

## Choice and Change of Measures in Performance Measurement Models

Mary A. Malina  
Naval Postgraduate School  
208 Ingersoll Hall  
555 Dyer Road  
Monterey, CA USA 93943  
1-831-656-1187  
1-831-656-3407 Fax  
[mamalina@nps.navy.mil](mailto:mamalina@nps.navy.mil)

Frank H. Selto  
Leeds School of Business  
University of Colorado at Boulder  
Boulder, CO USA 80309-0419  
1-303-492-1549  
[frank.selto@colorado.edu](mailto:frank.selto@colorado.edu)

November 2003

**Acknowledgements:** We appreciate constructive comments from two anonymous referees and workshop participants at the University of Colorado at Boulder and the University of Notre Dame, and in particular Steve Rock, Phil Shane, Michael Morris, Margaret Shackell-Dowell, Tom Stober, Lisa Sedor, and Sandra Vera-Munoz.

**Data Availability:** This research is supported by data generously provided by the anonymous host company. Proprietary data beyond that disclosed in this paper are not available.

**Company Approval:** The host company has approved the dissemination and publication of this research.

## Choice and Change of Measures in Performance Measurement Models

### Abstract

This paper uses management control, resource-based and systems-based strategy theories to describe a large U.S. manufacturing company's efforts to improve profitability by designing and using a performance measurement model (PMM). This PMM includes multiple performance measures relevant to its distribution channel for products, repair parts and maintenance services. The PMM is intended to reflect the company's understanding of performance relations among strategic resources, operational capabilities, and desired financial outcomes. The PMM also reflects its intended distribution strategy, the types of performance necessary to achieve that strategy by its distributors, and its desired financial outcomes. Furthermore, the company uses the model to evaluate its North American distributors and intends to use these evaluations as a partial basis for annual and long-term rewards. Thus, the PMM embodies the measurable portion of the firm's management control system of its distribution channel.

The study addresses two research questions: (1) Which criteria does a firm base its choices of performance weights and measures? (2) Why does a firm change performance weights and measures? The questions are investigated using qualitative and quantitative analyses of archival documents and interviews with top managers and distributors. Principal findings are that measures were chosen consistently with emerging theory but some changes to the PMM were inconsistent with a naïve interpretation of that theory.

Key Words: performance measurement model, management control, non-financial performance measures

# Choice and Change of Measures for Performance Measurement Models

## 1. Introduction

Improving performance measurement at key parts of the value chain is one of management accounting's major roles. Valid performance measurement allows a firm to effectively describe and implement strategy, to guide employee behavior, assess managerial effectiveness, and provide the basis for rewards. Managers and researchers from diverse disciplines have sought to improve management of the value chain by building and using performance measurement models (PMM). PMM are comprehensive models of the firm as a system, which reflects organizational knowledge of the relations among various value-chain performance measures. Many organizations reportedly have created PMM that model performance relations among key value-chain activities and valued outcomes [e.g., the balanced scorecard of Kaplan and Norton, 1996].

Consulting reports, normative studies, and descriptive theories predict that these comprehensive models lead to superior performance. Magretta [2002] argues that business models are essential to tying insights to financial results. Furthermore, knowledge-based and systems theories of the firm hypothesize that superior performance results from systemic management policies, rather than myopic focus on elements of the value chain [e.g., Huff and Jenkins, 2003; Morecroft, 2002; Sanchez and Heene, 1996]. Empirical evidence supporting normative claims or theoretical hypotheses is scant and usually is uncritical self-reports [e.g., Rucci et al., 1998; Barabba et al. 2002].

Systematic management requires a comprehensive management control system, but not all of a management control system need be measurable. However, the portion that is feasibly measured should be considered for the PMM; otherwise the organization might lose valuable performance information. The choice of performance measures is critical in reflecting the organization as a system. Since an organization is always adapting to its environment, it must be able to change its performance measures to reflect current conditions. This study describes the determinants of a particular PMM and investigates the relatively unexplored issue of choice and change of a functioning PMM.

### ***Prior Work on PMM***

The DuPont ROI formula is an early and enduring PMM that disaggregates financial performance into manageable elements [e.g., Zimmerman, 1997; p. 187]. EVA is a similar, more current and complex approach to identifying the incremental contribution to shareholder wealth and the manageable elements of periodic income [e.g., Adimando et al, 1994]. Rappaport's [1999] approach to building shareholder value recognizes incentive effects of over-reliance on periodic financial results and seeks to mitigate disincentives. Because all of these models focus primarily on financial

outcomes, they do not qualify as systems models; that is, they do not model the determinants of financial performance even within the boundaries of the firm.

More comprehensive PMM include Otley's [1999] performance management model, Ittner and Larcker's [2001] value-based management model, Epstein et al.'s [2000] APL model, Kanji's business scorecard [Kanji and Moura e Sa, [2002] and the balanced scorecard (BSC) [Kaplan and Norton, 1996, 2001]. These models describe links among business decisions and outcomes, and serve to guide strategy development, communication, implementation, and feedback at multiple points along the value chain. Because these comprehensive PMM are business models, reflecting inputs and both intermediate and final outputs, they generally include measures of operational, strategic, financial and non-financial performance. These models do represent efforts to use organizational knowledge to model the firm as a system and implement management control. This study investigates whether management control, knowledge-based and systems-based theories of the firm are descriptive in a particular case. Although a study such as this cannot generalize to the population of firms using PMM, the findings of this study can illustrate the theories applied to this investigation [Yin, 1994] and can provide a foundation for theory improvement, replications, and large sample tests.

This study addresses two research questions:

- Which criteria does a firm use to choose initial PMM weights and measures?
- Why does a firm change PMM weights and measures?

The questions are investigated both qualitatively and quantitatively using (1) archival documents that describe the company's distribution PMM and (2) interviews with top managers and distributors to understand the nature of the business and the objectives and dynamic structure of the PMM. Finally, the paper reflects on the implications of this study for performance measurement and management control. Principal findings are that measures were chosen consistently with emerging theory but some changes to the PMM were inconsistent with theory. Subsequent sections of this paper address theories supporting predicted answers to the research questions, the research site, analysis of qualitative data, discussion of results, and conclusions.

## **2. Management Control Theory, Knowledge-Based and Systems-Based Theories and PMM**

Management control theory argues that management control systems (MCS), which include PMM, are intended to insure that employees (1) know what is expected of them, (2) will exert effort to do what is expected, (3) are capable of doing what is expected, and (4) accomplish what is expected [e.g., Merchant, 1998]. For more than 30 years, researchers have known that firms choose a portfolio of controls and performance measures [e.g., Khandwalla, 1972]. However, subsequent research on

firms' choices of performance measures often has focused on broad dichotomies of measures, such as financial vs. non-financial measures and mechanistic vs. organic controls. The theory commonly used in that research likewise characterizes the contingencies affecting choices of measures and controls as broad dichotomies (e.g., high vs. low environmental uncertainty; old vs. new technology).<sup>1</sup>

One particularly popular research stream predicts that firms operating in complex and risky environments rely heavily on qualitative controls and non-financial performance measures and to a much lesser degree (if at all) on quantitative, financial-performance measures. Contingency research on choice of performance measures has yielded mixed results, perhaps because most of the reported studies are based on cross-sectional survey data, which can obscure the idiosyncrasies of firm-level definitions and implementations of performance measurements [e.g., Anderson and Young, 1999; Chenhall, 2003; Luft and Shields, 2002b]. Enough evidence exists, however, to suggest that most firms rely to some degree on financial performance measures and many use both quantitative and qualitative controls. In other words, firms evidently have great flexibility to choose the portfolio of measures and controls (especially when characterized as broad dichotomies) that they expect to work best in their situations. This equivocal result provides some motivation to search for additional theoretical explanation for the choice of performance measures.<sup>2</sup>

### ***Strategy and PMM***

Theories that explain management policies based on strategic resources, capabilities, learning, and systems offer guidance and predictions for the choice of performance measures. Recent strategic management literature has evolved the concept of a firm's product strategy beyond Porter's [1985] depiction of managing the value chain to achieve competitive advantage (e.g., through product cost leadership or differentiation). Porter's work on the importance of strategic positioning has greatly influenced later work that seeks to explain how firms might use their resources to attain strategic positioning. Research that followed Porter explains how firms reach and maintain the positions of strategic advantage that he described.

