

Choice and change of measures in performance measurement models

Mary A. Malina^{a,1}, Frank H. Selto^{b,*}

^a *Naval Postgraduate School, 208 Ingersoll Hall, 555 Dyer Road, Monterey, CA 93943, USA*

^b *Leeds School of Business, University of Colorado at Boulder, Boulder, CO 80309-0419, USA*

Received 30 May 2003; accepted 24 August 2004

Abstract

This paper uses management control, resource-based, systems-based and contingency-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 four research questions: (1) Are measure attributes important considerations for performance measure choice? (2) Does the importance of attributes differ according to firm strategy? (3) Does the importance of attributes for design and use differ according to firm strategy? (4) Does a company trade-off some individual attributes for others? The questions are investigated using qualitative and quantitative analyses of archival documents and interviews with top managers and distributors. Principal findings are that measure attributes are important considerations for choice and change of performance measures, design attributes are more important than use attributes, the importance of attributes does not appear to differ according to strategy, and some individual attributes are traded-off for other attributes.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Performance measurement model; Management control; Non-financial performance measures; Strategy

* Corresponding author. Tel.: +1 303 492 1549; fax: +1 303 492 5962.

E-mail addresses: mamalina@nps.edu (M.A. Malina), frank.selto@colorado.edu (F.H. Selto).

¹ Tel.: +1 831 656 1187; fax: +1 831 656 3407.

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, 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 reflect 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-based 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 et al., 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 (MCS), but not all of a MCS 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.

1.1. Knowledge-based and systems-based theories 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; 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 MCS. Furthermore, the measurable part of the MCS should itself be systemic, in the form of a PMM.

1.2. 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 four research questions:

- RQ1:* Are measure attributes important considerations for performance measure choice?
- RQ2:* Does the importance of attributes differ according to firm strategy?
- RQ3:* Does the importance of attributes for design and use differ according to firm strategy?
- RQ4:* Does a company trade-off some individual attributes for others?

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 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. Subsequent sections of this paper address performance measure attributes, the research questions, the research site, analysis of qualitative and quantitative data, discussion of results, and conclusions.

2. Performance measure attributes

Management control theory argues that 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 versus non-financial measures and mechanistic versus 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 versus low environmental uncertainty; old versus new technology).²

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, 2003). 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.³

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, 2002) 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 versus 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.

2.1. 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 and Roberts (1995) argue that, if a diverse set of performance measures is a complete and complementary set

² See Chenhall (2003) for a comprehensive review of management control research and contingency theory.

³ 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 MCS and, hence, PMM are idiosyncratic but not random.

(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 and complementarity.

2.2. *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.⁴ 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 MCS 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.

2.3. *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.

2.4. *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.

2.5. *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

⁴ Discussions of many current accounting and performance measurement issues can be traced back to Ijiri's classic work.

reduce the cognitive complexity of understanding and using multiple measures of performance (Luft and Shields, 2002). Strategy theorists predict significant benefits from building causal models of firms' strategic resources and capabilities. Huff (1990) and Huff and Jenkins (2003) describe these models as knowledge-based, cognitive maps, which can connect and organize dispersed organizational knowledge.

2.6. 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 Shilling, 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., Morecroft et al., 2002).

2.7. 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).

2.8. 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, 2002). Huff and Jenkins (2003) 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, 2003).

Management control and strategy theories identify eight desirable attributes of performance measures. Measures should be:

- A1. Diverse and complementary.
- A2. Objective and accurate.
- A3. Informative.
- A4. More beneficial than costly.
- A5. Causally related.
- A6. Strategic communication devices.
- A7. Incentives for improvement.
- A8. Supportive of improved decisions.

3. Research questions

3.1. *RQ1—Are measure attributes important considerations for performance measure choice?*

The discussions in Section 2 provide extensive support for the prediction that an organization chooses measures for their PMM (initially and subsequently) based on at least the eight attributes summarized in this study. The initial choice of performance measures may be based on perceived expected attributes since little history exists upon which to base the assessment. However, changes may occur after experience proves the perceptions correct or incorrect. Therefore, this study predicts that an organization will delete previously chosen PMM measures, which after experience, are perceived to not exhibit these attributes. Whether for initial choice or subsequent change, the theoretical basis for attribute importance remains the same. We expect that a company chooses measures that it expects to have the eight attributes and keeps only those that display the attributes in practice.

3.2. *RQ2—Does the importance of attributes differ according to firm strategy?*

More specifically, contingency-based research has shown that firm strategy can affect the design of PMM (Chenhall, 2003). These studies typically use one of several taxonomies to describe firms' strategies, including entrepreneurial or conservative (Miller and Freisen, 1982); prospector, analyzer, or defender (Miles and Snow, 1978); build, hold, or harvest (Gupta and Govindarajan, 1984); and product differentiation or cost leadership (Porter, 1980). More conservative strategies include defender, harvest, and cost leadership, while more entrepreneurial strategies are prospector, build, and product differentiation. Guilding (1999) found that companies following a prospector or build strategy are more likely to use a broad scope of PMM information. Govindarajan and Gupta (1985) found that companies with a build strategy are more inclined to use subjective controls than those pursuing a harvest strategy. Several studies find that firms following a more conservative strategy place more emphasis on cost control than those following a more entrepreneurial strategy (Chenhall and Morris, 1995; Dent, 1990; Simons, 1987). Bouwens and Abernethy (2000) found that firms going through a strategic change process, typically categorized as an entrepreneurial activity, place more importance on integrated PMM information. Finally, Abernethy and Brownell (1999) found that hospitals following a prospector strategy focused more attention on dialogue, communication and learning. No prior strategic-fit work was found leading to a strategy-based preference for attributes of informativeness and incentives for improvement. Table 1 maps the prior strategic-fit research to the eight performance measure attributes in Table 1.

