths are standardized regression weights.

**Figure 4-4 IOA and Organizational Performance – Supplemental Model**



**Figure 4-5 IOA and Org. Performance – Supplemental Model SEM Results**



Indicator variables refer to question numbers from Appendix F.

Numbers on paths are standardized regression weights.

**Table 4-1 Correlations between IOA Scales**

	MAT-50	Walk	AT-20
MAT-50			
Walk	0.54		
AT-20	0.82	0.62	
Budner	0.47	0.44	0.57

(Source: Furnham 1994, 406)

**Table 4-2 Prior Studies Using IOA as an Independent Variable**

Author(s)	Year	Dependent Variable	Scale	IOA Significant?	Interaction Significant?
Dermer	1973	Importance of information	Budner	Yes	N/A
McGhee et al.	1978	Additional information sought / decision confidence	Budner	No	N/A
Oliver and Flamholtz	1978	Decision confidence	Budner	Yes	N/A
Faircloth and Ricchiute	1981	Financial reporting alternatives / disclosure	MAT-50	No	N/A
Gul	1984	Decision confidence	AT-20	No	Yes
Gul	1986	Decision confidence	AT-20	No	Yes
Bennett et al.	1990	Feedback seeking behavior	MAT-50	Yes <sup>(a)</sup>	N/A
Tsui	1993	Interest rate	AT-20	Yes	N/A
Ghosh and Ray	1997	Decision confidence	Budner	Yes	N/A
Lal and Hassel	1998	Perceived usefulness of information	Budner	No	Yes
Wright and Davidson	2000	Interest rate / loan approval	AT-20	Yes	Yes

<sup>(a)</sup> Significance found only for the job-related construct.



**Table 4-3 Budner Scale Factor Analysis (Varimax Rotation), Pilot Study 1  
(n = 764)**

Appendix D Question No.	Factor 1 (Predictability)	Factor 2	Factor 3	Factor 4	Factor 5
9	<b>0.67</b>	-0.02	0.13	-0.09	-0.09
10	<b>0.61</b>	0.03	0.06	-0.06	-0.05
8	<b>0.58</b>	0.03	-0.12	0.01	-0.04
1	<b>0.50</b>	0.03	0.07	-0.13	-0.09
13	<b>0.42</b>	-0.08	-0.15	-0.04	0.03
2	0.23	-0.01	-0.02	-0.04	0.01
12	-0.07	<b>0.99</b>	0.09	0.04	0.11
15	-0.03	0.08	<b>0.57</b>	0.00	0.13
16	0.12	0.00	<b>0.54</b>	0.06	0.11
7	-0.13	0.09	0.27	0.20	0.09
14	0.17	0.01	-0.22	-0.04	0.12
11	0.15	0.15	0.16	0.11	0.08
4	-0.12	0.11	0.02	<b>0.70</b>	0.06
3	-0.09	-0.04	0.09	<b>0.38</b>	0.11
6	0.02	0.07	0.15	0.07	<b>0.53</b>
5	0.17	-0.04	-0.06	-0.15	<b>-0.46</b>

**Table 4-4 Cross Reference of Items Used in Variable Development**

Appendix D No. (Budner Scale) Or Other Scale	Appendix E Question No.	Appendix F Question No.
Appendix D, No. 8	1	33
AT-20	2	
MAT-50	3	30
AT-20	4	
MAT-50	5	36
MAT-50	6	29
Appendix D, No. 10	7	32
Walk	8	
MAT-50	9	
Appendix D, No. 13	10	
Walk	11	31
AT-20 / MAT-50	12	34
Appendix D, No. 9	13	28
Appendix D, No. 1	14	35

**Table 4-5 IOA Factor Analysis (Varimax Rotation), Pilot Study 2  
(n = 36)**

Appendix E Question No.	Factor 1 (Predictability)	Factor 2 (Job-Related)
14	<b>0.68</b>	0.23
12	<b>0.66</b>	0.10
11	<b>0.63</b>	0.17
7	<b>0.56</b>	0.18
13	<b>0.50</b>	-0.01
5	-0.17	<b>0.97</b>
6	0.44	<b>0.67</b>
1	0.21	<b>0.48</b>
3	0.14	<b>0.47</b>

**Table 4-6 IOA Factor Analysis, Main Study – Sample 1  
(n = 166)**

Appendix F Question No.	Factor Loading
32	0.62
34	0.60
33	0.56
36	0.47
35	0.46
29	0.44
31	0.43

**Table 4-7 Risk Factor Analysis, Main Study – Sample 1  
(n = 166)**

Appendix F Question No.	Factor Loading
22	0.86
25	0.75
23	0.55
24	0.54

**Table 4-8 Descriptive Statistics, Main Study – Sample 2**  
(n = 247)

Appendix F Question No.	Theoretical Range	Actual Range	Mean	Standard Deviation	Factor Loadings	Cronbach Alpha
<b>IOA</b>						0.67
29	1 - 7	1 - 7	4.32	1.815	0.53	
31	1 - 7	1 - 6	2.56	1.366	0.47	
32	1 - 7	1 - 7	3.18	1.565	0.49	
33	1 - 7	1 - 7	3.07	1.608	0.53	
34	1 - 7	1 - 7	2.78	1.578	0.30	
35	1 - 7	1 - 7	2.30	1.206	0.41	
36	1 - 7	1 - 7	4.42	1.739	0.58	
<b>Risk</b>						0.77
22	1 - 7	1 - 7	3.62	1.830	0.85	
23	1 - 7	1 - 7	5.80	1.277	0.58	
24	1 - 7	1 - 7	2.60	1.185	0.56	
25	1 - 7	1 - 7	3.09	1.453	0.73	
<b>Performance</b>						
1	1 - 10	3 - 10	7.53	1.403	1.00	

