BEHAVIORAL RESEARCH IN ACCOUNTING
Volume 18, 2006
pp. 185–205

Strategy, Choice of Performance Measures, and Performance

Wim A. Van der Stede
University of Southern California

Chee W. Chow
San Diego State University

Thomas W. Lin
University of Southern California

ABSTRACT: We examine the relationship between quality-based manufacturing strategy and the use of different types of performance measures, as well as their separate and joint effects on performance. A key part of our investigation is the distinction between financial and both objective and subjective nonfinancial measures. Our results support the view that performance measurement diversity benefits performance as we find that, regardless of strategy, firms with more extensive performance measurement systems—especially those that include objective and subjective nonfinancial measures—have higher performance. But our findings also partly support the view that the strategy-measurement "fit" affects performance. We find that firms that emphasize quality in manufacturing use more of both objective and subjective nonfinancial measures. However, there is only a positive effect on performance from pairing a quality-based manufacturing strategy with extensive use of subjective measures, but not with objective nonfinancial measures.

INTRODUCTION

Performance measures play a key role in translating an organization's strategy into desired behaviors and results (Campbell et al. 2004; Chenhall and Langfield-Smith 1998; Kaplan and Norton 2001; Lillis 2002). They also help to communicate expectations, monitor progress, provide feedback, and motivate employees through performance-based rewards (Banker et al. 2000; Chenhall 2003; Ittner and Larcker 1998b; Ittner et al. 1997; Ittner, Larcker, and Randall 2003). Traditionally, firms have primarily used financial measures for these purposes (Balkcom et al. 1997; Kaplan and Norton 1992). But with the ''new'' competitive realities of increased customization, flexibility, and responsiveness, and associated advances in manufacturing practices, both academics and practitioners have argued that traditional financial performance measures are no longer adequate for these functions (Dixon et al. 1990; Fisher 1992; Ittner and Larcker 1998a; Neely 1999). Indeed, many

We acknowledge the helpful suggestions by Tom Groot, Jim Hesford, Ranjani Krishnan, Fred Lindahl, Helene Loning, Michal Matejka, Ken Merchant, Frank Moers, Mark Peecher, Mike Shields, Sally Widener, workshop participants at the University of Illinois, the 2002 AAA Management Accounting Meeting in Austin, the 2002 World Congress of Accounting Educators in Hong Kong, and the 2003 AAA Annual Meeting in Honolulu. An earlier version of this paper won the best paper award at the 9th World Congress of Accounting Educators in Hong Kong (2002).

accounting researchers have identified the continued reliance on traditional management accounting systems as a major reason why many new manufacturing initiatives perform poorly (Banker et al. 1993; Ittner and Larcker 1995).

In light of this development in theory and practice, the current study seeks to advance understanding of the role that performance measurement plays in executing strategy and enhancing organizational performance. It proposes and empirically tests three hypotheses about the performance effects of performance measurement diversity; the relation between quality-based manufacturing strategy and firms' use of different types of performance measures; and the joint effects of strategy and performance measurement on organizational performance.

The distinction between objective and subjective performance measures is a pivotal part of our investigation. Prior empirical research has typically only differentiated between financial and nonfinancial performance measures. We go beyond this dichotomy to further distinguish between nonfinancial measures that are quantitative and objectively derived (e.g., defect rates), and those that are qualitative and subjectively determined (e.g., an assessment of the degree of cooperation or knowledge sharing across departmental borders). Making this finer distinction between types of nonfinancial performance measures contributes to recent work in accounting that has begun to focus on the use of subjectivity in performance measurement, evaluation, and incentives (e.g., Bushman et al. 1996; Gibbs et al. 2004; Ittner, Larcker, and Meyer 2003; MacLeod and Parent 1999; Moers 2005; Murphy and Oyer 2004).

Using survey data from 128 manufacturing firms, we find that firms with more extensive performance measurement systems, especially ones that include objective and subjective nonfinancial measures, have higher performance. This result holds regardless of the firm's manufacturing strategy. As such, our finding supports the view that performance measurement diversity, per se, is beneficial.