Although each of the eight attributes is desirable (*ceteris paribus*), strategic-fit research implies that some attributes may be more relevant than others depending upon company strategy. It is expected that firms following an entrepreneurial strategy will be more likely to prefer measures that are diverse and complementary (A1), causally related (A5), strategic communication devices (A6), and improve decision-making (A8). It is also expected that firms following a more conservative strategy will tend to prefer performance measures that are objective and accurate (A2) and more beneficial than costly (A4).

3.3. *RQ3—Does the importance of attributes for design and use differ according to firm strategy?*

Although all eight are desirable attributes, a company may deem some more relevant than others when designing or using measures in a PMM. The exploratory third question explores subjective classifications

Table 1
Framework for performance measure attributes

Performance measure attribute	Strategy		Study
	Entrepreneurial firm (prospector, build, product differentiation, entrepreneur)	Conservative firm (defender, harvest, cost leader, conservative)	
A1. Diverse and complementary	×		Guilding (1999) Govindarajan and Gupta (1985)
A2. Objective and accurate		×	
A3. Informative	?	?	No studies of informativeness related to strategy
A4. Benefit exceeds cost		×	
A5. Causally related	×		Chenhall and Morris (1995), Dent (1990), Simons (1987) Bouwens and Abernethy (2000)
A6. Strategic communication device	×		
A7. Incentive for improvement	?	?	Abernethy and Brownell (1999) No studies of incentives related to strategy
A8. Improved decision-making	×		

of the eight attributes as being primarily being driven by either design or use concerns. Performance measure attributes reflecting primarily design considerations are diverse and complementary (A1), objective and accurate (A2), informative (A3), more beneficial than costly (A4), and causally related (A5). Although all performance measures are at least minimally designed, we conjecture that some are selected as means to achieve specific ends. They can be *used* to communicate strategy (A6), provide incentives for improvement (A7), and improve decision-making (A8). We test whether the importance differs according to this partition of design and use attributes.

3.4. RQ4—Does a company trade-off some individual attributes for others?

Because it is possible that a measure can be useful and kept in a PMM even if it does not exhibit all attributes, we also explore whether an organization trades-off one attribute for another. The strategic-fit literature cited previously can lead one to expect trade-offs when measures' attributes are not consistently high. For example, a company following conservative strategies, given costly measurement, might (not) choose or retain a measure that is high (low) on accuracy and objectivity but low (high) on strategic communication. A more entrepreneurial firm, on the other hand, might (not) choose or retain a measure that is high (low) on improved decision-making but low (high) on the benefit–cost ratio. Although one generally might expect these trade-offs, they would be quite difficult to predict for a specific firm without having more precise attribute measurements. Thus, the fourth research question also is exploratory. Fig. 1 summarizes the four research questions.

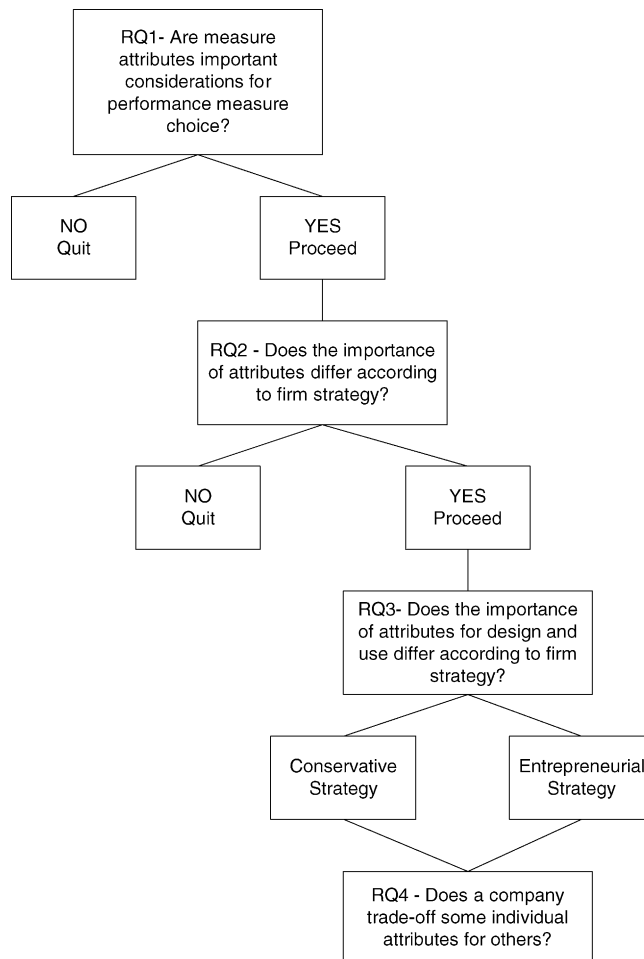


Fig. 1. Research questions.

4. Research site

The research takes place within a single firm to benefit from close investigation of (1) a PMM 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, 2003](#)). This study had access to company documents, archival data, and employees at various levels.

This study focuses on a PMM developed by a U.S. FORTUNE 500 equipment manufacturer for its distribution channel.⁵ The company employs over 25,000 people in its domestic and foreign operations. Competition is from similarly large domestic and international firms. 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.