**Table 4-9 Pearson Correlations, Main Study – Sample 2**  
(n = 247)

Ques. No.		29	31	32	33	34	35	36	22	23	24	25	1
<b><u>IOA</u></b>													
29	Corr.	1	.230(**)	.233(**)	.309(**)	.181(**)	.141(*)	.370(**)	-.138(*)	.010	-.068	-.012	.092
	Sig.		.000	.000	.000	.004	.026	.000	.030	.872	.288	.851	.148
31	Corr.	.230(**)	1	.290(**)	.190(**)	.139(*)	.246(**)	.226(**)	-.207(**)	-.106	-.157(*)	-.159(*)	.175(**)
	Sig.	.000		.000	.003	.029	.000	.000	.001	.095	.013	.012	.006
32	Corr.	.233(**)	.290(**)	1	.273(**)	.164(**)	.172(**)	.223(**)	-.246(**)	-.180(**)	-.181(**)	-.175(**)	.167(**)
	Sig.	.000	.000		.000	.010	.007	.000	.000	.005	.004	.006	.009
33	Corr.	.309(**)	.190(**)	.273(**)	1	.115	.228(**)	.320(**)	-.178(**)	-.017	-.122	-.145(*)	.095
	Sig.	.000	.003	.000		.071	.000	.000	.005	.789	.055	.022	.135
34	Corr.	.181(**)	.139(*)	.164(**)	.115	1	.146(*)	.225(**)	.043	.091	.048	.033	-.030
	Sig.	.004	.029	.010	.071		.022	.000	.505	.154	.455	.605	.644
35	Corr.	.141(*)	.246(**)	.172(**)	.228(**)	.146(*)	1	.259(**)	-.125(*)	-.075	-.129(*)	-.149(*)	.026
	Sig.	.026	.000	.007	.000	.022		.000	.049	.242	.043	.019	.681
36	Corr.	.370(**)	.226(**)	.223(**)	.320(**)	.225(**)	.259(**)	1	-.121	-.138(*)	-.149(*)	-.109	.023
	Sig.	.000	.000	.000	.000	.000	.000		.057	.030	.019	.087	.717
<b><u>Risk</u></b>													
22	Corr.	-.138(*)	-.207(**)	-.246(**)	-.178(**)	.043	-.125(*)	-.121	1	.512(**)	.446(**)	.618(**)	-.349(**)
	Sig.	.030	.001	.000	.005	.505	.049	.057		.000	.000	.000	.000
23	Corr.	.010	-.106	-.180(**)	-.017	.091	-.075	-.138(*)	.512(**)	1	.288(**)	.390(**)	-.274(**)
	Sig.	.872	.095	.005	.789	.154	.242	.030	.000		.000	.000	.000
24	Corr.	-.068	-.157(*)	-.181(**)	-.122	.048	-.129(*)	-.149(*)	.446(**)	.288(**)	1	.487(**)	-.180(**)
	Sig.	.288	.013	.004	.055	.455	.043	.019	.000	.000		.000	.005
25	Corr.	-.012	-.159(*)	-.175(**)	-.145(*)	.033	-.149(*)	-.109	.618(**)	.390(**)	.487(**)	1	-.256(**)
	Sig.	.851	.012	.006	.022	.605	.019	.087	.000	.000	.000		.000
<b><u>Performance</u></b>													
1	Corr.	.092	.175(**)	.167(**)	.095	-.030	.026	.023	-.349(**)	-.274(**)	-.180(**)	-.256(**)	1
	Sig.	.148	.006	.009	.135	.644	.681	.717	.000	.000	.005	.000	

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

**Table 4-10 SEM Standardized Regression Weights, Main Study – Sample 2  
(n = 247)**

Path	Research Question / Hypothesis	Standardized Regression Weight ( $\beta$ )	p-value
IOA ➔ Risk	Q1 / H1	- 0.34	0.000
Risk ➔ Performance	H2	- 0.40	0.000



**Table 4-11 Pct. Important Frequencies, Main Study – Sample 2  
(n = 247)**

	Frequency
Pct. Important <= 85%	82
85% < Pct. Important < 100%	62
Pct. Important = 100%	103
Total	247

**Table 4-12 Supplemental SEM Regression Weights, Main Study – Sample 2  
(n = 185)**

Path	Research Question / Hypothesis	Standardized Regression Weight ( $\beta$ )	p-value (2-tailed)
IOA → Pct. Important	Q1	0.14	0.128
Pct. Important → Risk	H1	-0.16	0.049
Risk → Performance	H2	-0.42	0.000

Note, excludes 62 cases where Pct. Important was between 85% and 100%

**Table 4-13 Additional Descriptive Statistics, Main Study – Sample 2  
(n = 247)**

Variable	Theoretical Range	Actual Range	Mean	Standard Deviation
IOA	7 – 49	7 – 46	22.63	6.337
Pct. Important	0% - 100%	25% - 100%	87.87%	15.582%
Risk	4 – 28	6 – 28	15.11	4.484
Performance	1 – 10	3 – 10	7.53	1.403

**Table 4-14 Pearson / Spearman Correlations, Main Study – Sample 2  
(n = 247)**

Variable		IOA	Pct. Important	Risk	Performance
IOA	Correlation	1	.079	-.227(**)	.133(*)
	Sig. (2-tailed)		.215	.000	.036
Pct. Important	Correlation	.131(*)	1	-.057	.175(**)
	Sig. (2-tailed)	.040		.376	.006
Risk	Correlation	-.236(**)	-.135(*)	1	-.351(**)
	Sig. (2-tailed)	.000	.033		.000
Performance	Correlation	.124	.174(**)	-.341(**)	1
	Sig. (2-tailed)	.051	.006	.000	

Note, Pearson correlations are above the diagonal and Spearman correlations are below the diagonal.