But we also find evidence that firms adjust their use of performance measures to strategy. Firms that emphasize quality in manufacturing tend to use more of both objective and subjective nonfinancial measures, but without reducing the number of financial measures. Interestingly, however, combining quality-based strategies with extensive use of objective nonfinancial measures is not associated with higher performance. This set of results is consistent with Ittner and Larcker (1995) who found that quality programs are associated with greater use of nontraditional (i.e., nonfinancial) measures and reward systems, but combining nontraditional measures with extensive quality programs does not improve performance. However, by differentiating between objective and subjective nonfinancial measures—thereby going beyond Ittner and Larcker (1995) and much of the extant accounting literature—we find that performance is higher when the performance measures used in conjunction with a quality-based manufacturing strategy are of the subjective type.

Finally, we find that among firms with similar quality-based strategies, those with *less* extensive performance measurement systems have lower performance, whereas those with *more* extensive performance measurement systems do not. In the case of subjective performance measures, firms that use them more extensively than firms with similar quality-based strategies actually have significantly higher performance. Thus, a "*mis*match" between performance measurement and strategy is associated with lower performance only when firms use fewer measures than firms with similar quality-based strategies, but not when they use more.

The paper proceeds as follows. The next section builds on the extant literature to formulate three hypotheses. The third section discusses the method, sample, and measures.

The fourth section presents the results. The fifth section provides a summary, discusses the study's limitations, and suggests possible directions for future research.

HYPOTHESES

Although there is widespread agreement on the need to expand performance measurement, two different views exist on the nature of the desirable change (Ittner, Larcker, and Randall 2003; Ruddle and Feeny 2000). In this section, we engage the relevant literatures to develop three hypotheses. Collectively, the hypotheses provide the basis for comparing the two prevailing schools of thought on how performance measurement should be improved; that of performance measurement *diversity* regardless of strategy versus that of performance measurement *alignment* with strategy (Ittner, Larcker, and Randall 2003).

The Performance Measurement Diversity View

A number of authors have argued that broadening the set of performance measures, per se, enhances organizational performance (e.g., Edvinsson and Malone 1997; Lingle and Schiemann 1996). The premise is that managers have an incentive to concentrate on those activities for which their performance is measured, often at the expense of other relevant but non-measured activities (Hopwood 1974), and greater measurement diversity can reduce such dysfunctional effects (Lillis 2002). Support for this view is available from economics-based agency studies. Datar et al. (2001), Feltham and Xie (1994), Hemmer (1996), Holmstrom (1979), and Lambert (2001), for example, have demonstrated that in the absence of measurement costs, introducing incentives based on nonfinancial measures can improve contracting by incorporating information on managerial actions that are not fully captured by financial measures. Analytical studies have further identified potential benefits from using performance measures that are subjectively derived. For example, Baiman and Rajan (1995) and Baker et al. (1994) have shown that subjective measures can help to mitigate distortions in managerial effort by ''backing out'' dysfunctional behavior induced by incomplete objective performance measures, as well as reduce noise in the overall performance evaluation.

However, the literature also has noted potential drawbacks from measurement diversity. It increases system complexity, thus taxing managers' cognitive abilities (Ghosh and Lusch 2000; Lipe and Salterio 2000, 2002). It also increases the burden of determining relative weights for different measures (Ittner and Larcker 1998a; Moers 2005). Finally, multiple measures are also potentially conflicting (e.g., manufacturing efficiency and customer responsiveness), leading to incongruence of goals, at least in the short run (Baker 1992; Holmstrom and Milgrom 1991), and organizational friction (Lillis 2002).

Despite these potential drawbacks, there is considerable empirical support for increased measurement diversity. For example, in a study of time-series data in 18 hotels, Banker et al. (2000) found that when nonfinancial measures are included in the compensation contract, managers more closely aligned their efforts to those measures, resulting in increased performance. Hoque and James (2000) and Scott and Tiessen (1999) also have found positive relations between firm performance and increased use of different types of performance measures (e.g., financial and nonfinancial). These results are consistent with nonfinancial performance measures containing additional information not reflected in financial measures.