4.1. Company strategy

Per archival documents, the company follows five strategic initiatives, which 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.
5. Attract, train, challenge and fully utilize people at all levels in order to achieve these objectives.

The company's strategy can be characterized along the lines of conservative, defender, harvest, and cost leader. Competitive advantage is gained by aggressively preventing competitors from entering the primary product market. Product performance, quality and efficiency have been the company's means to success.

Historically, the company has managed its distributors top-down and strictly by the impact on the company's "bottom-line". However, recent changes in the marketplace have affected the company's approach to measuring distributor performance. Because both domestic and foreign competitors now have similar products, competition for market share has shifted to customer service. In response to this increased competition, the company revised its distributor strategy to focus attention and resources on improving customer service and customer satisfaction.

The company's top-down approach to distributor management is consistent with the theory that firms following a conservative strategy rely more heavily on centralized control systems and on feed-forward control (Langfield-Smith, 1997). Its main distribution goal is improving the primary market share for its 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 in this market. Again, this supports the classification of the company strategy as more conservative. Conservative companies tend to use control systems that signal the need for innovation by highlighting substantial drops in market share and declining profitability (Miller and Freisen, 1982).

⁵ 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.

4.2. Distributor performance model

A major step in the company's management of its distribution system was to develop a strategically oriented PMM for its 31 North American distributorships that communicates the new customer focus, guides distributors, and provides additional bases for evaluation. This study will refer to the company's *distributor performance model* as the DPM.⁶

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 strategic initiatives, to guide and measure total distributor performance.
- Use as the starting point for the 3-year distributor contract renewal process.
- 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 company's key success factors do not perfectly mirror those of its distributors.⁷

The data for this study come from interviews with managers, 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° view of the DPM process.⁸ The length of the telephone interview was determined by how much the interviewee had to say and all lasted between 45 and 75 min. The researchers asked each interviewee the following open questions:⁹

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.

⁶ 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.

⁷ For a discussion of the conflict resulting from this initial mismatch, see [Malina and Selto \(2001\)](#).

⁸ For a complete discussion of the sampling technique, see [Malina and Selto \(2001\)](#).

⁹ 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.

Table 2
Coding scheme

	Code
1.1	A1 ₊ : Diverse and complementary
1.2	A1 ₋ : Not diverse and complementary
2.1	A2 ₊ : Objective and accurate
2.2	A2 ₋ : Not objective and accurate
3.1	A3 ₊ : Informative
3.2	A3 ₋ : Not informative
4.1	A4 ₊ : Strategic communication devices
4.2	A4 ₋ : Not strategic communication devices
5.1	A5 ₊ : Incentives for improvement
5.2	A5 ₋ : No incentives for improvement
6.1	A6 ₊ : Supportive of improved decisions
6.2	A6 ₋ : Not supportive of improved decisions
7.1	A7 ₊ : More beneficial than costly
7.2	A7 ₋ : Not more beneficial than costly
8.1	A8 ₊ : Causally related
8.2	A8 ₋ : Not causally related
9.1	Measure dropped from DPM
9.2	Measure kept on DPM

Interviewees freely revealed these factors during the interviews. Their unprompted responses were used to support or deny the predictions.¹⁰

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 interview text referring to a choice or a change in DPM measures according to each of the eight attributes 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 an attribute, an additional code was attached regarding whether the measure was dropped from the DPM or remained on the DPM during the time period tested.¹¹ This restrictive approach to coding qualitative data is designed to illustrate or test theory.¹² A complete list of codes used in this research is shown in Table 2.

¹⁰ 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.

¹¹ 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% 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.

¹² Miles and Huberman (1994) provide extensive explanations of alternative coding methods.

Table 3
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 percentage of sales	4.0	PBIT as percentage of sales	10
Cash flow as percentage 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 financial statement submission	1.0		
		Image	10
DPM total weight	100	DPM total weight	100

5. Data analysis

This section is divided into two subsections. The first subsection describes evidence regarding the initial choice of performance measures. The second subsection uses both qualitative and quantitative analyses to evaluate interview evidence related to the eight attributes regarding changes of performance measures.

Throughout the 4 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 3. Table 4 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

Table 4
DPM measure definitions

Measure	Definition
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	
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 h
Service cycle time	Percent of service repairs diagnosed within 1 h and completed within 6 h
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 percentage of sales	PBIT as a percentage of sales
Cash flow as percentage 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	Twelve quarter rolling regression over the most recent 12 quarters parts sales growth
Service sales growth	Twelve quarter rolling regression over the most recent 12 quarters non-warranty sales growth
New market 1 sales growth	Twelve quarter rolling regression over the most recent 12 quarters new market 1 sales growth
New market 2 sales growth	Twelve quarter rolling regression over the most recent 12 quarters new market 2 sales growth
Other sales growth	Twelve 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 h 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 vs. planned expenditure on non-warranty repairs
Days to financial 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)

measures (performance reviews, industry involvement, and training) were dropped from the DPM. Of the 15 measures dropped from the initial DPM, 14 were eliminated at the end of 1999.

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 5 chronicles the revisions in the weights assigned to the measures dropped from the DPM over the study period.

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 and changes to them.

5.1. Initial choice of performance measures

Company designers and managers volunteered eleven comments regarding how and why the initial measures appeared on the DPM. The most cited reason (five 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. The company's conservative tradition can be seen in its emphasis on objectivity and accuracy in performance measures. However, the change in the approach to measuring distributor performance appears as emphasis was also placed on the attributes of strategic communication and causality.