Barney [1991] argues that successful firms achieve competitive advantage by acquiring and using unique resources to build inimitable capabilities that create strategic advantages [see also Amit and Schoemaker, 1993 and Kogut and Kulatilaka, 2001]. Organizational learning theory by Nonaka and Takeuchi [1995] and Senge [1990] predicts that successful firms create strategic advantages by learning dynamically to use their resources effectively. This learning is realized through development and deployment of the firm's capabilities, processes, or competencies to use resources [e.g., Prahalad and Hamel, 1990]. Morecroft et al. [2002] hypothesize that successful firms manage strategic resources and capabilities through holistic management systems; that is, creating **and** maintaining strategic advantages are enhanced by systemic management. Our accounting interpretation of current

management theories is that firms create and maintain strategic advantages or positions by efficiently creating, deploying, and using performance-based management control systems. Furthermore, the measurable part of the management control system should itself be systemic, in the form of a PMM.

### ***Management Control and PMM***

Recent management control research addresses specific factors that might explain firms' choices of performance measures to achieve and maintain strategic advantages. Laboratory experiments [e.g., Libby et al, 2002; Lipe and Salterio, 2000; Luft and Shields, 2001, 2002a] and surveys of management control practice [e.g., Ittner et al. 2002; Cavaluzzo and Ittner, 2002; Ittner and Larcker, 1998] have identified attributes of performance measures that are associated with use, usefulness, and performance. When combined with current resource-based and systems-based strategy theories, what emerges is a focus on performance measures' attributes that supercede the popular financial vs. non-financial dichotomy. In all cases the literature cited in the following subsections presumes that the organization seeks to improve performance relative to its strategic goals. Implications of these complementary management theories for choices of PMM are considered as follows.

**Measures are diverse and complementary.** Firms' management controls can benefit from greater diversity of performance measures (i.e., operational, strategic, financial, and non-financial measures) if operational measures reflect the current drivers of future financial performance and are early in the value chain [Ittner and Larcker, 2002]. Milgrom & Roberts [1995] argue that, if a diverse set of performance measures is a complete and complementary set (or system), using a subset of measures leads to inferior performance. From a similar systems perspective, Warren [2002] argues that successful management policies (e.g., PMM) reflect resource interdependence, complementarity, and temporal causality between resources and outcomes of uses (discussed in more detail later).

**Measures are objective and accurate.** Ijiri [1967] long ago re-established the theoretical importance of (accounting) performance measure accuracy and objectivity. This topic has not lost relevance.<sup>3</sup> More recently, Libby et al. [2002] find that experimental subjects in management-control tasks rely on performance measures that have been verified by third parties, which might create demand for accurate and objective measures. Other studies have found that low-quality measurement is associated with low management control system use or impact [Cavaluzzo and Ittner, 2002; Ittner and Larcker, 1998]. However, it is unclear ex ante if investing in measurements is superior to measuring the wrong things or the right things poorly, or avoiding unreliable measures altogether [e.g., Cavaluzzo and Ittner, 2002; Gates, 1999]. Objectivity (or verifiability) and accuracy (or error free) are theoretically independent concepts yet are often coincident in practice with reference to performance measurement.

**Measures are informative** Performance measures that differentiate managers facing similar, uncontrollable factors are informative. Informative measures can improve evaluations, even if they

are not completely controllable by managers [e.g., Antle and Demski, 1988]. In particular, early value-chain measures can be valuable if they are informative about managers' leading actions [Ittner and Larcker, 2001] in sufficient time to take corrective control actions.

**Measures communicate strategy.** Models such as PMM facilitate communication, learning, and creation of new knowledge and can be the key tool to building a learning organization [Huff and Jenkins, 2003]. The right performance measures align actions and strategy by reducing managers' financial myopia [McKenzie and Schilling, 1998], and effectively communicate strategy [Kaplan and Norton, 2001; Malina and Selto, 2001]. Systemic management understands and exploits knowledge of dynamic interrelations among resources and capabilities. The elements of a PMM are intended to reflect the strategic use of resources and deployment of efficient processes [e.g., Sanchez et al., 2002].

**Measures create incentives for improvement.** Using performance measures that capture inherent time delays between certain decisions (e.g., investing in R&D and employee development) can lead to improved incentives [e.g., Rappaport, 1999; Cloutier and Boehlje, 2002]. Ittner and Larcker [2001] also observe that operational measures, which have good "line of sight," can increase the expectancy of rewards based on those measures [e.g., Green, 1992].

**Measures improve decision-making.** Organization of measures into distinct categories can affect decision making, perhaps by reflecting the structure of knowledge about the firm's value chain [Lipe and Salterio, 2000]. Measures with tangible connections to processes being managed also might activate more knowledge and promote better learning and decision making compared to relying on financial measures alone [Luft and Shields, 2001, 2002a]. Huff and Jenkins [2002] argue that models (e.g., PMM) organize and express the rationale of complex systems, which aid planning and evaluation activities. Furthermore, such models can represent micro- or macro-levels of knowledge of activities, processes, and systems, thus aiding individuals at all levels of the organization. PMM might improve decision making by identifying actions and impacts that heretofore have been hidden by traditional measurement systems [e.g., Huff and Jenkins, 2002].

**Benefits outweigh costs of collection.** Monitoring employee behavior through a PMM is a costly activity. Generating, organizing, and reporting performance information consume scarce company resources [Merchant, 1998; Simons, 2000]. As management accounting researchers have known since the early days of the field [e.g., Horngren, 1967], the perceived benefits of using performance measures should outweigh the associated costs.

**Measures reflect system causality.** Some academics and consultants have prescribed forms of causal PMM [e.g., Kaplan and Norton, 1996; Epstein et al., 2000; Kanji and Moura e Sa, 2002]. Regardless of the sources of business models, causal relations among firms' multiple performance measures often are neither specified nor measured well [Ittner et al., 2002]. Quantifying cause-and-effect

relations between actions and outcomes at key points in the value chain could help predict future effects of current actions [e.g., Eccles, 1991]. A functioning causal PMM also might free managers to focus more on strategy and evaluation issues [e.g., Kaplan and Norton, 2001] than on information processing. Furthermore, a comprehensive, causal PMM might reduce the cognitive complexity of understanding and using multiple measures of performance [Luft and Shields, 2002a]. Strategy theorists predict significant benefits from building causal models of firms' strategic resources and capabilities. Huff [1990] and Huff and Jenkins [2002] describe these models as knowledge-based, cognitive maps, which can connect and organize dispersed organizational knowledge.

### ***Predicted Attributes of Measures***

Organizations choose performance measures that reflect the use of resources and capabilities that are critical to organizational success. The options for performance measures are limitless and perhaps idiosyncratic to firms and industries; the possible contingencies and tradeoffs, therefore, also might be limitless. Nonetheless, the foregoing discussions lead to predictions of the attributes of measures that organizations choose for (or, alternatively, delete from) PMM. This study hypothesizes that an organization constructs PMM based on performance measure attributes. We recognize that organizations make tradeoffs among these attributes, but theory does not support tradeoff predictions at this time. We later provide descriptive evidence about tradeoffs made in this case. Formally, the hypothesis and testable predictions are as follows:

**Hypothesis**: Organizations choose performance measures based on measure attributes.

- P1<sub>0</sub>: Organizations choose performance measures that are diverse and complementary.
- P2<sub>0</sub>: Organizations choose performance measures that are objective and accurate.
- P3<sub>0</sub>: Organizations choose performance measures that are informative.
- P4<sub>0</sub>: Organizations choose performance measures that are strategic communication devices.
- P5<sub>0</sub>: Organizations choose performance measures that are incentives for improvement.
- P6<sub>0</sub>: Organizations choose performance measures that support improved decisions.
- P7<sub>0</sub>: Organizations choose performance measures that are more beneficial than costly.
- P8<sub>0</sub>: Organizations choose performance measures that are causally related.