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

**Table 4-15 Correlations after Median Split, Main Study – Sample 2**  
**Median Split Performed on IOA <sup>(a)</sup>**

Variable		IOA	Pct. Important
<b>Panel A – Low on IOA (n =118)</b>			
IOA	Correlation	1	-.160
	Sig. (2-tailed)		.084
Pct. Important	Correlation	-.094	1
	Sig. (2-tailed)	.313	
<b>Panel B – High on IOA (n = 109)</b>			
IOA	Correlation	1	.235(*)
	Sig. (2-tailed)		.014
Pct. Important	Correlation	.320(**)	1
	Sig. (2-tailed)	.001	

Note, Pearson correlations are above the diagonal and Spearman correlations are below the diagonal.

<sup>(a)</sup> Subjects with the median score of 23 were excluded rather than arbitrarily assigning them to either the low or high group.

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

**Table 4-16 Perceived Risk T-Test**  
**Between Those Low on Percentage of Performance Measures Perceived to be**  
**Important (Pct. Important) to Those High on Pct. Important**  
**After Median Split on Pct. Important <sup>(a)</sup>**  
**Main Study – Sample 2**

	n	Mean	Standard Deviation	t	df	Sig.
Pct. Important < 95%	113	15.56	4.549			
Pct. Important > 95%	103	14.21	4.249			
Difference		1.35		2.238	214	.026 <sup>(b)</sup>

<sup>(a)</sup> Subjects with the median score of 95% were excluded rather than arbitrarily assigning them to either the low or high group.

<sup>(b)</sup> Nonparametric tests yielded similar results.

## **5 CONCLUSION**

### **5.1 Summary and Conclusions**

This dissertation examines the utilization of organizational performance measurement systems in three separate research studies. The first study (chapter 2) examines the influence of strategy on the underlying components of a performance measurement system and provides encouraging results. Chapter 2 provides evidence that subjects have internalized the theory behind the current movement away from traditional financial control and management control systems towards systems of strategic control. Additional evidence is provided on relative weightings of performance measures. While bottom-line financial measures are perceived as more important than nonfinancial measures; evidence suggests that they were perceived to be influenced by driver-type nonfinancial measures.

The second study (chapter 3) examines the relative weights placed on performance measures. The results suggest that managers can count on a measure being perceived to have the same relative weight in a performance measurement system regardless of its position relative to its target (i.e. measures will not be over or under-weighted because of relative performance to target). The second

study also examines the relative weights on financial versus nonfinancial measures. It is possible that individuals place too much emphasis on financial measures and too little emphasis on the more forward-looking nonfinancial measures. Anthony and Govindarajan (2001, 451) state,

... not only are most senior managers well trained and adept with financial measures, they also keenly feel pressure regarding the financial performance of their companies. Shareholders are vocal, and boards of directors frequently apply pressure on the shareholders' behalf. This pressure may overwhelm the long-term, uncertain payback of the nonfinancial measures.

The first study (chapter 2) provides encouraging evidence that subjects perceive a link between driver-type measures, which are often nonfinancial, and outcome measures, many of which are financial. Financial measures, however, are perceived by the subjects of the first study as more important than nonfinancial measures. The results presented in chapter 3 provide additional evidence that users of performance measures systems do rely heavily on financial measures. Evidence of a halo effect is also provided in chapter 3. The participants allowed the financial results to bias their view of the nonfinancial results and potentially overstate the performance of the organization.

The third study (chapter 4) examines the influence of individual differences on the evaluation of organizational performance. Individuals rate the performance of the organization differently even when provided with the same performance measurement information to make the evaluation. Chapter 4 provides evidence that individuals intolerant of ambiguity seek out additional performance measurement information on which to base their evaluation. Evidence is also provided that the more intolerant the individual, the more information that is utilized. Individuals more tolerant of ambiguity tend to behave in a fashion that is counter to those more intolerant. The results tend to suggest



that the relationship between intolerance of ambiguity and the importance of information is not a linear one.

Chapter 4 also provides evidence of the positive effect on evaluation of organizational performance provided by reliance on additional performance measurement information. The more performance measurement information that is perceived to be important, the less risk that is perceived by the individual. The study further provides support for the theory of cognitive consistency by providing evidence of the negative correlation between perceived risk and evaluation of organizational performance.

## **5.2 Implications for Future Research**

Modern performance measurement systems are intended to provide a more “balanced” approach to the evaluation of organizational performance. The potential negative effects of focusing primarily on bottom-line financial measures like return-on-investment are well documented. This dissertation provides evidence that potential cognitive difficulties inherent in the use of a system containing multiple measures may mitigate the proposed advantages of using multiple measures. Future research should continue to document the relative weights placed on financial and nonfinancial measures by using different subject pools and methodologies. Potential methods to reduce a financial measures bias should also be explored. Future research should also explore the extent to which the halo effect (financial performance affecting the perceived performance on nonfinancial measures) discussed in chapter 3 exists in other settings.

This dissertation provides evidence on the impact of a personality variable, intolerance of ambiguity, on the utilization of performance measurement information. The dissertation also provides evidence that the utilization of performance measurement information influences the evaluation of organizational

performance. Future research should explore other potential factors that may have an impact on the usage of information.

Prior research using intolerance of ambiguity (IOA) as an independent variable has produced inconsistent findings, due in part to the use of multidimensional measures with poor reliability (Ho and Rodgers 1993). This dissertation uses unidimensional constructs for IOA to gain insight into its influence on information usage. The evidence suggests that the relationship between IOA and information usage might not be a linear one. Future IOA research should be careful not to rely heavily on prior inconclusive findings as they may be driven by the lack of linearity or poor measurement reliability rather than the absence of the expected relationship.