5.2. Analysis of RQ1—Are measure attributes important considerations for performance measure choice?

5.2.1. Qualitative analysis of RQ1

The following subsections describe the results of the qualitative analysis of changes in measures summarized in Table 6. Table 6 presents the distributions of codes, by positive or negative form of the attributes, 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 positive and negative perceptions. The code frequencies in columns 1 and 2 are meant to reassure the reader that the authors did not selectively focus on unrepresentative comments for the analyses that follow. Column 3 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.

5.2.1.1. Diverse and complementary (A1). 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

Table 5
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—traditional market	27	28	28	28	28	18	18	18	10	10	10	10	10			
Market share—new market 1	3	3	3	3	3	7	7	7								
Market share—new market 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 financial statement submission	1	1	1	1	1	1	1	1								

Table 6
Changes in DPM measures code frequency

	Measure dropped	Measure retained	Number of respondents
Design attributes			
A1 ₊ : Diverse and complementary	0	3	3
A1 ₋ : Not diverse and complementary	11	4	8
Total	11	7	
A2 ₊ : Objective and accurate	14	12	13
A2 ₋ : Not objective and accurate	46	19	13
Total	60	31	
A3 ₊ : Informative	0	0	0
A3 ₋ : Not informative	0	0	0
Total	0	0	
A4 ₊ : Benefits exceed costs	3	2	3
A4 ₋ : Costs exceed benefits	17	0	9
Total	20	2	
A5 ₊ : Causally related	13	23	13
A5 ₋ : Not causally related	4	0	4
Total	17	23	
Use attributes			
A6 ₊ : Strategic communication devices	17	6	13
A6 ₋ : Not strategic communication devices	1	0	1
Total	18	6	
A7 ₊ : Incentives for improvement	18	6	14
A7 ₋ : No incentives for improvement	2	3	5
Total	20	9	
A8 ₊ : Supportive of improved decisions	20	7	9
A8 ₋ : Not supportive of improved decisions	3	0	3
Total	23	7	
Total positive comments	85	59	
Total negative comments	84	26	
Total comments	169	85	

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 over-represented 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 mar-

ket 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 attribute, 15 supported the negative form that measures are not diverse and complementary.

5.2.1.2. Objective and accurate (A2). Objectivity, accuracy, reliability, and auditability appear to have been the company's and distributors' primary concerns for DPM measures. Interviewees commented on this attribute 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 negative form of the attribute that measures are not objective and accurate.

5.2.1.3. Informative (A3). 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., 7th 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, this attribute cannot be analyzed at the individual-measure level and will be withheld from the quantitative analyses that follow.

5.2.1.4. Costs versus benefit (A4). The company perceived a number of DPM measures to be misleading or unreliable and also perceived the costs of resolving disputes about these subjective measures to be greater than the benefits derived. Seventeen comments stated 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.

Agreement existed among the interviewees that the cost of some measures exceeded their benefits, providing support for the negative form of this attribute.

5.2.1.5. Causally related (A5). 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 this attribute supported the positive form that DPM measures reflect causality.

5.2.1.6. Strategic communication devices (A6). At the time of the DPM rollout, the company also launched its new customer-oriented objective. 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 strategy to sustain market share. 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 positive form of A6 that measures are strategic communication devices. Distributors and designers all agreed that DPM measures helped communicate corporate strategy to the distribution system.

5.2.1.7. Incentives for improvement (A7). The comments regarding incentive effects of DPM measures were generally in support of the positive form of this attribute 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 who saw 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 3-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 supported the positive form of A7 that measures provide incentives for improvement.

5.2.1.8. Supportive of improved decisions (A8). 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.

Table 7

Research question 1, positive vs. negative comments: analysis of attribute codes and measure retention

	Actual distribution		No effect distribution		Chi-square	p-value
	Measure dropped	Measure retained	Measure dropped	Measure retained		
Positive comments	85	59	96	48	8.42	<0.01
Negative comments	84	26	73	37		
Total comments	169	85				

The service cycle time measure, when it was part of the DPM, did affect 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 positive form of the attribute.

5.2.2. Quantitative analysis of RQ1—Are measure attributes important considerations for performance measure choice?

If theories of performance measurement choice are descriptive, comments from interviewees should tend to reflect favorable perceptions of the attributes of retained measures (the positive form of the attributes), whereas comments should reflect unfavorable perceptions of the attributes of dropped measures (the negative form of the attributes). Table 7 presents the distributions of codes, by positive or negative form of the attributes, for measures dropped from the DPM and for those retained on the DPM. A Chi-square test compares the observed distribution of comments to the expected distribution, which assumes that dropping or keeping a measure is unrelated to positive or negative forms of the attributes.

The Chi-square test of RQ1 indicates that the comment pattern is not random, but is significantly different ($p < 0.01$) from the expected (no effect) distribution. This result reinforces the qualitative results that measure attributes are important considerations for dropping or keeping DPM measures. Retained measures are more likely to have positive forms of the attributes, and dropped measures are more likely to have negative forms of the attributes. This result, which is consistent with the theory presented in Section 2, demonstrates that measure attributes are important considerations for performance measure choice.