Similarly, this study predicts that an organization will delete previously chosen PMM measures, which after experience, are perceived to not exhibit these attributes. Because it is possible that a measure can be useful even if it does not exhibit all attributes, this study explores whether an organization trades-off some attributes.

### **3. Research Site**

The research takes place within a single firm to benefit from close investigation of (1) a performance model developed by the firm, not imposed by the researchers (or external consultants), (2) measures



of performance relevant to that firm, not generic measures that might or might not apply to the firm, and (3) access to multiple levels of managers to enrich the understanding of the origins and uses of the PMM. Thus, this study offers some advantages compared to cross-sectional analysis at the firm level, particularly given the difficulty of comparably describing performance measurement in many firms simultaneously [e.g., Luft and Shields, 2002b].

This study focuses on a PMM developed by a U.S. *FORTUNE* 500 equipment manufacturer for its distribution channel.<sup>4</sup> The company employs over 25,000 people in its domestic and foreign operations. The company's competition is from similarly large domestic and international firms. The primary bases of competition for these firms are price, quality, customer service, and speed of delivery of both goods and services. The company sells its major equipment assemblies through OEM contracts and independent, exclusive distributors, who also sell repair parts and maintenance services. The distribution system is the company's primary contact with retail customers. Each distributor operates within an assigned geographic area. The distributors may not compete with each other or sell competitors' products or services. While the distributorships are independent entities, most are owned by individuals with prior company experience.

Historically, the company has managed top-down and strictly by the "bottom-line." Its primary distribution goal was improving its share of the market for its primary products. Profitability at the distributor level was important (particularly to distributors) but less important than primary market share because most of the company's profit derived from sales of its primary products and parts. This narrow performance focus had caused considerable tension between management and some distributors who saw more profitable, regional opportunities in the company's secondary markets. This tension persisted through early versions of the PMM but was resolved, as will be explained later.

For many years, the company's success was based on the performance and cost of its major products. Because competitors now have similar products, competition for market share has shifted to customer service. In response to increased competition from both domestic and foreign firms, the company revised its strategy to focus attention and resources on improving quality, customer service, and customer satisfaction. A major step in this change was to develop a strategically oriented PMM for its 31 North American distributorships that communicates the new strategy, guides distributors, and provides additional bases for evaluation. This study will refer to the company's *distributor performance model* as the DPM.<sup>5</sup>

According to company documents, the purposes of the DPM are to:

- Highlight areas within distributorships that need improvement to enhance customer relations;
- Provide an objective set of criteria, consistent with the company's new strategic initiatives, to guide and measure total distributor performance;

- Use as the starting point for the three-year distributor contract renewal process; and
- Use for comparing and ranking distributors and may be used for performance-based compensation.

The DPM was developed internally by company employees, without the aid of external consultants. As was customary in this company, a top-down approach was used. The DPM designers created the initial DPM with selective input from distributor personnel, although the designers had the final say. The DPM was designed to focus on outcomes that the company felt were important for the distributors to accomplish in order to meet company goals, not necessarily distributor goals. Therefore, the distributor key success factors do not perfectly mirror those of the company.<sup>6</sup>

### ***Sources of Data***

The data for this study come from interviews with DPM designers and administrators and distributors, from company documents and from archival performance data. During the second quarter of 1999, sixteen interviews were conducted: nine distributor-owners and seven DPM designers and administrators. Distributors, designer and administrator views were sought in order to have a 360 degree view of the DPM process.<sup>7</sup> The length of the telephone interview was determined by how much the interviewee had to say and all lasted between 45 and 75 minutes. The researchers asked each interviewee the following open questions:<sup>8</sup>

1. In your own words, what is the DPM?
2. What do you think is the objective of the DPM?
3. What are the nine measures that distributors report really measuring?
4. What are the measures that are filled out by the company really measuring?
5. How do the measures that distributors report relate to the company's measures?
6. Do the measures help you in any way?
7. Are there any benefits from the DPM itself?
8. Do you have any (other) recommendations for improving the DPM?

Note that the questions do not directly prompt respondents to discuss the factors that are predicted to determine choices or changes in either measures or weights. The questions were purposefully generic, providing a framework for discussing DPM measures, but not directly asking about choice or change. Interviewees freely revealed these factors during the interviews. Their unprompted responses were used to support or deny the predictions.<sup>9</sup>

Company documents provided archival background for the structure of the DPM and quantitative data for each quarterly DPM from the first quarter of 1998 to the fourth quarter of 2001. The interviews were analyzed using qualitative data software (Atlas.ti). Qualitative data software is used to systematically code the qualitative data. A predetermined set of codes was used to identify

portions of interviews text referring to a choice or a change in DPM measures according to each of the eight attribute predictions developed from the literature. Although all respondents had multiple, coded comments, the coding procedure gave each respondent only a single code for each intersection of performance measure and attribute. Thus, if one interviewee, for example, offered a paragraph describing the accuracy of a specific measure, this response generated a single set of codes – one code for the measure and one for the attribute. For each coded comment regarding a prediction, an additional code was attached regarding whether the measure was dropped from the DPM or remained on the DPM during the time period tested.<sup>10</sup> This restrictive approach to coding qualitative data is designed to illustrate or test theory.<sup>11</sup> A complete list of codes used in this research is shown in Table 1.

Table 1 Coding Scheme
--------------------------

#### 4. Data Analysis

Section 4 is divided into two sections. The first section uses qualitative analysis to analyze evidence regarding the initial choice of performance measures and weightings of those measures. The second section uses both qualitative and quantitative analysis to evaluate interview evidence related to the eight attribute predictions regarding changes of performance measures.

Throughout the four years covered by this research, the company made many changes to the DPM. Weightings changed, measures were added, and measures were dropped. The initial DPM, implemented in the first quarter of 1998, and the revised DPM are shown in table 2. Table 3 contains DPM measure definitions. The most obvious change in the revised DPM is a reduction in the number of measures included. The first DPM contained 29 measures while the most recent has 14. Also, all three of the people measures (performance reviews, industry involvement, and training) were dropped from the DPM. Of the fifteen measures dropped from the initial DPM, fourteen were eliminated at the end of 1999.

Table 2 Comparison of DPMs Table 3 DPM Measure Definitions
---

Other obvious changes in the DPM are the weights assigned to measures. In the first DPM, weightings ranged from 1.0 to 27 points. As the company shortened the DPM, remaining measures

ultimately were given either a 5- or 10-point weight. Table 3 chronicles the revisions in the weights assigned to the measures dropped from the DPM over the study period.

Table 4 Revised Measure Weights
------------------------------------

The following subsections discuss evidence from the interviews that indicates support or lack of support for predicted reasons behind the company's initial choices of all measures, weights, and changes to them.

***Research Question 1: Choice of Initial Performance Measures and Weights***

The measures chosen for the initial DPM reflect the five strategic initiatives of the company. The five initiatives, which were obtained from company documents, are:

1. Demonstrate a comparative advantage in each of our markets worldwide, measured in product performance, economic value to the customer and all aspects of customer support,
2. Achieve an average return on equity of at least x% over economic cycles in order to afford the investment required to sustain a comparative advantage in each market into the future, and afford investment in new business,
3. Grow in order to provide superior total return to our shareholders over time,
4. Demonstrate our commitment to help improve the community in which we operate and be a responsible citizen of society, and
5. Attract, train, challenge and fully utilize people at all levels in order to achieve these objectives.

***Choice of measures.*** Interviewees made eleven comments regarding how and why the initial measures appeared on the DPM. The most cited reason (5 comments, or almost half of total comments) for including the initial measures was that they are objective and accurate. A company employee stated:

*The objective of the DPM is to have an objective, documented, factual measurement system rather than a subjective one.*

Several interviewees commented that the initial measures were chosen for strategic communication and causality. Three comments were made that the initial measures were designed to communicate the corporate strategy down to the distributor level. A DPM designer stated:

*The measures were chosen to dovetail the core objectives.*

The measures were also chosen with cause and effect relations in mind. Three comments were made that, by design, the initial DPM measures were intuitively related to each other. A DPM designer commented:

*The theory of how the measures tie together is good. There are logical correlations.*

The coded interviews show objectivity and accuracy, strategic communication devices and causality as the primary bases for initial choices of DPM measures.