Although prior studies like these have advanced our understanding of performance measurement design, they share the limitation of not distinguishing among measures that are determined objectively and ones that are based on subjective judgment. A key difference between objective and subjective measures is that the latter are often less accurate and

reliable, and are more open to raters' biases (Campbell 1990; Fulk et al. 1985; Hawkins and Hastie 1990; Heneman 1986). Prendergast and Topel (1993), for example, have suggested that allowing subjective judgment in performance evaluation can reduce employee motivation. This result is due to the latitude for evaluators to ignore performance measures that are included in the performance plan, and to use measures that differ from those originally planned. Moreover, when evaluations are subjective, employees may divert job effort toward influencing their supervisors' evaluations (Milgrom 1988; Prendergast 1993; Prendergast and Topel 1996). Whether limitations like these outweigh the benefits of subjective measures is an important question.

To date, only a few studies have empirically examined the use and effects of subjectivity in performance measurement. Of particular relevance to the current study, Ittner, Larcker, and Meyer (2003) analyzed how subjective performance measures were used in a leading international financial services provider. They found that leaving room for subjectivity allowed supervisors to ignore many performance measures, and short-term financial measures became the *de facto* determinants of bonus awards. Employee dissatisfaction with the system was so high that the firm reverted to basing bonuses strictly on revenue. Moers (2005) also found that the use of subjective performance measures induces performance evaluation bias. In a study of managerial compensation plans in a privately owned Dutch maritime firm, he found that the use of multiple objective and subjective performance measures was related to more compressed and more lenient performance ratings. These findings by Ittner, Larcker, and Meyer (2003) and Moers (2005) are useful for understanding the uses and limitations of subjective performance measures; however, their generalizability is limited by each study having focused on only one company.

In sum, prior studies have argued for performance measurement to extend beyond just financial measures, and importantly for this study, a case has been made for differentiating between objective and subjective nonfinancial measures. Although arguments have been proffered about the downsides of performance measurement diversity, in general, and about the inclusion of subjective performance measures, in particular, the expectation is that:

H1: Firm performance will be positively associated with performance measurement diversity, particularly when the performance measurement system is extended to include (more) objective and subjective nonfinancial performance measures.

The Performance Measurement Alignment View

In contrast to proponents of performance measurement diversity, contingency theory maintains that the optimal design of performance measurement systems depends on the organization's strategy (and other organizational characteristics), and that performance will be higher only if both are *aligned* (for reviews, see Chenhall 2003; Fisher 1995; Langfield-Smith 1997). In this study, we specifically consider the performance measurement implications of quality-based manufacturing strategy.

The primary reason for this focus is that quality-based initiatives have challenged the relevance of traditional financial measures (Abernethy and Lillis 1995; Dixon et al. 1990), and "one way in which this challenge has been met is by performance measurement system expansion" (Lillis 2002, 498). For example, nonfinancial measures can help ensure that quality improvement results, which often take considerable time to materialize, are not subjugated to financial results (Daniel and Reitsperger 1991; Ittner and Larcker 1995). Moreover, some of the key dimensions of quality-focused strategies, such as those focused

on knowledge sharing and cooperativeness (Lillis 2002), are difficult to quantify and, thus, may need to be assessed subjectively.

Prior studies in this area, however, have not distinguished between objective and subjective nonfinancial performance measures. Moreover, although some prior studies have found associations between the choice of performance measures and the type of strategy pursued (e.g., Abernethy and Lillis 1995; Daniel and Reitsperger 1991; Ittner and Larcker 1995; Perera et al. 1997), they have rarely investigated the performance effects of such choices. Or, as Ittner and Larcker (1998a, 221) and Chenhall (2003, 142) have emphasized, the scant empirical evidence that has been reported on the strategy-measure-performance relationship—only some of which relates to manufacturing strategy—is equivocal at best. In sum, considering the limited evidence in conjunction with the failure of prior work to distinguish between objective and subjective performance measures, there is room for further investigation. Next, we review related prior studies as the basis for developing two hypotheses.

Abernethy and Lillis (1995) found that reliance on traditional (cost efficiency-based) measures was positively associated with performance for firms pursuing a nonflexible manufacturing strategy, and negatively associated with performance for firms pursuing a flexible manufacturing strategy. However, their study did not examine the performance of firms that combine nontraditional performance measures with a flexible manufacturing strategy. Furthermore, discriminatory power was low because most firms in the study extensively used virtually all of the performance measures.