5.3. Quantitative analysis of RQ2—Does the importance of attributes differ according to firm strategy?

Research question 2 asks if organizational strategy is related to the relative importance placed on performance measure attributes. As described in the previous section, the company itself follows a rather conservative strategy and is expected to prefer conservative performance measures. Table 8 presents the distributions of codes, by positive or negative form, of attributes expected to be con-

Table 8

Research question 2, conservative vs. entrepreneurial strategy attributes: analysis of attribute codes and measure retention

	Actual distribution		No Effect distribution		Chi-square	<i>p</i> -value
	Measure dropped	Measure retained	Measure dropped	Measure retained		
Conservative attributes						
Positive comments—conservative attributes ^a	17	14 ^b	22	9 ^b	5.26	<0.025
Negative comments—conservative attributes	63	19	58	24		
Total comments	80	33				
Entrepreneurial attributes						
Positive comments—entrepreneurial attributes ^c	50	39	55	34	5.40	<0.025
Negative comments—entrepreneurial attributes	19	4 ^b	14	9 ^b		
Total comments	69	43				

^a Attributes predicted to most applicable for a conservative strategy are A2 (objective and accurate) and A4 (cost vs. benefit).^b Distribution value that contributes the most to the Chi-square statistic. The 5-comment variance represents a 55.5% difference between the actual and expected distribution.^c Attributes predicted to be most applicable for an entrepreneurial strategy are A1 (diverse and complementary), A5 (causally related), A6 (strategic communication device), and A8 (improved decision-making).

sistent with a conservative strategy and for those expected to be consistent with an entrepreneurial strategy.

The Chi-square tests of the conservative and entrepreneurial strategy attributes indicate that the patterns are not random, and significant differences exist between the observed and expected (no effect) distributions ($p < 0.025$). For the conservative attributes test, the largest impact on the Chi-square statistic results from a larger than expected number of positive comments about retained measures. Conversely, for the entrepreneurial attributes test, the largest impact on the Chi-square statistics stems from a smaller than expected number of negative comments that are associated with retained measures. These results indicate that both conservative and entrepreneurial attributes have been important to the company's choices, in subtly different ways. Measures consistent with a conservative strategy were retained if they were positively perceived while measures consistent with an entrepreneurial strategy appear to be retained if comments were not overly negative.

5.4. Quantitative analysis of RQ3—Does the importance of attributes for design and use differ according to firm strategy?

Table 9 presents the distributions of comment codes, by positive or negative form, of the attributes divided into design and use. This exploratory Chi-square test indicates that the overall pattern is not random, but is significantly different ($p < 0.001$) from the expected (no effect) distribution. The largest impact on the Chi-square statistic comes from positive comments about design attributes. Next, strategy is introduced to investigate whether firm strategy affects the importance of design and use attributes. Since attributes associated with a more conservative strategy are exclusively design attributes, the Chi-square

Table 9

Research question 3, design vs. use attributes: analysis of attribute codes and measure retention

	Actual distribution		No effect distribution		Chi-square	<i>p</i> -value
	Measure dropped	Measure retained	Measure dropped	Measure retained		
Overall						
Positive comments—design attributes ^a	30	40 ^b	47	23 ^b	24.61	<0.001
Negative comments—design attributes	78	23	67	34		
Positive and negative comments—use attributes ^c	61	22	55	28		
Total comments	169	85				
Conservative						
Positive comments—design attributes	17	14 ^d	22	9 ^d	5.26	<0.025
Negative comments—design attributes	63	19	58	24		
Total comments	80	33				
Entrepreneurial						
Positive comments—design attributes	13	26 ^f	24	15 ^f	20.28	<0.001
Negative comments—design attributes	15	4	12	7		
Positive and negative comments—use attributes ^e	41	13	33	21		
Total comments	69	43				

^a Design attributes include A1 (diverse and complementary), A2 (objective and accurate), A4 (cost vs. benefit), and A5 (causally related).

^b Distribution value that contributes the most to the Chi-square statistic. The 17-comment variance represents a 73.9% difference between the actual and expected distribution.

^c Use attributes include A6 (strategic communication devices), A7 (incentive for improvement), and A8 (improved decision-making). Cells with small expected counts ($n < 5$) have been collapsed.

^d Distribution value that contributes the most to the Chi-square statistic. The 5-comment variance represents a 55.5% difference between the actual and expected distribution.

^e Use attributes associated with an entrepreneurial strategy include A6 (strategic communication devices) and A8 (improved decision-making). Cells with small expected counts ($n < 5$) have been collapsed.

^f Distribution value that contributes the most to the Chi-square statistic. The 11-comment variance represents a 73.3% difference between the actual and expected distribution.

test is run solely for design attributes. Thus, this test is identical to that reported in Table 8; the firm appears to have considered positive design attributes identified with a more conservative strategy when deciding to retain a performance measure ($p < 0.025$).

The final Chi-square test indicates that the firm also considered design attributes associated with an entrepreneurial strategy to be important considerations ($p < 0.001$). Thus, design attributes of measures appear to dominate usage attributes, regardless of the strategic orientation of measures.

5.5. Quantitative analysis of RQ4—Does a company trade-off some attributes for others?

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 10. The first column of numbers in Table 10 reproduces comment frequencies shown in Table 6 for entrepreneurial use attributes. The

Table 10
Research question 4, trade-off of individual attributes

Comments in unexpected direction		Associated unfavorable comments						
Measures	Favorable comments on dropped measures	A1_	A2_	A4_	A5_	A6_	A8_	None
A6 ₊ : Strategic communication devices	17	1	4	2	0	–	1	9
A8 ₊ : Supportive of improved decisions	20	0	6	9	0	0	–	4
Total	37	1	10	11	0	0	1	13
Proportions of total comments	1.000	0.027	0.270	0.297	0.000	0.000	0.027	0.350

A1_–: not diverse and complementary; A2_–: not objective and accurate; A4_–: not more beneficial than costly; A5_–: not causally related; A6_–: not strategic communication devices; A8_–: not supportive of improved decisions.

company dropped measures that are strongly perceived as strategic communication devices (A6₊, 17 comments) and supportive of improved decisions (A8₊, 20 comments), when straightforward predictions indicate that the measures should have been retained.