**Choice of weights.** A total weight of 100 points is allocated across the individual measures on the DPM. The relative weights reflect the company's view of the most important areas of distributor performance. Seven comments were coded with regard to choice of initial weighting. Although designers felt that all measures were important to good management, some measures should receive the most attention. For example, consider the traditional market share weighting. Designers put the 27% weight on traditional market share because the company wanted to communicate to the distributors that the measure was of highest importance to the company. The importance was meant to be obvious.

*It [the early DPM] is intended to score poorly if [traditional] market share is low but others are still good. If the distributors are not good [traditional] market share people then they should be starving. The company will benefit by driving behavior toward higher weights [relative weights assigned to performance measures].*

Interviewees mentioned three times that initial weights were based on the most important aspects of the company's strategy.

Another factor in determining initial weights was the accuracy and objectivity of the measures (4 comments). DPM designers aimed to create a model with high-quality measures but recognize that some measures are more credible than others. Anticipated high quality measures, such as service cycle time, received high weights. Low quality measures, such as new market shares, received low weights even if they might have been otherwise important.

*We have to consider the quality of the data we're getting when doing the weighting. Hardness of the numbers definitely affects the weights.*

It appears that the initial weighting of DPM measures was a function of communicating corporate strategy with consideration for perceived quality of the measure.

## **Research Question 2: Changes in Performance Weights and Measures - Qualitative Analysis**

**Changes in weights.** Experience with the DPM measures led to more or less satisfaction with the quality of measurements and induced behaviors. The evolution of this satisfaction is reflected in the trajectories of measures' weights. Consider the evolution of the traditional market share measure. The measure's weight remained at 28 points for four quarters as the company tried to force distributors to focus on the company's primary market. As disputes grew about the exclusion of other market share measures and as the company became more aware of the other market opportunities, the company dropped the weight to 18 points for a three-quarter period, then to 10 points for another five quarters. The company clung to its traditional market share until it became convinced that this measure was creating too many disputes and was actually impeding company growth. Eventually all market share measures were replaced with equally weighted sales growth measures (for reasons described below).

**Changes in measures.** The following subsection describes the results of the qualitative analyses of changes in measures summarized in table 5. The code frequencies in column 1 are meant to reassure the reader that the authors did not selectively focus on unrepresentative comments for the analyses that follow. Column 2 shows the number of respondents comprising the frequency of comment and are meant to reassure the reader that one or two interviewees did not dominate the results.

Table 5 Changes in DPM Measures Code Frequency
---

**P1: Measures are diverse and complementary.** The DPM was intended to be a broad set of measures that encompasses the activities that distributors must manage well to create a successful distributorship. The original intent was to reinforce that distributors needed to look at all aspects of the business, from hiring and training employees to winning market share. Distributors recognized the complementarity of the DPM's diverse measures. For example,

*We can't have good results without good people who are trained. We can't have good customer satisfaction without good fill rate and cycle time results. They all relate to sales growth and profitability. I see them definitely (as complementary).*

Most distributors, however, complained about the lack of diversity in original market share measures. Distributors felt that traditional market share measures were overrepresented and over-weighted on the DPM. Distributors face diverse opportunities. Some make most of their profit from the traditional, mature market, but many can earn more from the new, growing markets. Although

diverse market share measures were included on the initial DPM, only the traditional measure carried significant weight.

*They aren't measuring anything but [traditional] business.  
[sarcastically:] If they are only interested in about one-third of their  
business, then it's good.*

The measures chosen to be on the DPM were to be diverse and complementary. As noted above, this was not perceived always to be the case. Of the 18 comments coded to this prediction, 15 supported the alternative that measures are not diverse and complementary.

P2: Measures are objective and accurate. Objectivity, accuracy, reliability, and auditability appear to have been the company's and distributors' primary concerns for DPM measures. Interviewees commented on this prediction most frequently, with 91 comments in total.

The traditional market share measure dominated the original DPM in part because it was highly objective and accurate.

*We know every single [product in the traditional market] that gets  
sold to the tenth of a percentage point.*

In contrast, the new market share measures were perceived to be less accurate and less objective.

*How we would measure [new] market share is strictly information  
we would generate ourselves.*

Almost every distributor commented on certain measures being inaccurate or subject to manipulation. In general, the distributors felt that the people measures were not well defined or verifiable. Six of the nine distributors mentioned that these measures involved some guessing and that there was no rigorous audit process in place to verify the data reported. Early in the life of the DPM, the distributors manually collected and reported service cycle time, which the company dropped after a few years. Six of the nine distributors commented on the lack of accuracy in reporting this measure.

*I'm going to make that number look as good as I can without  
outright lying or cheating.*

DPM administrators also were aware that the measure might not be accurate.

*I wouldn't put a lot of stock in either the [service cycle time  
measures] we got before or the ones on this scorecard.*

Nearly every measure dropped from the DPM lacked perceived objectivity and/or accuracy. The exception was the traditional market share measure which was replaced with the almost equally objective sales growth measure. The majority of the comments supported the alternative form of the prediction that measures are not objective and accurate.

P3: Measures are informative. Relatively few respondents indicated that informativeness was an important design criterion. The company uses the DPM to compare, benchmark, and rank distributorships and as a stimulus to peer communication. Each distributor receives its own report and its relative numerical ranking (e.g., 7<sup>th</sup> out of 31). To promote information exchange and competition, the names of distributors achieving top ratings are posted on the company's intranet for all distributors to see.

*Individuals can determine their performance against someone else in the corporation performing the same function so they would know how effective they were at accomplishing their goals and their tasks.*

Several distributors recognized the influence of peer pressure on their behavior.

*Anytime you publish a report and there are 31 entities being measured using the same metric, you create competition. We are competitive, so it matters what rank you are. Even if no one looks at the rank, I want to be #1.*

Qualitative evidence exists to support the importance of DPM informativeness, but not necessarily at the individual measure level. That is, the company appears to use the overall DPM score to rank distributor performance more than the scores on individual measures. Therefore, P3 cannot be supported at the individual measure level.

P4: Measures are strategic communication devices. At the time of the DPM roll-out, the company also launched a new customer-oriented strategy. The original DPM measures covered the value-chain activities that the company felt were critical to creating a distributorship that competes successfully on customer satisfaction, the keystone of the company's revised strategy. DPM designers observed:

*Those [measures and categories] are right in line with the strategic initiatives, aligned with our customer and quality objectives. We try to dovetail any plans the company has into those objectives.*

All but one comment were in support of the prediction that measures are strategic communication devices. Distributors and designers all agreed that DPM measures helped communicate corporate strategy to the distribution system.

P5: Measures are incentives for improvement. The comments regarding incentive effects of DPM measures were generally in support of P5 that measures provided incentives for improvement. Service cycle time, best practices, and people measures were all cited as creating incentives for distributors to improve performance.



*Best practice is a positive measure because it forces you to take a look at other ways to run your business.*

Comments regarding market share were split. Distributors commented that the emphasis on traditional market was discouraging to those distributors with more growth potential in the newer markets.

*If you did poorly on [the traditional market share] and well on [the new market share measures] then you're still viewed as a poor distributor.*

The ultimate control over distributor behavior has been the three-year contract review, when the company evaluates distributor performance and either renews or terminates the relationship. The DPM in total is intended to provide a constructive and evaluative structure and to be an objective basis for the contract review. Designers stated:

*The only incentive is losing the distributorship and [the DPM] is the centerpiece of the contract review.*

The stick might be more evident than any carrot, because as one top manager candidly stated,

*It [the DPM] helps when it comes time to terminate a distributor. If you've set up goals and a distributor has failed to achieve them, you can get around sales representation and franchise laws for wrongful termination. It's not the primary objective but it's a benefit.*

Nearly every comment supports the null form of P5 that measures provide incentives for improvement.