This dissertation contributes to the existing literature in several areas and provides a foundation for future research. Overall, this dissertation adds to our knowledge of organizational performance measurement system utilization by examining the relative weightings of performance measures, the judgmental effects from utilization of performance measurement systems, the impact of intolerance of ambiguity on the importance of performance measurement data, and the link between performance measurement data and the perception of risk in the evaluation of organizational performance.

### 5.3 References

- Anthony, R. N., and V. Govindarajan. 2001. *Management Control Systems, 10<sup>th</sup> ed.*, New York, NY: The McGraw-Hill Companies, Inc.
- Ho, J. L. and W. Rodgers. 1993. A review of accounting research on cognitive characteristics. *Journal of Accounting Literature* (12): 101-130.

## **APPENDICES**

### **Appendix A Case Materials for Chapter 2 and Chapter 3**

#### *(a) Company Background*

Adapted from *Coors Case: Balanced Scorecard* (Grove et al., 2000)

### **COORS BREWING COMPANY**

#### **Company Background**

Coors had been a family owned and operated business from its inception in 1873 until 1993 when the first non-family member became President and Chief Operating Officer. However, Coors family members still held the Chairman of the Board of Directors and Chief Executive Officer positions and all voting stock. Only nonvoting, Class B common stock was publicly traded. Coors has been financed primarily by equity and has only borrowed capital twice in its corporate history. The first long-term debt, \$220 million, 8.5% notes, was issued in 1991 and the final \$40 million of principal was repaid by the end of 1999. The second long-term debt, \$100 million, 7% unsecured notes, was issued in a 1995 private placement. \$80 million of this principal is due in 2002 and the last \$20 million is due in 2005.

In the mid-1970's Coors was a regional brewery with an eleven-state market, selling one brand in a limited number of packages through approximately two hundred distributors. Traditionally, Coors beer had been a non-pasteurized, premium beer. (However, with a recently developed sterilization process, its products now have the

same shelf life as its competitors' pasteurized products.) Coors plant in Golden, Colorado was its only production facility and it had no other distribution centers.

Over the next 25 years, Coors changed dramatically by expanding into all fifty states and various foreign markets. By the end of the twentieth century, Coors had production facilities in Golden, Colorado, Memphis, Tennessee, Elkton, Virginia, and Zaragoza, Spain. It had expanded to using twenty-one "satellite redistribution centers" in the United States before a special project reduced this number to eight. Beer shipments were made by both truck and railroad cars. Coors had approximately 650 domestic beer distributors although about 200 of them accounted for 80% of Coors total sales. Coors also had several joint ventures and international distributors in Canada, the Caribbean, Latin American, Europe, and the Pacific.

Coors had sixteen beer brands, including a specialty line, Blue Moon, that competed with the domestic micro brewing industry. However, Coors continued to focus upon its four key premium brands, Coors Light, Original Coors, Killian's Irish Red, and Zima. Coors Light was the fourth largest selling beer in the U.S. In packaging, Coors had to compete with the major competitors' value packaging, such as twelve-packs and thirty-packs. In 1959, Coors introduced the nation's first all-aluminum beverage can and in the late 1990's, it had introduced a baseball bat bottle and a football pigskin bottle. There were also numerous state labeling laws to meet, such as returnable information, and packaging graphics to reinforce the Rocky Mountains image for Coors beer.

Competition in the beer industry was strong, especially in the United States. Anheuser-Busch (A/B) was the market leader with approximately 44% of the U.S. market, 80 million barrels sold, \$8 billion beer sales and \$1 billion net profit. Due to its size, A/B was the acknowledged price leader in the industry. A/B also had thirteen domestic production plants, including one in Ft. Collins, Colorado, to achieve its customer service goal of having no major domestic distributor more than 500 miles away from one of its beer production plants.

Number two in this market was Miller, owned by Philip Morris, with approximately 22% market share, 40 million barrels sold, \$4 billion beer sales, and \$460 million net profit. Miller also had seven domestic production plants. Coors was number three with an 11% market share, 20 million barrels sold, \$2 billion beer sales, and \$80 million net profit. Coors had three production plants in the United States. Its Colorado plant was the largest brewery in the world and served 70% of the U.S. market with its ten can lines, six bottle lines, and two keg lines.

There were no other domestic brewers with market share in excess of 5%. In the late 1990's, there had been consolidation of the larger companies in the domestic beer industry. The most recent example was Stroh Brewing Company (SBC) with about 5% market share. SBC had signed agreements to sell its major brands to Miller and the remaining brands to Pabst Brewing Company. SBC would then exit the beer industry by 2000.

From 1983 through 1998, Coors was the only major U.S. brewer to increase its sales volume each year although industry sales had grown only about 1% per year in the 1990's. Coors had outpaced the industry volume growth rate by one or two percentage points each year. Coors had accomplished this growth by building its key premium brands in key markets and strengthening its distributor network, recently with improved supply chain management.

## **Coors Vision Statement and Business Strategies**

Coors vision statement was as follows:

Our company has a proud history of visionary leadership, quality products and dedicated people which has enabled us to succeed in a highly competitive and regulated industry. We must continue to build on this foundation and become even more effective by aligning and uniting the human, financial and physical aspects of our company to bring great tasting beer, great brands and superior service to our distributors, retailers and consumers and to be a valued neighbor in our communities. Our continued success will require teamwork and an even stronger dedication by every person in our organization to a common purpose, our Vision. Achieving our Vision requires that we begin this journey immediately and with urgency for it will require significant change for us to thrive and win in our industry.