To determine whether the company trades off these apparently beneficial 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.¹³ The columns of Table 10 labeled A1_– to A8_– present the number of times that favorable comments in the first column are associated with negative comments for each of the other attributes. For example, interviewees said that the people-related performance measures, among others, help distributors make improved decisions (i.e., support for A8₊). However, the people measures (among others) were dropped from the DPM.

The Atlas.ti query tool found co-occurring comments (27% of the total) stating that the dropped measures, which ought to support improved decisions, are not objective and accurate (A2_–). This indicates that the DPM attribute of improved decision-making is subordinate to the attribute of objectivity and accuracy. More co-occurring comments (27.9%) stated that dropped measures were not cost-beneficial. Almost no other types of co-occurring negative comments appeared. When the company unexpectedly dropped measures, it consistently made choices consistent with the proposition that the attributes of objectivity and accuracy (A2) and benefits versus costs (A4) were more important than attributes of strategic communication (A6) and support of improved decisions (A8).

6. Conclusions and discussion

This study focuses on the efforts of one large firm to model drivers of its distribution performance. It 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

¹³ Co-occurrence or proximity rules include coded quotations of one type that *enclose*, *are enclosed by*, *overlap*, *are overlapped by*, *preceded 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.

in a live setting. The results of this qualitative investigation augment previous survey, experimental, and normative studies that collectively are refining a theory of performance measurement.

6.1. Theory-based analyses

Management control and strategy theories identify at least eight desirable attributes of performance measures. Measures should be diverse and complementary, objective and accurate, informative, more beneficial than costly, causally related, strategic communication devices, incentives for improvement, and supportive of improved decisions. Analysis of interviews with company PMM designers and users find that, taken as a whole, these attributes are relevant to decisions to delete or retain performance measures. This study finds that the studied organization's choices of performance measures appear to be strongly influenced by measures' attributes, consistent with theory.

Further analysis yields two results related to the attributes' relative importance for retaining or dropping performance measures. First, the relative importance of performance measure attributes in decisions to drop or retain measures differs subtly according to firm strategy. Positive attributes of conservative-strategy measures and relative lack of negative attributes of entrepreneurial-oriented measures are associated with the measures retained on the PMM. Because the company itself followed a rather conservative strategy, we expected conservative-strategy measures to be dropped only if they had an abundance of negative attributes. This appeared to be the case. Likewise, because of the top-down nature of distribution management, we expected that the company would keep only the entrepreneurial-strategy measures (urged by entrepreneurial distributors) that had a relative abundance of positive attributes. However, in this case, entrepreneurial measures tended to be kept as long as they were perceived to not have excessive negative attributes. This result appears somewhat inconsistent with our interpretations of prior contingency theory strategic-fit research on attributes.

We identify three possible reasons for this mild contradiction. First, prior research investigates strategy-fit at the PMM level, not at the individual-measure level. Our predictions of which attributes are more applicable for the two broad definitions of firm strategy are based on the applicability of these PMM-level results to measure-level analysis. The measurement-level of analysis may be contributing to the unexpected result.

Second, the classification of firm strategy does not lend itself to a clear-cut conservative or entrepreneurial dichotomy. As cautioned by [Chenhall \(2003\)](#), the extant typologies might be overly simplistic for modern, dynamic firms. Most aspects of the company strategy reflect a conservative strategy yet not all follow neatly in line. For example, the recent change to focus on customer satisfaction is more in line with an entrepreneurial, differentiation strategy than a conservative strategy.

Finally, and related to the difficulty of neatly categorizing a company's strategy, two related entities are being studied. The company might follow a conservative strategy while the distributors, although still fitting mostly into a conservative strategy classification, have distinct entrepreneurial tendencies. The company's narrow performance focus had caused considerable tension between management and regional distributors who saw more profitable opportunities in the company's secondary markets. Unlike the company, many distributors' success stems from tapping the less established and growing secondary markets. This strategic tension played out in the evolution of a PMM for the distribution channel and may be contributing to the subtle difference in criteria for keeping or dropping performance measures. In other words, distributors, because they have direct impacts on customers and company sales, might have more influence on PMM design than we anticipated.

6.2. *Exploratory analyses*

Attributes that we associated with design appear to be more influential than those associated with use. This finding led to an investigation of trade-offs among attributes. Although all of the attributes might be desirable, the two attributes that appear most associated with keeping or dropping individual DPM measures are (a) the objectivity and accuracy of measures and (b) the cost versus benefit balance of measurement. A measure might promote good decision-making and communicate strategy, for example, but the company dropped it if experience showed it could not be measured inexpensively, accurately and objectively. Trade-offs also permit a more detailed analysis of the effect of strategy on PMM choice. At an aggregate level, only a subtle difference was found in the importance of attributes consistent with either a conservative versus an entrepreneurial strategy. However, at the individual-attribute level, we find that objectivity and accuracy as well as cost versus benefit, both consistent with the company's conservative strategy, accompany changes in DPM measures.