P6: Measures are supportive of improved decisions. The company chose measures for the DPM to reflect the activities and types of financial and non-financial performance believed to be effective to achieve increased company profitability. Most distributors agreed:

*The elements and the structure [of the DPM] are outstanding and they have a lot of potential to help us all improve... I grew up working for a CPA and he ingrained in me that, if you can't measure it, you can't improve it.*

Several of the measures on the initial DPM that were subsequently eliminated also were perceived to be helpful for decision making. In regard to training, one distributor mentioned:

*As a result of the measure, we more formally measure this. I like that because it lets me see which areas of the company are doing training and which aren't. I can chat with those not doing training. They get so tied up doing other work that it's hard to put time aside.*

The service cycle time measure, when it was part of the DPM, did effect decision making in the service process at most distributorships. Distributors often mentioned that it helped them to redesign processes and re-schedule work.

*I wasn't an advocate [of the service cycle time measure] at the start but now I am. It tells us to quickly figure out what's wrong so we can make an intelligent statement to the customer, so they can say go ahead or not. We have been able to flow more work through our show by getting the quick, easy stuff through the shop ... It's helping us.*

All but three comments supported DPM measures as being helpful in day-to-day managing and decision making, therefore providing support for the null form of the prediction.

P7: Benefits exceed costs. There were concerns with a number of DPM measures because they were perceived to be misleading or unreliable, and the company perceived the costs of resolving disputes about these subjective measures to be greater than the benefits derived. There were seventeen comments that the cost to compile or the time spent resolving disputes outweighed the benefits received from collecting the measure. The performance evaluation and service cycle time measures in particular were seen as consuming too many resources.

*To do [performance evaluations] four times a year for everyone in the company, there aren't enough hours in the day. They need to find something more appropriate for technicians and clerks.*

*It's just adding time and effort to report [service cycle time] to the company. If we're doing this, we're not doing things for the customer.*

There was agreement among the interviewees that the cost of some measures exceeded their benefits, providing support for the alternative form of P7.

P8: Measures are causally related. The causal nature of the DPM was intended by designers and generally perceived by distributors. The DPM measures were grouped and saliently displayed to guide decisions and recognize the company's beliefs about the relations among measures within the group. Representative comments from distributors reflect the intuition and general belief in the DPM's causal relations.

*A lot of business tends to run with financial and market share measures. Those are pretty crude handles. You have to get underneath to things like quality and cycle times, and softer things like employee development. That's where the leverage of the business is. The others are the results of what you've done.*

*My intuition is that the growth goals help drive market levels, but we don't know the quantitative consequences. Some measures are more tightly correlated, but I don't have a sense of which of these are the key ones that seem to have the most leverage compared to others.*

Ninety percent of the comments coded to P8 support the prediction that the DPM reflects causality among measures. Interestingly, the distributors and company believed the measures deleted from the DPM are causally related to other measures on the DPM. The nature of measures that were dropped is considered next.

***Research Question 2: Changes in Performance Weights and Measures - Quantitative Analysis***

If theories of performance measurement choice and retention are descriptive, comments from interviewees should tend to reflect favorable perceptions of the attributes of retained measures (the null form of the predictions), whereas comments should reflect unfavorable perceptions of the attributes of dropped measures (the alternate form of the predictions). Table 6 shows several tests of the distributions of codes, by predictions, for measures dropped from the DPM and for those retained on the DPM. Columns record the frequency of comments associated with dropped and kept measures. Rows record the frequency of comments associated with favorable (null) and unfavorable (alternate) perceptions. Chi-square tests compared the observed distribution of comments to the expected distribution, which assumes that dropping or keeping a measure is unrelated to null or alternate predictions.

Table 6 Analysis of Attribute Codes and Measure Retention
---

The chi-square test of the overall hypothesis that measures are kept or dropped because of the eight attributes indicates that the pattern is not random, but is significantly different ( $p < 0.005$ ) from the predicted distribution, conservatively generated from actual frequencies. This result indicates that measures' attributes are important considerations for dropping or keeping DPM measures. Retained measures are more likely to have favorable perceptions, and dropped measures are more likely to have unfavorable perceptions.

Eight forms of the predictions (P1<sub>a</sub>, P2<sub>0</sub>, P2<sub>a</sub>, P4<sub>0</sub>, P5<sub>0</sub>, P6<sub>0</sub>, P7<sub>a</sub> and P8<sub>0</sub>) had sufficient cell size to properly analyze using Chi-square analysis. These tests compare actual comment distributions to equal, "no effects" distributions. Significant  $p$ -values were found for seven of the eight. When measures were not diverse and complementary (P1<sub>a</sub>), not objective and accurate (P2<sub>a</sub>) or perceived benefits exceeded the perceived costs (P7<sub>a</sub>), measures were dropped from the DPM, as expected.

Also as expected, measures perceived to be causally related (P8<sub>0</sub>) were retained. Three tests of the predictions were significant in an unexpected direction. Contrary to expectations, measures that significantly communicated company strategy (P4<sub>0</sub>), provided incentives for improvement (P5<sub>0</sub>), or supported decision making (P6<sub>0</sub>) were dropped from the DPM. This suggests complex tradeoffs of attributes in decisions to retain or drop measures.

### ***Attribute Trade-offs***

To explore whether the company trades-off some attributes for others, comments contrary to expectations for dropped measures were investigated further, as shown in table 7. The first column of numbers in table 7 reproduces comment frequencies used in table 6 that drive the unexpected significant result. The company dropped measures that are perceived as strategic communication devices (P4<sub>0</sub>), incentives for improvement (P5<sub>0</sub>), and supportive of improved decisions (P6<sub>0</sub>), when straightforward predictions indicate that the measures should have been retained. To determine whether the company trades off these attributes, we investigated comments referring to other attributes within the same section of text. A qualitative database query using Atlas.ti identified co-occurrence or associations between coded sections of text.<sup>12</sup>

The columns of table 7 labeled P1<sub>a</sub> to P8<sub>a</sub> present the number of times that favorable comments in the first column are associated with negative comments supporting alternative forms of each of the other predictions. For example, comments on the people-performance measures are helpful to distributors in making decisions (i.e., support for P6<sub>0</sub>). However, the people measures were dropped from the DPM. The Atlas.ti query tool found several co-occurring comments supporting P2<sub>a</sub> that the measures are not objective or accurate. This indicates that the DPM attribute of improved decision making is subordinate to the attribute of objectivity or accuracy. By far the greatest concentrations of negative comments reflect concerns for objectivity and accuracy (P2<sub>a</sub>) – 29.1 percent of associated comments – and benefits versus costs (P7<sub>a</sub>) – 30.9 percent of associated comments. The company consistently made choices that the attributes of strategic communication (P4), incentives for improvement (P5), and support of improved decisions (P6) were less important than the attributes of objectivity and accuracy (P2) and benefits versus costs (P7).

Table 7 Interactions among Attributes of Unexpectedly Dropped Measures
--

## **5. Discussion and Conclusions**

This study focuses on the efforts of one large firm to model drivers of its distribution performance. The present study complements recent research and offers an analytic generalization to an emerging

theory of performance measurement and management control. This study contributes to analytic generalization about performance measurement by empirically examining choices and revisions of diverse performance measures in a live setting. The results of this qualitative investigation support previous survey, experimental, and normative studies that collectively are refining a theory of performance measurement.

### ***Drivers of Performance and Management Control***

Theory of performance measurement and management control identify eight desirable attributes of performance measures. Measures should be:

1. Diverse and complementary
2. Objective and accurate
3. Informative
4. Strategic communication devices
5. Incentives for improvement
6. Supportive of improved decisions
7. Benefits exceed costs
8. Causally related

Analysis of interviews with company PMM designers and users find that all of these attributes were important to initial choices and/or subsequent revisions of a PMM, thereby supporting the hypothesis that organizations choose performance measures based on measure attributes. Although all of the attributes might be important, the two attributes that appear most responsible for keeping or dropping individual DPM measures are (a) the objectivity and accuracy of measurements and (b) the cost versus benefit balance of measurement. A measure might be strategically important, for example, but the company dropped it if experience showed it could not be measured inexpensively, accurately or objectively.