Using this vision statement, top management had decided to focus on four fundamentals: improving quality, improving service, boosting profitability, and developing employee skills. In the 1997 Coors annual report, both the CEO and the President discussed the following general business strategies or “six planks” to drive these fundamentals in the future:

1. baseline growth: we will profitably grow key brands and key markets,
2. incremental growth: we will selectively invest to grow high potential markets, channels, demographics, and brands,
3. product quality: we will continuously elevate consumer perceived quality by improving taste, freshness, package integrity, and package appearance at point of purchase,
4. distributor service: we will significantly enhance distributor service as measured by improved freshness, less damage, increased on-time arrivals, and accurate order fill at a lower cost to Coors,
5. productivity gains: we will continuously lower total company costs per barrel so Coors can balance improved profitability, investments to grow volume, market share, and revenues, and funding for the resources needed to drive long-term productivity and success, and
6. people: we will continuously improve our business performance through engaging and developing our people.

*(b) Descriptions of Performance Measures*

1. **Market share:** increase in Coors market share of the domestic beer market. Reported monthly.
2. **Barrels produced per labor hour:** total barrels packaged per labor hour worked. Reported daily.
3. **Baseline growth:** percentage of sales growth for key brands and key markets. Reported annually.
4. **Beer waste and package scrap:** waste and scrap as a percent of total production. Reported weekly.
5. **Community involvement:** number of volunteer hours per employee.
6. **Customer complaints:** total customer complaints related to taste, freshness, package integrity, appearance, and foreign objects per 100,000 barrels sold. Reported weekly.
7. **Incremental growth:** percentage of sales growth for high potential brands and markets. Reported annually.
8. **Load item accuracy:** percent beer line items shipped exact as compared to the commitment to the distributor. Reported daily.
9. **Load schedule performance:** truck or rail car loaded on time (within two hours of scheduled lead time). Reported daily.
10. **Manufacturing cost per barrel:** total plant cost (brewing materials, production labor, support labor, operating supplies, manufacturing overhead, maintenance materials, and packing materials) on a per barrel basis. Reported monthly.
11. **Net profit per barrel:** net income, excluding all special charges and special credits, on a per barrel basis. Reported monthly.
12. **New products:** number of new brands introduced each year. Reported annually.
13. **Plant productivity:** actual production hours divided by total production hours including run time, unplanned downtime and changeovers. Reported daily.
14. **Production stability:** total quantity of correct product from the beer lines as scheduled within a four-hour window as a percent of total production. Reported daily.
15. **Quality index:** weighted roll-up of component quality measures concerning plant audits, microbiology, and chemistry on a scale of 1 to 100. Reported quarterly.
16. **Return on invested capital:** after-tax income before interest expense and any special charges or credits divided by the sum of average total debt and shareholders' equity. Reported annually.
17. **Revenues per barrel:** total net revenues after reducing gross revenues by the excise taxes imposed by federal laws. Reported monthly.
18. **Safety (lost work case incident rate):** total recordable cases that resulted in lost work as a ratio to total labor hours worked. Reported quarterly.
19. **Safety (total case incident rate):** total OSHA recordable case incidents as a ratio to total number of labor hours worked. Reported quarterly.
20. **Selling, general and administrative (S, G & A) cost per barrel:** all S, G & A costs (includes selling, advertising, outbound transportation, distribution, and all general and administrative costs; excludes interest expenses, special charges, and income taxes) on a per barrel basis. Reported monthly.
21. **Skills:** inventory of cross-functional employee skills. Reported annually.
22. **Throughput:** total barrels packaged each period. Reported monthly.
23. **Training:** number of training hours per employee. Reported annually.
24. **Warehouse moves:** actual shipments as a percent of planned shipments. Reported Weekly.

Adapted from Coors Case: Balanced Scorecard (Grove et al., 2000).

## Appendix B Performance Measures for Chapter 2 Survey

No.	Performance Measure	Year 2000		Circle the number indicating importance of the measure in evaluating Coors.	Circle the number indicating your view of Coors' performance on this measure.
		Actual	Target	1 = Not important; 6 = Critical	1 = Very Poor; 6 = Very Good
1	Annual Market Share Increase	0.05%	0.05%	1 2 3 4 5 6	1 2 3 4 5 6
2	Barrels Produced per Labor Hour	6.0	6.0	1 2 3 4 5 6	1 2 3 4 5 6
3	Baseline Growth (for key brands and markets)	2.1%	2.0%	1 2 3 4 5 6	1 2 3 4 5 6
4	Beer Waste & Package Scrap (as a % of total production)	0.5%	0%	1 2 3 4 5 6	1 2 3 4 5 6
5	Community Involvement (volunteer hours per employee annually)	28	30	1 2 3 4 5 6	1 2 3 4 5 6
6	Customer Complaints (per 100,000 barrels sold)	0.05%	0.05%	1 2 3 4 5 6	1 2 3 4 5 6
7	Incremental Growth (for high potential brands and markets)	3.2%	3.0%	1 2 3 4 5 6	1 2 3 4 5 6
8	Load Item Accuracy	96%	100%	1 2 3 4 5 6	1 2 3 4 5 6
9	Load Schedule Performance	60%	100%	1 2 3 4 5 6	1 2 3 4 5 6
10	Manufacturing Cost per Barrel	\$54	\$53	1 2 3 4 5 6	1 2 3 4 5 6
11	Net Profit per Barrel	\$6	\$6	1 2 3 4 5 6	1 2 3 4 5 6
12	New Products (new brands introduced each year)	7	6	1 2 3 4 5 6	1 2 3 4 5 6
13	Plant Productivity	75%	80%	1 2 3 4 5 6	1 2 3 4 5 6
14	Production Stability	65%	100%	1 2 3 4 5 6	1 2 3 4 5 6
15	Quality Index (out of 100)	91	90	1 2 3 4 5 6	1 2 3 4 5 6
16	Return on Invested Capital	11.9%	12.0%	1 2 3 4 5 6	1 2 3 4 5 6
17	Revenue per Barrel	\$100	\$100	1 2 3 4 5 6	1 2 3 4 5 6
18	Safety (lost work incident rate)	0.0001	0	1 2 3 4 5 6	1 2 3 4 5 6
19	Safety (total case incident rate)	0.0002	0	1 2 3 4 5 6	1 2 3 4 5 6
20	Selling, General & Administrative Cost per Barrel	\$26	\$27	1 2 3 4 5 6	1 2 3 4 5 6
21	Skills (inventory of cross-functional employee skills)	7	7	1 2 3 4 5 6	1 2 3 4 5 6
22	Throughput per Month (millions of barrels)	1.74	1.75	1 2 3 4 5 6	1 2 3 4 5 6
23	Training (hours per employee annually)	39	40	1 2 3 4 5 6	1 2 3 4 5 6
24	Warehouse Moves (actual shipments as a % of plan)	95%	100%	1 2 3 4 5 6	1 2 3 4 5 6
Please evaluate the overall performance of Coors Brewing Company in the year 2000. ==>				1 2 3 4 5 6	1 2 3 4 5 6