6.3. *Discussion*

The 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, primary market share. Primary market share was important to corporate strategy, but the company also had the ability to measure primary market share extremely accurately and objectively. Many distributors complained that the traditional market did not afford them growth opportunities and that more weight should attach to new market shares. Unfortunately, new market shares are difficult to measure accurately and objectively. The company ultimately eliminated all the market share measures from the DPM. The company decided to move the weight assigned to primary market share to sales growth measures. Measurement of sales growth is easily obtained, and relatively accurate and objective for both the traditional market and the new markets, although sales growth does not speak directly to the company's strategy to grow market share. However, adding sales growth measures retained diversity among measures and recognized distributors' entrepreneurial opportunities.

Apparently this company, with its long history of conservative, 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.

6.4. *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 trade-offs among measure attributes that might be necessary to implement feasible PMM?
- Do all firms or only those with a conservative strategy stress cost of measurement, accuracy and objectivity of performance measures above all other attributes?
- Is complementarity an issue beyond initial design? Do returns or trade-offs to scale and scope of measurement exist?
- What are the performance opportunity costs of trading off performance measure attributes?

6.5. Epilogue

Three years after the introduction of the DPM, as reported in [Malina and Selto \(2001\)](#), distributors' measured financial performance improved significantly. The DPM is alive and well as of May 2004, 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.

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. This research is supported by data generously provided by the anonymous host company, which has granted permission to publish this research. Proprietary data beyond that disclosed in this paper are not available.

References

- Abernethy, M.A., Brownell, P., 1999. The role of budgets in organizations facing strategic change: an exploratory study. *Acc. Organ. Society* 24 (3), 189–204.
- Adimando, C., Butler, R., Malley, S., Ravid, S.A., Shepro, R., Vernon, R.H., Willett, J., Zimmerman, J., 1994. Stern Stewart EVA Roundtable. *J. Appl. Corp. Finance* 7 (2), 46–70.
- Amit, R., Schoemaker, P., 1993. Strategic assets and organizational rent. *Strategic Manage. J.* 14, 33–46.
- Anderson, S.W., Young, S.M., 1999. The impact of contextual and process factors on the evaluation of activity-based costing systems. *Acc. Organ. Society* 24, 525–559.

- Antle, R., Demski, J.S., 1988. The controllability principle in responsibility accounting. *Acc. Rev.* 63 (4), 700–718.
- Barabba, V., Huber, C., Cooke, F., Pudar, N., Smith, J., Paich, M., 2002. A multimethod approach for creating new business models: the General Motors OnStar project. *Interfaces* 32 (1), 20–34.
- Barney, J., 1991. Firm resources and sustained competitive advantage. *J. Manage.* 17 (1), 99–120.
- Bouwens, J., Abernethy, M., 2000. The consequences of customization on management accounting system design. *Acc. Organ. Society* 25 (3), 221–259.
- Cavaluzzo, K., Ittner, C., 2002. Implementing performance measurement innovations: evidence from government. Georgetown University Working Paper.
- Chenhall, R.H., 2003. Management control systems design within its organizational context: findings from contingency-based research and directions for the future. *Acc. Organ. Society* 28, 127–168.
- Chenhall, R.H., Morris, D., 1995. Organic decision and communication processes and management accounting systems in entrepreneurial and conservative business organizations. *Omega, Int. J. Manage. Sci.* 23 (5), 485–497.
- Cloutier, L.M., Boehlje, M., et al., 2002. Innovation management under uncertainty: a system dynamics model of R&D investments in biotechnology. In: Morecroft (Ed.), *Systems Perspectives on Resources, Capabilities, and Management Processes*. Pergamon, Amsterdam.
- Dent, J.F., 1990. Strategy, organization and control: some possibilities for accounting research. *Acc. Organ. Society* 15, 3–25.
- Eccles, R., 1991. The performance measurement manifesto. *Harvard Bus. Rev.* 69, 131–137 (*Economics* 19, 45–48).
- Epstein, M.J., Kumar, P., Westbrook, R.A., 2000. The drivers of customer and corporate profitability: modeling, measuring, and managing the causal relationships. *Advances Manage. Acc.* 9 (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*. Greenwood Publishing Group, Westport, CT.
- Govindarajan, V., Gupta, A.K., 1985. Linking control systems to business unit strategy: impact on performance. *Acc. Organ. Society*, 51–66.
- Guilting, C., 1999. Competitor-focused accounting: an exploratory note. *Acc. Organ. Society* 24 (7), 583–595.
- Gupta, A.K., Govindarajan, V., 1984. Business unit strategy, managerial characteristics, and business unit effectiveness at strategy implementation. *Acad. Manage. J.*, 25–41.
- Huff, A. (Ed.), 1990. *Mapping Strategic Thought*. Wiley, New York.
- Huff, A., Jenkins, M. (Eds.), 2003. *Mapping Strategic Knowledge*. Sage Publications, Thousand Oaks, CA.
- Horngren, C.T., 1967. *Cost Accounting: A Managerial Emphasis*, second ed. Prentice-Hall, Englewood Cliffs, NJ.
- Ijiri, Y., 1967. *The Foundations of Accounting Measurement*. Scholars Book Co., Houston, TX.
- Ittner, C.D., Larcker, D.F., 2001. Assessing empirical research in managerial accounting: a value-based management perspective. *J. Acc. Econ.* 32, 349–410.
- Ittner, C.D., Larcker, D.F., 2002. Determinants of performance measure choices in worker incentive plans. *J. Lab. Econ.* 20 (2), 58.
- Ittner, C.D., Larcker, D.F., 1998. Are non-financial measures leading indicators of financial performance? An analysis of customer satisfaction. *J. Acc. Res.* 36 (Suppl.), 1–35.
- Ittner, C.D., Larcker, D.F., Randall, T., 2002. Performance implications of strategic performance measurement in financial services firms. Wharton School Working Paper.
- Kanji, G., Moura e Sa, P., 2002. Kanji's business scorecard. *Total Quality Manage.* 13 (1), 13–27.
- Kaplan, R.S., Norton, D.P., 2001. *The Strategy-Focused Organization*. Harvard Business School Press, Boston, MA.
- Kaplan, R.S., Norton, D.P., 1996. *The Balanced Scorecard*. Harvard Business School Press, Boston, MA.
- Khandwalla, P.N., 1972. The effect of different types of competition on the use of management controls. *J. Acc. Res.* 10 (2), 275–285.
- Kogut, B., Kulatilaka, N., 2001. Capabilities as real options. *Organ. Sci.* 12 (6), 744–758.
- Langfield-Smith, K., 1997. Management control systems and strategy: a critical review. *Acc. Organ. Society* 22 (2), 207–232.
- Libby, T., Salterio, S., Webb, A., 2002. The balanced scorecard: the effects of assurance and process accountability on managerial judgment. Wilfred Laurier University Working Paper.
- Lipe, M., Salterio, S., 2000. The balanced scorecard: judgmental effects of common and unique performance measures. *Acc. Rev.* 75, 283–298.
- Luft, J., Shields, M., 2001. Why does fixation persist? Experimental evidence on the judgment performance effects of expensing intangibles. *Acc. Rev.* 76 (4), 561–587.