This trade-off of attributes can be seen clearly in the treatment of market share measures. As mentioned previously, the company initially emphasized the importance of the traditional market share. This was due in part to corporate strategy and in part to the ability to measure market share in this market to the penny. Distributors complained that the traditional market did not capture the growth opportunities for many distributors. Unfortunately, in these new markets, market share is difficult to measure accurately. The dispute ultimately was resolved by eliminating all the market share measures from the DPM. The company decided to move the weight assigned to market share to sales growth measures. Measurement of sales growth is easily obtained, accurate and objective for both the traditional market and the new markets. Furthermore, diversity among measures was retained.

Apparently this company, with its long history of bottom-line financial management, could not tolerate subjectivity in measurement or high costs of improved information, nor could it tolerate the costly disputes that inevitably follow the use of subjective measures for performance evaluations and contracting. This finding has implications for the use of performance measures (e.g., some non-financial measures) that cannot be measured or audited as objectively or accurately as financial measures. For example, many organizations currently measure employee capabilities and some might use these measures in performance evaluations. A full understanding of the “softness” of such measures might preclude non-productive disputes that can arise if they are used to evaluate and reward performance, particularly in firms that are accustomed to using hard, financial measures. The creative task facing PMM designers is to find measures that are: diverse and complementary, informative, strategic communication devices, incentives for improvement, supportive of improved decisions, causally related, but are also objective and accurate and not too costly. Without doubt, this is a challenging task.

### ***Future Research***

Future research might replicate this study in similar or dissimilar firms or to a large sample of diverse firms to obtain useful contrasts or further support for the theory of performance measurement models.

Future research questions include:

- What are the measurement tradeoffs among measure attributes that might be necessary to implement feasible PMM? Are cost of measurement and accuracy and objectivity always trump cards?
- Do all firms or only those with strong bottom-line, top-down management (or other attributes) stress cost of measurement, accuracy and objectivity of performance measures above all other attributes?
- Is complementarity an issue beyond initial design? Are there not tradeoffs but instead returns to scale and scope of measurement?
- What are the performance opportunity costs of trading off performance measure attributes?

### ***Epilogue***

As reported in Malina and Selto [2001], three years after the introduction of the DPM distributors’ measured financial performance improved significantly. The DPM is alive and well as of November 2003, and it is used for managing and evaluating distributors and distributorships. Top management commitment to using and refining the DPM was and continues to be strong. Distributors continue to adapt their behavior to the DPM’s guidance, and the DPM has undergone several more changes since the completion of this research. An example of adaptation is that all distributors now meet the DPM’s challenging safety goal, despite many early complaints that it was an impossible standard [see Malina

and Selto, 2001]. Changes to the DPM include dropping formal reporting of profit (PBIT/S), because of unforeseen claims by labor unions for a greater share of profits, and inclusion of several new measures that had been proposed earlier but not yet developed. The DPM continues to be the center point of the contract renewal process, and its increasing objectivity and relevance has added to its acceptance. The DPM has become an integral part of management control at this company and undoubtedly will continue to evolve.

## References

- Adimando, C., R. Butler, S. Malley, S. A. Ravid, R. Shepro, R. H. Vernon, J. Willett and J. Zimmerman. 1994. Stern Stewart EVA Roundtable, *Journal of Applied Corporate Finance*, 7(2), 46-70.
- Amit, R. and P. Schoemaker. (1993). Strategic assets and organizational rent. *Strategic Management Journal*, 14: 33-46.
- Anderson, S.W. and S.M. Young. 1999. The impact of contextual and process factors on the evaluation of activity-based costing systems. *Accounting, Organizations and Society* 24: 525-559.
- Antle, R. and J. S. Demski. 1988 The controllability principle in responsibility accounting. *The Accounting Review* 63(4): 700-718.
- Barabba, V., C. Huber, F. Cooke, N. Pudar, J. Smith, M. Paich. 2002. A multimethod approach for creating new business models: The General Motors OnStar project. *Interfaces*, 32(1): 20-34.
- Cavaluzzo, K. and C. Ittner. 2002. Implementing performance measurement innovations: Evidence from government. Georgetown University working paper.
- Cloutier, LM. And M. Boehlje. 2002. Innovation management under uncertainty: A system dynamics model of R&D investments in biotechnology. In Morecroft et al. (eds). *Systems Perspectives on Resources, Capabilities, and Management Processes*. Amsterdam: Pergamon.
- Eccles, R. 1991. The performance measurement manifesto. *Harvard Business Review* 69: 131-7
- Economics* 19: 45-48.
- Epstein, M. J., P. Kumar and R. A. Westbrook. "The Drivers Of Customer And Corporate Profitability: Modeling, Measuring, And Managing The Causal Relationships," *Advances in Management Accounting*, 2000, v9(1), 43-72.
- Gates, S. 1999. *Aligning Strategic Performance Measures and Results*. The Conference Board, New York, NY.
- Green, T. B. 1992. *Performance and Motivation Strategies for Today's Workforce: A Guide to Expectancy Theory Applications*. Westport, CT: Greenwood Publishing Group.
- Huff, A. and M. Jenkins (eds.) 2003. *Mapping Strategic Knowledge*. Thousand Oaks, CA: Sage Publications.
- Horngrén, C. T. 1967. *Cost Accounting: A Managerial Emphasis*. 2d ed., Englewood Cliffs, N.J.: Prentice-Hall.
- Ijiri, Y. 1967. *The Foundations of Accounting Measurement*. Houston, TX: Scholars Book Co.



- Ittner, C.D. and D.F. Larcker. 1996. Measuring the impact of quality initiatives on firm financial performance. In *Advances in the Management of Organizational Quality*, Vol. 1. Eds. S. Ghosh and D. Fedor. Greenwich, CT: JAI, 1-37.
- Ittner, C.D. and D.F. Larcker. 2001. Assessing empirical research in managerial accounting: A value-based management perspective. *Journal of Accounting and Economics*. v32 (1-3,Dec): 349-410.
- Ittner, C.D. and D.F. Larcker. 2002. Determinants of performance measure choices in worker incentive plans. *Journal of Labor Economics* 20(2): S58.
- Ittner, C.D. and D.F. Larcker. 1998. Are non-financial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research* 36 (supplement): 1-35.
- Ittner, C.D., D.F. Larcker and T. Randall. 2002. Performance implications of strategic performance measurement in financial services firms. Wharton School working paper.
- Kanji, G. and P. Moura e Sa. 2002. Kanji's business scorecard. *Total Quality Management*. 13(1): 13-27.
- Kaplan, R.S. and D.P. Norton. 2001. *The Strategy-Focused Organization*. Boston, MA: Harvard Business School Press.
- Kaplan, R.S. and D.P. Norton. 1996. *The Balanced Scorecard*. Boston, MA: MA: Harvard Business School Press.
- Kogut, B. and N. Kulatilaka. 2001. Capabilities as real options. *Organization Science*, (12):6: 744-58.
- Libby, T., S. Salaterio, and A. Webb. 2002. The balanced scorecard: The effects of assurance and process accountability on managerial judgment. Wilfred Laurier University working paper.
- Lipe, M. and S. Salterio. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review* 75: 283-98.
- Lipe, M. and S. Salterio. 2002. A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society* 27:531-40.
- Luft, J. and M. Shields. 2001. Why does fixation persist? Experimental evidence on the judgment performance effects of expensing intangibles. *The Accounting Review*. 76(4): 561-87.
- Luft, J. and M. Shields. 2002a. Learning the drivers of financial performance: Judgment and decision effects of financial measures, nonfinancial measures, and statistical models. Michigan State University working paper.
- Luft, J. and M. Shields. 2002b. Mapping management accounting: Graphics and guidelines for theory-consistent empirical research. Forthcoming *Accounting, Organizations and Society*.
- Magretta, J. 2002. Why business models matter. *Harvard Business Review*, 80(5).