Ittner and Larcker (1995) focused on firms using advanced quality programs (TQM). Their sample consisted of 249 firms from the automobile and computer industries in four countries. Using data from a consulting firm, they found ''no support for the proposition that, holding other determinants of performance constant, the highest performance levels should be achieved by organizations making the greatest use of both TQM practices and nontraditional information and reward systems'' (Ittner and Larcker 1995, 2). On the contrary, among firms with extensive use of advanced quality programs, those with a strong reliance on nontraditional performance measures had lower performance than those with less extensive use of such measures. In another study based on the same sample, Ittner and Larcker (1997) examined the relation between quality-focused strategies and strategic control practices, including the importance placed on quality performance in determining managerial compensation. The performance effects (on ROA, ROS, sales growth, and self-reported performance) of a match between the two were generally insignificant.

While the findings of both their 1995 and 1997 studies are informative, Ittner and Larcker (1995) caution that both industries in their sample may be so competitive that there was insufficient variation in the variables for identifying significant performance effects. A further limitation is that the studies only had access to measures of the overall importance placed on nonfinancial and quality performance. There was no information about the kinds and numbers of measures used of each type, nor was it possible to distinguish between objective and subjective nonfinancial performance measures.

Said et al. (2003) also examined the fit between operational and competitive circumstances and firms' choice of performance measures, as well as the performance effects of including nonfinancial measures in compensation contracts. A notable advance over prior studies is the use of a large sample (1,441 firm-years for firms using nonfinancial performance measures in managerial bonus plans, matched against the same number of firm-years for firms using exclusively financial measures). The results indicated that firms with a greater quality focus made greater use of nonfinancial measures. Furthermore, consistent

with the contingency-based approach to performance measurement, firms that relied on nonfinancial measures either more or less than a benchmark model had lower performance. Notably, these performance effects diverged from those of Ittner, Larcker, and Randall (2003) where negative deviations from the benchmark had no detectible effect and positive deviations were associated with higher performance. However, Said et al. (2003) only used crude indicators (either a dummy variable or the total weight) to reflect a firm's use or nonuse of nonfinancial measures and, hence, like the other studies of quality initiatives, they did not distinguish among different types of nonfinancial performance measures.

Taken as a whole, the contingency theory-based studies have made a case for the need to align performance measures with quality strategy. Yet they have produced only limited, and at times mixed, findings on the relation between quality initiatives and performance measure use and their joint effect on performance. Our next two hypotheses provide focus to further empirical investigation and extend prior work by incorporating the distinction between objective and subjective nonfinancial measures.

H2: Firms that place more emphasis on quality in the manufacturing strategy will extend their performance measurement system to include (more) objective and subjective nonfinancial performance measures.

H3: Firms that place more emphasis on quality in the manufacturing strategy will have higher performance only when they match their performance measurement system to their strategy, that is, when they extend their performance measurement system to include (more) objective and subjective nonfinancial performance measures.

Together, these two hypotheses posit that firms pursuing a quality-based manufacturing strategy will tend to use more extensive performance measurement systems that include objective and subjective nonfinancial performance measures, and when they do—thus aligning their choice of performance measures with their manufacturing strategy—they will exhibit superior performance.

METHOD

Sample

We employ the survey method because the data sought are not available from public sources.[1] Given the central role of manufacturing strategy in our study, we target managers or directors of manufacturing in manufacturing firms.

To obtain a large enough sample for statistical tests, we include both U.S. and European firms. Using the Explore database from CorpTech, we restrict the U.S. sample to Southern California, where the authors' schools are likely to have some goodwill. We select manufacturing firms with at least 50 employees and U.S.$2.5 million in annual sales (678 firms) to ensure that the sample firms would have sufficient sophistication and scale to use advanced management practices (performance measurement systems, in particular). We

[1] Based on recommendations by the *Total Design Method* (Dillman 1978, 1999) and other sources on survey design (see Van der Stede et al. [2005] for a review), the authors went through several iterations of the survey before submitting it to the scrutiny of four management accounting faculty colleagues for pre-testing. We also searched the literature for existing scales whenever available. Finally, we did follow-ups and conducted nonresponse bias analyses (see below).

received 106 replies, 19 of which were unusable for a range of reasons, leaving a usable sample of 87.[2]

In Europe, we geographically restrict our target sample to Belgium where we partnered with the Vlerick Leuven Ghent Management School (VLGMS), the premier Belgian business school. The VLGMS maintains a database of Belgian firms from which we select a sample using the same employment and sales criteria as in the U.S. We obtained 48 replies from a sample of 447 firms. Of these, 41 were usable.[3]