Adapted from *Coors Case: Balanced Scorecard* (Grove et al., 2000).



## Appendix C Performance Measures for Chapter 3 Survey

### Version 1

No.	Performance Measure	Year 2000		Circle the number indicating importance of the measure in evaluating Coors.	Circle the number indicating your view of Coors' performance on this measure.
		Actual	Target	1 = Not important; 6 = Critical	1 = Very Poor; 6 = Very Good
1	Annual Market Share Increase	0.04%	0.05%	1 2 3 4 5 6	1 2 3 4 5 6
2	Barrels Produced per Labor Hour	5.0	6.0	1 2 3 4 5 6	1 2 3 4 5 6
3	Baseline Growth (for key brands and markets)	2.0%	2.0%	1 2 3 4 5 6	1 2 3 4 5 6
4	Beer Waste & Package Scrap (as a % of total production)	0.5%	0%	1 2 3 4 5 6	1 2 3 4 5 6
5	Community Involvement (volunteer hours per employee annually)	25	30	1 2 3 4 5 6	1 2 3 4 5 6
6	Customer Complaints (per 100,000 barrels sold)	0.06%	0.05%	1 2 3 4 5 6	1 2 3 4 5 6
7	Incremental Growth (for high potential brands and markets)	2.5%	3.0%	1 2 3 4 5 6	1 2 3 4 5 6
8	Load Item Accuracy	93%	100%	1 2 3 4 5 6	1 2 3 4 5 6
9	Load Schedule Performance	55%	100%	1 2 3 4 5 6	1 2 3 4 5 6
10	Manufacturing Cost per Barrel	\$55	\$53	1 2 3 4 5 6	1 2 3 4 5 6
11	Net Profit per Barrel	\$3	\$6	1 2 3 4 5 6	1 2 3 4 5 6
12	New Products (new brands introduced each year)	5	6	1 2 3 4 5 6	1 2 3 4 5 6
13	Plant Productivity	76%	80%	1 2 3 4 5 6	1 2 3 4 5 6
14	Production Stability	50%	100%	1 2 3 4 5 6	1 2 3 4 5 6
15	Quality Index (out of 100)	85	90	1 2 3 4 5 6	1 2 3 4 5 6
16	Return on Invested Capital	7.1%	12.0%	1 2 3 4 5 6	1 2 3 4 5 6
17	Revenue per Barrel	\$100	\$100	1 2 3 4 5 6	1 2 3 4 5 6
18	Safety (lost work incident rate)	0.0001	0	1 2 3 4 5 6	1 2 3 4 5 6
19	Safety (total case incident rate)	0.0002	0	1 2 3 4 5 6	1 2 3 4 5 6
20	Selling, General & Administrative Cost per Barrel	\$28	\$27	1 2 3 4 5 6	1 2 3 4 5 6
21	Skills (inventory of cross-functional employee skills)	6	7	1 2 3 4 5 6	1 2 3 4 5 6
22	Throughput per Month (millions of barrels)	1.72	1.75	1 2 3 4 5 6	1 2 3 4 5 6
23	Training (hours per employee annually)	42	40	1 2 3 4 5 6	1 2 3 4 5 6
24	Warehouse Moves (actual shipments as a % of plan)	95%	100%	1 2 3 4 5 6	1 2 3 4 5 6
Please evaluate the overall performance of Coors Brewing Company in the year 2000. ==>					1 2 3 4 5 6

Adapted from *Coors Case: Balanced Scorecard* (Grove et al., 2000).

## Appendix D Survey Materials for Chapter 4, Pilot Study 1

### IOA Scale (Budner 1962)

Select the letter which ranks the extent of your agreement with each statement, according to the following scale:

- a Agree strongly
- b Agree
- c Agree slightly
- d Disagree slightly
- e Disagree
- f Disagree strongly

Please attempt to choose the rating that best describes how you believe or feel.

No.	Question	Circle one
1	An expert who doesn't come up with a definite answer probably doesn't know too much.	a b c d e f
2	In the long run it is possible to get more done by tackling small, simple problems rather than by tackling large and complicated ones.	a b c d e f
3	Many of our most important decisions are based upon insufficient information.	a b c d e f
4	People who fit their lives to a schedule probably miss most of the joy of living.	a b c d e f
5	There is really no such thing as a problem that can't be solved.	a b c d e f
6	It is more fun to tackle a complicated problem than to solve a simple one.	a b c d e f
7	People who insist upon a yes or no answer just don't know how complicated things really are.	a b c d e f
8	A good job is one where what is to be done and how it is to be done are always clear.	a b c d e f
9	The sooner we all acquire similar values and ideals the better.	a b c d e f
10	A person who leads an even, regular life in which few surprises or unexpected happenings arise, really has a lot to be grateful for.	a b c d e f
11	I would like to live in a foreign country for awhile.	a b c d e f
12	Teachers or supervisors who hand out vague assignments give a chance for one to show initiative and originality.	a b c d e f
13	What we are used to is always preferable to what is unfamiliar.	a b c d e f
14	I like parties where I know most of the people more than ones where all or most of the people are complete strangers.	a b c d e f
15	A good teacher is one who makes you wonder about your way of looking at things.	a b c d e f
16	Often the most interesting and stimulating people are those who don't mind being different and original.	a b c d e f

## Appendix E Survey Materials for Chapter 4, Pilot Study 2

Instructions: Select the item which ranks the extent of your agreement with each statement. Please attempt to choose the rating that best describes how you believe or feel.