- Malina, M.A. and F.H. Selto. 2001. Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard. *Journal of Management Accounting Research* 13: 48 - 90.
- McKenzie, F.C. and M.D. Shilling. 1998. Avoiding performance measurement traps: Ensuring effective incentive designs and implementation. *Compensation and Benefits Review* 30(4): 57-65.
- Merchant, K. 1998. *Modern Management Control Systems: Text and Cases*. Upper Saddle River, NJ: Prentice Hall.
- Miles, M.B. and A.M. Huberman. 1994. *Qualitative data analysis*. Thousand Oaks, CA: Sage Publications.
- Morecroft, J. 2002. Resource management under dynamic complexity. In Morecroft et al. (eds). *Systems Perspectives on Resources, Capabilities, and Management Processes*. Amsterdam: Pergamon.
- Morecroft, J., R. Sanchez, and A. Heene (eds). 2002. *Systems Perspectives on Resources, Capabilities, and Management Processes*. Amsterdam: Pergamon.
- Nonaka, I. and H. Takeuchi. (1995). *The Knowledge-Creating Company*. New York: Oxford University Press.
- Norreklit, H. 2000. The balance on the balanced scorecard – A critical analysis of some of its assumptions. *Management Accounting Research* 11: 65-88.
- Otley, D.T. 1999. Performance management: A framework for management control systems research. *Management Accounting Research*: 363-382.
- Prahalad, C. and G. Hamel, 1990. The core competence of the corporation. *Harvard Business Review*, 90(3): 79-91.
- Rappaport, A. 1999. New thinking on how to link executive pay to performance. *Harvard Business Review*. Mar-Apr: 91-101.
- Rucci, A.J., S.P. Kirn and R.T. Quinn. 1998. The employee-customer-profit chain at Sears. *Harvard Business Review* 76(1): 82-97.
- Sanchez, R., A. Heene and H. Thomas (eds.). 1996. *Dynamics of Competence-Based Competition*. Oxford: Pergamon.
- Senge, P. 1990. *The Fifth Discipline: The Art and Practice of the Learning Organization*. New York: Doubleday.
- Simons, R. 2000. *Performance Measurement & Control Systems for Implementing Strategy: Text & Cases*. Upper Saddle River, NJ: Prentice Hall.

- Van de Ven, A., and Drazin. 1985. The concept of fit in contingency theory. *Research in Organizational Behavior* 7: 333-65.
- Warren, K. 2002. Operationalizing the impact of competence building on the performance of firms' resource systems. In Morecroft et al. (eds). 2002. *Systems Perspectives on Resources, Capabilities, and Management Processes*. Amsterdam: Pergamon.
- Yin, R.K. 1994. *Case Study Research Design and Methods*. Newbury Park, CA: Sage Publications, Inc.
- Zimmerman, J. 1997. *Accounting for Decision Making and Control*. Burr Ridge, IL: Irwin-McGraw-Hill.

---

**Table 1**  
**Coding Scheme**

---

Code	
1.1	P1 <sub>0</sub> : Diverse and complementary
1.2	P1 <sub>a</sub> : Not diverse and complementary
2.1	P2 <sub>0</sub> : Objective and accurate
2.2	P2 <sub>a</sub> : Not objective and accurate
3.1	P3 <sub>0</sub> : Informative
3.2	P3 <sub>a</sub> : Not informative
4.1	P4 <sub>0</sub> : Strategic communication devices
4.2	P4 <sub>a</sub> : Not strategic communication devices
5.1	P5 <sub>0</sub> : Incentives for improvement
5.2	P5 <sub>a</sub> : No incentives for improvement
6.1	P6 <sub>0</sub> : Supportive of improved decisions
6.2	P6 <sub>a</sub> : Not supportive of improved decisions
7.1	P7 <sub>0</sub> : More beneficial than costly
7.2	P7 <sub>a</sub> : Not more beneficial than costly
8.1	P8 <sub>0</sub> : Causally related
8.2	P8 <sub>a</sub> : Not causally related
9.1	Measure dropped from DPM
9.2	Measure kept on DPM

---

Table 2  
Comparison of Distributor Performance Models  
Q1 1998 to Q4 2001

Q1 1998		Q4 2001	
Measure	Weight	Measure	Weight
Customer Satisfaction	3.0	Customer Satisfaction	10
Market Share – Traditional Market	27.0		
Market Share – New Market 1	3.0		
Market Share – New Market 2	3.0		
Dealer Survey	2.0	Dealer Survey	10
Parts Fill Rate	3.0	Parts Fill Rate	5
Service Cycle Time	10.0		
CME	3.0		
Best Practices	1.0		
PBIT as % of Sales	4.0	PBIT as % of Sales	10
Cash Flow as % of Sales	2.0		
Parts Inventory Turns	2.0	Parts Inventory Turns	5
Whole Goods Inventory Turns	2.0	Whole Goods Inventory Turns	5
Days Sales Outstanding	2.0	Days Sales Outstanding	5
Service Utilization	2.0	Service Utilization	5
Parts Sales Growth (traditional)	2.0	Parts Sales Growth (traditional)	5
Service Sales Growth (traditional)	2.0	Service Sales Growth (traditional)	5
New Market 1 Sales Growth	2.0	New Market 1 Sales Growth	10
New Market 2 Sales Growth	1.0	New Market 2 Sales Growth	10
Other Sales Growth	2.0		
Environmental Assessment	2.0		
Safety	2.0	Safety	5
Performance Reviews	1.0		
Industry Involvement	1.0		
Training	2.0		
Warranty Audit	8.0		
Building Condition	3.0		
Policy Dollars	2.0		
Days to Fin. Statement Submission	1.0		
		Image	10
DPM Total Weight	100	DPM Total Weight	100

**Table 3**  
**DPM Measure Definitions**

<b>Measure</b>	<b>Definition</b>
Customer Satisfaction	Score on customer satisfaction event card
Market Share Measures	
Market Share – Traditional Market	Monthly percentage of distributor share of total traditional market
Market Share – New Market 1	Monthly percentage of distributor share of total new market 1
Market Share – New Market 2	Monthly percentage of distributor share of total new market 2
Dealer Survey	Annual survey of dealer satisfaction with distribution system
Parts Fill Rate	Percentage of parts orders filled within 24 hours
Service Cycle Time	Percent of service repairs diagnosed within one hour and completed within six hours
CME	Audit of distributor use of Company Marketing Excellence system
Best Practices	Number of best practices submitted or number of best practices implemented from data base in the last 12 month period
PBIT as % of Sales	PBIT as a percentage of sales
Cash Flow as % of Sales	Cash flow from operations as a percentage of sales
Parts Inventory Turns	Parts cost of sales divided by average parts inventory
Whole Goods Inventory Turns	Whole goods cost of sales divided by average parts inventory
Days Sales Outstanding	Average trade receivable balance divided by sales per day
Service Utilization	Technician hours billed divided by hours available to be billed
Parts Sales Growth	12 quarter rolling regression over the most recent 12 quarters parts sales growth
Service Sales Growth	12 quarter rolling regression over the most recent 12 quarters non-warranty sales growth

---

Table 3 (continued)

New Market 1 Sales Growth	12 quarter rolling regression over the most recent 12 quarters new market 1 sales growth
New Market 2 Sales Growth	12 quarter rolling regression over the most recent 12 quarters new market 2 sales growth
Other Sales Growth	12 quarter rolling regression over the most recent 12 quarters other sales growth
Environmental Assessment	If distributor has completed an environmental assessment per Company guidelines.
Safety	Lost-time accidents per 200,000 hours worked
People Measures	
Performance Reviews	Random sample of percent of employees evaluated annually
Industry Involvement	Evidence of membership in industry and trade associations
Training	Percent of hours spent on training per hours worked
Warranty Audit	Percent of compliance of warranty repairs
Building Condition	Company determined rating of distributor properties
Policy Dollars	Comparison of actual versus planned expenditure on non-warranty repairs
Days to Fin. Statement Submission	Number of days from close of accounting period to submission of financial statement to Company
Image	(Not implemented at the time of this research)