Hence, our total usable sample size is 128 (87 + 41). The sample firms have an average headcount in the manufacturing department (firm as a whole) of about 227 (3,216). Their average annual production value (sales) in the manufacturing department (firm as a whole) is about U.S.$86 million (U.S.$1.2 billion).[4] On average, respondents have been working for their current firms for about 11 years, and in their current positions for 6 years. Their average age is 46 years, and most have a college degree. Of the above firm and respondent characteristics, only the respondents' years at the firm (10 versus 13) and age (47 versus 41) are significantly different between the U.S. and Belgian samples, respectively ($p < 0.10$). We can discern no reason why these two characteristics would bias the results. More importantly, the U.S. and Belgian samples do not differ ($p > 0.10$) along any self-reported performance dimension (financial, operating, employee-related, and customer-related).

We perform two tests to assess the potential for nonresponse bias. First, we compare sales and employment (available in the databases) between the responding and nonresponding firms. Neither variable is significantly different between U.S. and Belgian respondents and nonrespondents. Second, we compare early versus late respondents (i.e., those who responded after the final follow-up) on all the variables used in the analyses (quality-based manufacturing strategy, number of each category of measures, and performance). We find no significant difference (at $p < 0.10$) on any of these variables in either the U.S. or Belgian sample. While these tests do not constitute definitive evidence, they do increase our confidence that there is neither significant nonresponse bias between the target and survey samples, nor important systematic differences between the U.S. and Belgian samples.

Variables

The survey contains questions on the performance measures in use, manufacturing strategy, performance, and the firm and respondent demographics reported above.

---

[2] The 106 replies represent 16 percent of the 647 valid mailings (31 of the 678 surveys were returned as undeliverable). Of the 106 replies, 19 were either partially incomplete, or uncompleted for reasons as diverse as: addressee deceased or no longer with the firm (7); interim manager (1); no time (2); firm in bankruptcy or reorganization (2); firm in start-up (1); nonmanufacturing (1); nondomestic operations (1); and policy of nonparticipation (4). Of the 87 usable responses (106 – 19), we received 38 (44 percent) immediately without follow-up, 20 (23 percent) after the first follow-up without replacement, and 29 (33 percent) after the second follow-up with replacement.

[3] The 48 replies represent 11 percent of the 438 valid mailings (9 of the 447 surveys were returned as undeliverable). Of the replies, 7 were not usable because operations had been stopped (3); the company had a policy of nonparticipation (3); and one addressee promised to respond later, but never did. Of the 41 usable responses (48 – 7), we received 27 (66 percent) immediately without follow-up and 14 (34 percent) after the first and only follow-up with replacement. We did not do an intermediate follow-up (just a reminder letter) in Belgium as we did in the U.S. because of cost considerations and (international) administrative complexity.

[4] The rather large difference between the company-wide and manufacturing-specific statistics is due to 45 percent of the manufacturing departments being part of a profit center, division, or business unit within the company. For the other 55 percent of the manufacturing departments in our sample, the next higher level is the firm as a whole.

Performance Measures in Use

Because of its relative complexity, we reproduce the performance measurement section of the survey in the Appendix. This section contains one subsection on financial measures, three subsections on objective nonfinancial measures (internal operating measures, employee-oriented measures, and customer-oriented measures), and one subsection on subjective performance measures.[5] We asked participants to indicate the specific performance measures currently used by upper management in their firms for performance measurement and evaluation of the manufacturing department.[6] Respondents could check off measures from the list and also write in additional measures. Table 1 reports the distributional statistics of these data.

The data reveal a wide range across the sample firms in the total number of performance measures used. We take this to imply that the respondents had tried to reflect the actual situations in their firms, and not simply checked off items on the list. Another indication that the respondents were seriously engaged is that 28 out of 128 respondents (22 percent) took the effort to write in a total of 38 additional measures.[7] Moreover, more than half of the respondents indicated that their firms had made changes to their performance measurement systems in the past three years, which we view as evidence that the firms consider issues of performance measurement important enough to merit continual scrutiny and adjustment.[8]

Table 1 shows that among the sample firms, the average performance measurement system contains about 25 percent financial measures, 56 percent objective nonfinancial measures, and 19 percent subjective measures. More specifically, the most numerous are internal operating measures (26 percent), followed by financial measures (25 percent), and

[5] We obtained this list of measures from a wide reading of both the academic and practitioner literatures, including Abernethy and Lillis (1995), Butler et al. (1997), Chow et al. (1997), Clinton and Hsu (1997), Epstein and Manzoni (1997), Fisher (1992), Hoffecker and Goldenberg (1994), Maisel (1992), Kaplan and Norton (1992, 1993, 1996a, 1996b, 1996c, 2000, 2001), Provost and Leddick (1993), Scott and Tiessen (1999), and Stivers et al. (1998).