1. A good job is one where what is to be done and how it is to be done are always clear.
  - ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
2. The best part of working a jigsaw puzzle is putting in that last piece.
  - ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
3. If I were a doctor, I would prefer the uncertainties of a psychiatrist to the clear and definite work of someone like a surgeon or X-ray specialist.
  - ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
4. A problem has little attraction for me if I don't think it has a solution.
  - ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
5. In a situation in which other people evaluate me, I feel a great need for clear and explicit evaluations.
  - ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
6. If I am uncertain about the responsibilities of a job, I get very anxious.
  - ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
7. A person who leads an even, regular life in which few surprises or unexpected happenings arise, really has a lot to be grateful for.
  - ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly

8. Nobody can have feelings of love and hate towards the same person.
- ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
9. I function very poorly whenever there is a serious lack of communication in a job situation.
- ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
10. What we are used to is always preferable to what is unfamiliar.
- ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
11. It is better to keep on with the present method of doing things than to take away that which might lead to chaos.
- ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
12. If I were a scientist, I might become frustrated because my work would never be completed (science will always make new discoveries).
- ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
13. The sooner we all acquire similar values and ideals the better.
- ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly
14. An expert who doesn't come up with a definite answer probably doesn't know too much.
- ☐ Agree strongly
  - ☐ Agree
  - ☐ Agree slightly
  - ☐ Disagree slightly
  - ☐ Disagree
  - ☐ Disagree strongly

## Appendix F Survey Materials for Chapter 4, Main Study (BSC Format)

### Background Information and Instructions

You are the new Vice-President of Operations of XYZ, Inc. Your first job as V.P of Operations is to evaluate the performance of one of the manufacturing divisions, Division A. Below are Division A's performance measures for 2003. Please review the performance measures and answer the questions that follow.

### XYZ, Inc. – Division A Performance Measures

Performance Measure	Year 2003	
	Actual	Target
<b>Customer Perspective</b>		
Customer Satisfaction Index (out of 100)	91	90
Market Share	20.2%	20%
On-time Delivery	95.5%	95%
No. of Warranty Claims per 1,000 Shipments	39.7	40
Average Lead Time in Days (Order to Delivery)	7.2	7
<b>Internal Business Perspective</b>		
Inventory Turns	20	20
Defects per 1,000 Units	49.5	50
Purchase Price Variance - Favorable (Unfavorable)	1.02%	1.0%
Labor Efficiency	89.5%	90%
Manufacturing Cycle Time in Days	4.1	4
<b>Learning &amp; Growth Perspective</b>		
Employee Satisfaction Index (out of 100)	91	90
Employee Turnover	1.9%	2%
Training Hours per Employee	39	40
Safety (OSHA Reportable Lost Time Accidents)	2	2
Percentage Involvement in Voluntary Quality Circles	81.6%	80%
<b>Financial Perspective</b>		
Operating Income (\$000)	\$10,019	\$10,000
Return on Investment	12.0%	12%
Revenue Growth Rate	10.1%	10%
Revenue per Unit	\$1,000	\$1,000
Free Cash Flow (\$000)	\$5,141	\$5,000

## Part I

**Instructions:** Please answer the following questions in Part I based on your review of Division A's performance measures for 2003 (see page 2).

1. On a scale of 1 – 10 where **1 is very poor** and **10 is very good**, please circle the number that indicates your opinion of Division A's overall 2003 performance.

1   2   3   4   5   6   7   8   9   10

**Instructions:** Please indicate the importance of each of the individual measures to you in your overall evaluation of Division A's 2003 performance. The scale is 1 – 10 where **1 is not at all important** and **10 is very important**. Circle the number from 1 to 10 that best indicates your opinion.

- |   |                      |
|---|----------------------|
| 2. Customer Satisfaction Index                          | 1 2 3 4 5 6 7 8 9 10 |
| 3. Market Share   | 1 2 3 4 5 6 7 8 9 10 |
| 4. On-time Delivery                                     | 1 2 3 4 5 6 7 8 9 10 |
| 5. No. of Warranty Claims per 1,000 Shipments           | 1 2 3 4 5 6 7 8 9 10 |
| 6. Average Lead Time in Days (Order to Delivery)        | 1 2 3 4 5 6 7 8 9 10 |
| 7. Inventory Turns                                      | 1 2 3 4 5 6 7 8 9 10 |
| 8. Defects per 1,000 Units                              | 1 2 3 4 5 6 7 8 9 10 |
| 9. Purchase Price Variance                              | 1 2 3 4 5 6 7 8 9 10 |
| 10. Labor Efficiency                                    | 1 2 3 4 5 6 7 8 9 10 |
| 11. Manufacturing Cycle Time in Days                    | 1 2 3 4 5 6 7 8 9 10 |
| 12. Employee Satisfaction Index                         | 1 2 3 4 5 6 7 8 9 10 |
| 13. Employee Turnover                                   | 1 2 3 4 5 6 7 8 9 10 |
| 14. Training Hours per Employee                         | 1 2 3 4 5 6 7 8 9 10 |
| 15. Safety (OSHA Reportable Lost Time Accidents)        | 1 2 3 4 5 6 7 8 9 10 |
| 16. Percentage Involvement in Voluntary Quality Circles | 1 2 3 4 5 6 7 8 9 10 |
| 17. Operating Income                                    | 1 2 3 4 5 6 7 8 9 10 |
| 18. Return on Investment                                | 1 2 3 4 5 6 7 8 9 10 |
| 19. Revenue Growth Rate                                 | 1 2 3 4 5 6 7 8 9 10 |
| 20. Revenue per Unit                                    | 1 2 3 4 5 6 7 8 9 10 |
| 21. Free Cash Flow                                      | 1 2 3 4 5 6 7 8 9 10 |