---

**Table 4**  
**Revised Performance Weights of Dropped Measures**

	Q198	Q298	Q398	Q498	Q199	Q299	Q399	Q499	Q100	Q200	Q300	Q400	Q101	Q201	Q301	Q401
Market Share Measures																
Market Share – Trad. Mkt	27	28	28	28	28	18	18	18	10	10	10	10	10			
Market Share – New Mkt 1	3	3	3	3	3	7	7	7								
Market Share – New Mkt 2	3	3	3	3	3	7	7	7								
Service cycle time	10	10	10	10	10	10	10	10								
CME	3	3	3	3	3	3	3	3								
Best Practices	1	1	1	1	1	3	3	3								
Cash Flow	2	2	2	2	2	2	2	2								
Environmental Assessment	2	2	2	2	2	2	2	2								
People Measures																
Performance Reviews	1	1	1	1	1	1	1	1								
Industry Involvement	1	1	1	1	1	1	1	1								
Training	2	2	2	2	2	2	2	2								
Warranty Audit	8	8	8	8	8	8	8	8								
Building Condition	3	3	3	3	3	3	3	3								
Policy Dollars	2	2	2	2	2	2	2	2								
Days to Fin. Stmt Sub.	1	1	1	1	1	1	1	1								



**Table 5**  
**Changes in DPM Measures**  
**Code Frequency**

	Total	Number of Respondents
P1 <sub>0</sub> : Diverse and complementary	3	3
P1 <sub>a</sub> : Not Diverse and complementary	<u>15</u>	8
Total	<u>18</u>	
P2 <sub>0</sub> : Objective and accurate	26	13
P2 <sub>a</sub> : Not Objective and accurate	<u>65</u>	13
Total	<u>91</u>	
P3 <sub>0</sub> : Informative	4	4
P3 <sub>a</sub> : Not Informative	<u>0</u>	0
Total	<u>4</u>	
P4 <sub>0</sub> : Strategic communication devices	23	13
P4 <sub>a</sub> : Not Strategic communication devices	<u>1</u>	1
Total	<u>24</u>	
P5 <sub>0</sub> : Incentives for improvement	24	14
P5 <sub>a</sub> : No Incentives for improvement	<u>5</u>	5
Total	<u>29</u>	
P6 <sub>0</sub> : Supportive of improved decisions	27	9
P6 <sub>a</sub> : Not Supportive of improved decisions	<u>3</u>	3
Total	<u>30</u>	
P7 <sub>0</sub> : Benefits exceed costs	5	3
P7 <sub>a</sub> : Costs exceed benefits	<u>17</u>	9
Total	<u>22</u>	
P8 <sub>0</sub> : Causally related	36	13
P8 <sub>a</sub> : Not Causally related	<u>4</u>	4
Total	<u>40</u>	
<b>Total Frequency</b>	<b>258</b>	

**Table 6**  
**Analysis of Attribute Codes and Measure Retention**

	Actual Distribution			No Effect Distribution		Chi-square	p-value <
	Measure Dropped	Measure Retained		Measure Dropped	Measure Retained		
Overall actual (expected) frequencies	87(98.09)	61(49.91)	148			8.73	<b>0.005</b>
	<u>84(72.91)</u>	<u>26(37.09)</u>	<u>110</u>				
	171	87	258				
P1 <sub>0</sub> : Diverse and complementary	0	3		1.5	1.5	n/a*	
P1 <sub>a</sub> : Not Diverse and complementary	11	4		7.5	7.5	2.40	<b>0.100</b>
P2 <sub>0</sub> : Objective and accurate	14	12		13	13	0.04	0.900
P2 <sub>a</sub> : Not Objective and accurate	46	19		32.5	32.5	10.40	<b>0.001</b>
P3 <sub>0</sub> : Informative	2	2		2	2	n/a	
P3 <sub>a</sub> : Not Informative	0	0		0	0	n/a	
P4 <sub>0</sub> : Strategic communication devices	17	6		11.5	11.5	4.35	<b>0.025</b>
P4 <sub>a</sub> : Not Strategic communication devices	1	0		0.5	0.5	n/a	
P5 <sub>0</sub> : Incentives for improvement	18	6		12	12	5.04	<b>0.025</b>
P5 <sub>a</sub> : No Incentives for improvement	2	3		2.5	2.5	n/a	
P6 <sub>0</sub> : Supportive of improved decisions	20	7		13.5	13.5	5.33	<b>0.025</b>
P6 <sub>a</sub> : Not Supportive of improved decisions	3	0		1.5	1.5	n/a	
P7 <sub>0</sub> : Benefits exceed costs	3	2		2.5	2.5	n/a	
P7 <sub>a</sub> : Costs exceed benefits	17	0		8.5	8.5	15.06	<b>0.001</b>
P8 <sub>0</sub> : Causally related	13	23		18	18	2.25	<b>0.100</b>
P8 <sub>a</sub> : Not Causally related	4	0		2	2	n/a	

\* No Effect (expected) distribution cells must contain at least 5 observations to use Chi-square analysis properly.

**Table 7**

**Interactions Among Attributes of Unexpectedly Dropped Measures**

Predictions Significant in Unexpected Direction		Frequency of Associated Negative Attribute Comments								
Measures are:	Favorable Comments on Dropped Measures	P1 <sub>a</sub>	P2 <sub>a</sub>	P3 <sub>a</sub>	P4 <sub>a</sub>	P5 <sub>a</sub>	P6 <sub>a</sub>	P7 <sub>a</sub>	P8 <sub>a</sub>	None
P4 <sub>0</sub> : Strategic communication devices	17	1	4	0	-	0	1	2	0	9
P5 <sub>0</sub> : Incentives for improvement	18	0	6	0	0	-	0	6	1	5
P6 <sub>0</sub> : Supportive of improved decisions	20	0	6	0	0	0	-	9	0	4
Total	55	1	16	0	1	0	1	17	1	18
Proportions of total comments	1.000	.018	<b>.291</b>	0	.018	0	.018	<b>.309</b>	.018	.364

P1<sub>a</sub>: Not diverse and complementary

P2<sub>a</sub>: Not objective and accurate

P3<sub>a</sub>: Not informative

P4<sub>a</sub>: Not strategic communication devices

P5<sub>a</sub>: Not incentives for improvement

P6<sub>a</sub>: Not supportive of improved decisions

P7<sub>a</sub>: Not more beneficial than costly

P8<sub>a</sub>: Not causally related

## Endnotes

- 
- <sup>1</sup> See Chenhall [2003] for a comprehensive review of management control research and contingency theory.
- <sup>2</sup> Van de Ven and Drazin [1985] discuss the concept of equifinality as a serious impediment to progress in understanding firm behavior via contingency theory. We agree with the concept of contingency theory that management control systems and, hence, PMM are idiosyncratic but not random. We do fear that equifinality practically nullifies tests of contingency theory because each firm, at a particular time, is its own standard.
- <sup>3</sup> Discussions of many current accounting and performance measurement issues can be traced back to Ijiri's classic work.
- <sup>4</sup> The company has created PMM for other elements of its value chain, but gave access to only the distribution PMM. Other PMMs reportedly are being used successfully.
- <sup>5</sup> Company employees refer to the DPM as a "balanced scorecard," but the term has acquired a generic label that might obscure the unique characteristics of this PMM.
- <sup>6</sup> For a discussion of the conflict resulting from this initial mismatch, see Malina and Selto [2001].
- <sup>7</sup> For a complete discussion of the sampling technique, see Malina and Selto [2001].
- <sup>8</sup> These interviews also were used to explore communication effectiveness in Malina and Selto [2001]. The present study reflects an independent use and extension of these qualitative data.
- <sup>9</sup> Perhaps respondents would have made more comments specific to the hypothesized attributes if we had posed direct questions. However, we wanted to avoid responses that were artifacts of leading questions. We do not have the luxury of a parallel case study where we could have asked leading questions to compare with our results.
- <sup>10</sup> Two researchers independently coded the qualitative data. The average coding reliability between the two coders was 87%, which falls above the minimum norm of 80 percent coding reliability (Miles and Huberman, 1994). Coding discrepancies were reconciled by consensus. The consensus coding supports the reported qualitative analyses. For a complete discussion of insuring coding reliability, see Malina and Selto, 2001.
- <sup>11</sup> Miles and Huberman [1994] provide extensive explanations of alternative coding methods.
- <sup>12</sup> Co-occurrence or proximity rules include coded quotations of one type that *enclose, are enclosed by, overlap, are overlapped by, precede by one line, or follow by one line* coded quotations of another type. See appendix B of Malina and Selto [2001] for a complete discussion of finding associations among codes using qualitative software.