[6] Similar to Ittner and Larcker (1995), we consider "use" to mean "the existence of a measure in the performance measurement and evaluation system" without getting into whether the specific measure is used for informational or reward purposes, or both. To curb a potential upward bias in the number of measures reported—i.e., situations where financial, objective nonfinancial, or subjective performances are tracked *but not used* (Stivers et al. 1998)—we explicitly direct respondents "*to only check (or write in) those measures that are reported, analyzed, and discussed on a regular basis for the purpose of performance measurement and evaluation*" (see the Appendix).

[7] Of these 28 respondents, 21 provided one write-in measure, six provided two write-in measures, and one wrote in a total of five measures. By measurement category, there are 13 financial, 19 objective nonfinancial, and six subjective write-ins. Some are idiosyncratic, such as *my ability to work cooperatively with the Japanese top management.* Less idiosyncratic measures are *capital expenditures* and *receivables on net sales* for financial; *order backlog*, *product return turnaround cycle time*, and *quoted lead time accuracy* for objective nonfinancial; and *my ability to maintain strong internal control* for subjective measures.

[8] Twenty-five percent of the respondents reported that their firms' performance measurement system had been modified in the last year, 34 percent in the last 2–3 years, and 41 percent reported no change in the last three years. The percentage of sample firms indicating a change in the performance measurement system in the last three years (59 percent) is almost identical to that reported in a survey by the Institute of Management Accountants (1996) (60 percent). Of the firms in our sample reporting a change during the last 3 years, the change involved fine-tuning financial measures in use (53 percent); dropping some financial measures (11 percent); adding financial measures (47 percent); fine-tuning nonfinancial measures in use (43 percent); dropping some nonfinancial measures (5 percent); or adding nonfinancial measures (64 percent). These percentages sum to over 100 percent, indicating that the changes are multidimensional. They also reveal that adding both nonfinancial and financial measures are common, with the former being more prevalent (64 percent versus 47 percent). Moreover, dropping nonfinancial measures (5 percent) is less common than dropping financial measures (11 percent). Overall, the direction of change seems to be toward greater use of nonfinancial performance measures.

TABLE 1
Performance Measurement System Composition
(n = 128)

	Min	Max	Mean	Std.Dev.
Number of financial measures (*#_FIN*)[a]	1	11	6	3
Number of objective nonfinancial measures (*#_NONFIN*)[a]	0	26	14	5
Number of internal operating measures (*#_OPS*)[a]	0	11	6	2
Number of employee-oriented measures (*#_EMPL*)[a]	0	8	4	2
Number of customer-oriented measures (*#_CUST*)[a]	0	8	4	2
Number of subjective measures (*#_SUBJ*)[a]	0	7	5	2
Total number of measures (*#_TOTAL*)[a]	6	41	25	7
Percent financial measures (*%_FIN*)[b]	0.08	1.00	0.25	0.09
Percent objective nonfinancial measures (*%_NONFIN*)[b]	0.00	0.83	0.56	0.09
Percent internal operating measures (*%_OPS*)[b]	0.00	0.67	0.26	0.07
Percent employee-oriented measures (*%_EMPL*)[b]	0.00	0.33	0.15	0.06
Percent customer-oriented measures (*%_CUST*)[b]	0.00	0.27	0.15	0.06
Percent subjective measures (*%_SUBJ*)[b]	0.00	0.43	0.19	0.08
			100%	
Categorical measure of measurement diversity (*C_DIVERSITY*)[c]	1.00	4.00	3.29	0.99