## Part II

**Instructions:** Please select the item which ranks the extent of your agreement with each statement regarding your evaluation of XYZ Inc. - Division A's performance. Please attempt to choose the rating that best describes how you believe or feel.

22. There was sufficient information within the given performance measures to evaluate Division A's overall performance.
  - a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
23. Additional information would have been beneficial in the evaluation of Division A's overall performance.
  - a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
24. The performance measures given for Division A were reliable measurements.
  - a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
25. I am confident in my evaluation of Division A.
  - a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
26. There is a great amount of subjectivity involved in evaluating a business unit's performance.
  - a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly

### Part III

**Instructions:** The questions in this section are unrelated to your evaluation of Division A. The questions in Part III are related to experience and decision-making style.

27. Please circle the selection that best indicates your years of business experience.
- a. Less than 10 years
  - b. 10 years to 19 years
  - c. 20 years to 29 years
  - d. 30 years or more

**Instructions:** For the remaining questions in Part III, please circle the item which ranks the extent of your agreement with each statement. Please attempt to choose the rating that best describes how you believe or feel.

28. The sooner we all acquire similar values and ideals the better.
- a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
29. If I am uncertain about the responsibilities of a job, I get very anxious.
- a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
30. If I were a doctor, I would prefer the uncertainties of a psychiatrist to the clear and definite work of someone like a surgeon or X-ray specialist.
- a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
31. It is better to keep on with the present method of doing things than to take away that which might lead to chaos.
- a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly

(Continued on next page.)



32. A person who leads an even, regular life, in which few surprises or unexpected happenings arise, really has a lot to be grateful for.
- a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
33. A good job is one where what is to be done and how it is to be done are always clear.
- a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
34. If I were a scientist, I might become frustrated because my work would never be completed (science will always make new discoveries).
- a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
35. An expert who doesn't come up with a definite answer probably doesn't know too much.
- a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly
36. In a situation in which other people evaluate me, I feel a great need for clear and explicit evaluations.
- a. Agree strongly
  - b. Agree
  - c. Agree slightly
  - d. Disagree slightly
  - e. Disagree
  - f. Disagree strongly

**Mailing Instructions:** Please return pages 3 through 6 to (*address omitted*). A return mail envelope has been provided. No postage is required.

## **VITA**

### **GERALD KENNETH DEBUSK**

#### **Education**

MBA from Liberty University, 1991

B.S. in Accounting from Virginia Polytechnic Institute and State University, 1983

#### **Certifications**

- Certified Public Accountant in North Carolina and Virginia
- Certified Management Accountant

#### **Publications / Conference Presentations**

DeBusk, G.K., R.M. Brown and L.N. Killough. Components and relative weights in utilization of performance measurement systems like the balanced scorecard. *British Accounting Review* 35 (3): 215-231.

DeBusk, G.K., L.N. Killough and R.M. Brown. Judgmental effects in the use of performance measurement systems. Paper presented (concurrent session) at the southeast regional meeting of the American Accounting Association, Charleston, SC, April 2003.

#### **Teaching Activities**

##### Courses Taught:

Cost Accounting, Managerial Accounting, Auditing, Fundamentals of Accounting (Graduate), Intermediate Accounting I & II, Accounting Information Systems, Principles of Accounting II, and Personal Finance

##### Teaching Positions Held:

- Graduate Teaching Assistant (2000 – 2004), Virginia Polytechnic Institute and State University
- Adjunct faculty positions (1991 – 2000) at Wilson College, Chambersburg, PA; Campbellsville University, Campbellsville, KY; Roanoke College, Salem, VA; New River Community College, Dublin VA

## **Related Experience**

### Xaloy, Inc.

Controller, Secretary-Treasurer, Pulaski, VA 1997 – 2000

- Responsible for worldwide accounting, treasury, credit, and information systems functions.
- Managed successful installation of new Enterprise Resource Planning software and Wide Area Network for all Xaloy facilities. The dollar value of the capital project (\$900,000) exceeded all preceding Xaloy projects.
- Developed and implemented a performance pay plan based on value added concepts.

### Ingersoll-Rand Company

Plant Controller, Shippensburg, PA / Campbellsville, KY 1992 – 1997

Various Manager and Supervisor Positions, Mocksville, NC / Shippensburg, PA 1988 – 1992

- Managed accounting department at a facility with \$200 million in annual sales.
- Designed and implemented a “pay-on-use” and “invoiceless” accounts payable system.
- Member of a 9-person team responsible for training more than 500 employees in Just-in-Time manufacturing.
- Designed Activity Based Costing System that was later implemented at a large manufacturing facility.

### Amsted Industries

Cost Accountant, Lynchburg, VA 1985 – 1988

- Managed actual costing system for a foundry with \$85 million in annual sales.

### Deloitte Haskins + Sells

Senior Assistant & Staff Accountant, Roanoke, VA 1983 – 1985

- Audited various SEC clients. Specialized in public utilities.

## **Awards**

- Floyd A. Beams Scholarships, 2003-4, 2002-3, 2001-2
- Institute of Management Accountants Outstanding Member, Winston-Salem Chapter, 1990