- Luft, J., Shields, M., 2002. 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., Shields, M., 2003. Mapping management accounting: graphics and guidelines for theory-consistent empirical research. *Acc. Organ. Society* 28 (2/3), 169–249.
- Magretta, J., 2002. Why business models matter. *Harvard Bus. Rev.* 80 (5).
- Malina, M.A., Selto, F.H., 2001. Communicating and controlling strategy: an empirical study of the effectiveness of the balanced scorecard. *J. Manage. Acc. Res.* 13, 48–90.
- McKenzie, F.C., Shilling, M.D., 1998. Avoiding performance measurement traps: ensuring effective incentive designs and implementation. *Compensation Benefits Rev.* 30 (4), 57–65.
- Merchant, K., 1998. *Modern Management Control Systems: Text and Cases*. Prentice-Hall, Upper Saddle River, NJ.
- Miles, M.B., Huberman, A.M., 1994. *Qualitative data analysis*. Sage Publications, Thousand Oaks, CA.
- Miles, R.W., Snow, C.C., 1978. *Organizational Strategy, Structure and Process*. McGraw-Hill, New York.
- Milgrom, P., Roberts, J., 1995. Complementarities and fit: strategy, structure, and organizational change in manufacturing. *J. Acc. Econ.* 19 (2/3), 179–208.
- Miller, D., Freisen, P.H., 1982. Innovation in conservative and entrepreneurial firms: two models of strategic momentum. *Strategic Manage. J.*, 1–25.
- Morecroft, J., et al., 2002. Resource management under dynamic complexity. In: Morecroft (Ed.), *Systems Perspectives on Resources, Capabilities, and Management Processes*. Pergamon, Amsterdam.
- Morecroft, J., Sanchez, R., Heene, A. (Eds.), 2002. *Systems Perspectives on Resources, Capabilities, and Management Processes*. Pergamon, Amsterdam.
- Nonaka, I., Takeuchi, H., 1995. *The Knowledge-Creating Company*. Oxford University Press, New York.
- Otley, D.T., 1999. Performance management: a framework for management control systems research. *Manage. Acc. Res.*, 363–382.
- Porter, M., 1980. *Competitive strategy*. The Free Press, New York.
- Porter, M., 1985. *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press, New York.
- Prahalad, C., Hamel, G., 1990. The core competence of the corporation. *Harvard Bus. Rev.* 90 (3), 79–91.
- Rappaport, A., 1999. New thinking on how to link executive pay to performance. *Harvard Bus. Rev.* March–April, 91–101.
- Rucci, A.J., Kirn, S.P., Quinn, R.T., 1998. The employee–customer–profit chain at Sears. *Harvard Bus. Rev.* 76 (1), 82–97.
- Sanchez, R., Heene, A., Thomas, H. (Eds.), 1996. *Dynamics of Competence-Based Competition*. Pergamon, Oxford.
- Senge, P., 1990. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday, New York.
- Simons, R., 1987. Accounting control systems and business strategy: an empirical analysis. *Acc. Organ. Society* 12, 357–374.
- Simons, R., 2000. *Performance Measurement & Control Systems for Implementing Strategy: Text & Cases*. Prentice-Hall, Upper Saddle River, NJ.
- Van de Ven, A., Drazin, 1985. The concept of fit in contingency theory. *Res. Organ. Behav.* 7, 333–365.
- Warren, K., et al., 2002. Operationalizing the impact of competence building on the performance of firms' resource systems. In: Morecroft (Ed.), *Systems Perspectives on Resources, Capabilities, and Management Processes*. Pergamon, Amsterdam.
- Yin, R.K., 1994. *Case Study Research Design and Methods*. Sage Publications Inc., Newbury Park, CA.
- Zimmerman, J., 1997. *Accounting for Decision-making and Control*. Irwin–McGraw-Hill, Burr Ridge, IL.