[a] See the Appendix for the list of performance measures and their breakdown by category (financial, internal operating, employee-oriented, customer-oriented, and subjective).
[b] As a percentage of the total number of performance measures used in each firm.
[c] The categorical measure of performance measurement *C_DIVERSITY* has the following values when the unit uses: 1 = financial measures only (i.e., measures checked in Section 1 of the Appendix only); 2 = financial and objective nonfinancial measures (i.e., measures checked in Sections 1 and 2 only); 3 = financial and subjective measures (i.e., measures checked in Sections 1 and 3 only); and 4 = financial and objective nonfinancial and subjective measures (i.e., measures checked in Sections 1, 2, and 3).

lower but still non-negligible proportions of subjective performance measures (19 percent), employee-oriented measures (15 percent), and customer-oriented measures (15 percent).[9]

We operationalize the notion of performance measurement "diversity" using two different approaches (see Table 1). Our first operationalization is a count measure, that is, the number of financial (*#_FIN*), objective nonfinancial (*#_NONFIN*), and subjective (*#_SUBJ*) measures. Our second operationalization is a categorical measure (*C_DIVERSITY*) where 1 = financial measures only (i.e., measures checked in Section 1 of the Appendix only); 2 = financial and objective nonfinancial measures (i.e., measures checked in Sections 1 and 2 only); 3 = financial and subjective measures (i.e., measures checked in Sections 1 and 3 only); and 4 = financial and objective nonfinancial and subjective measures (i.e., measures checked in Sections 1, 2, and 3). This categorical approach facilitates recognizing another dimension of performance measurement diversity, as in the case of comparing a firm that has two each of the three types of measures versus one that has three measures each from only two categories. Moreover, objective nonfinancial measures

[9] The distribution of measures across measurement types (financial, objective nonfinancial, and subjective) is not significantly different between the U.S. and Belgian samples. Hence, for parsimony, we only report the results based on the combined sample. Moreover, when we include a dummy variable for the sample firms' national origin in the subsequent tests, we find it to be uniformly insignificant, with no qualitative impact on the results.

such as productivity and other efficiency measures are in some sense ''proxies'' for financial measures (e.g., they are more timely measures, and thus, leading proxies for financial measures), whereas subjective measures add an extra dimension to performance measurement. Our *C_DIVERSITY* measure reflects this distinction by differentiating between measurement systems that use *financial and objective nonfinancial measures* (=2) versus those that use *financial and subjective measures* (=3).

Manufacturing Strategy

We derive our measure of quality-based manufacturing strategy primarily from Abernethy and Lillis (1995), Chenhall (1997), Ittner and Larcker (1995), and Perera et al. (1997), supplemented by a wider reading of the academic and practitioner literatures (Banker et al. 1993; Hackman and Wageman 1995; Saraph et al. 1989). In total, the quality strategy scale includes the seven items shown in Table 2. These items capture employee involvement, process improvements, and cross-departmental coordination, which have been

TABLE 2
Measures of Quality Strategy and Performance
(n = 128)

Quality Strategy (*QUALSTRAT*)
(Eigenvalue = 3.01; Cronbach α = 0.77; μ = 3.55, σ = 0.63)

In your department, to what extent:[a]	**Factor Loading**	**Alpha if Item Deleted**
1. Are nonmanagement employees evaluated for quality performance?	0.67	0.75
2. Do nonmanagement employees participate in quality improvement decisions?	0.70	0.74
3. Is building awareness about quality among non-management employees ongoing?	0.73	0.73
4. Are quality performance data displayed at employee work stations/areas?	0.56	0.76
5. Are suggestion programs for quality improvement among non-management employees used?	0.70	0.73
6. Are programs in place to improve cycle-times (e.g., by reducing time-delays or non-value-added activities in manufacturing)?	0.64	0.74
7. Are programs in place to coordinate quality improvements with other departments within the organization?	0.57	0.75

Performance (*ALLPERF*)
(Eigenvalue = 2.30; Cronbach α = 0.74; μ = 3.67, σ = 0.64)

Please rate the following performances of your department (relative to the industry average):[b]	**Factor Loading**	**Alpha if Item Deleted**
1. Financial performance of your department?	0.68	0.73
2. Operating performance of your department?	0.84	0.62
3. Employee-oriented performance of your department?	0.73	0.70
4. Customer-oriented performance of your department?	0.78	0.68

Output of factor analysis with principal component extraction and oblique rotation (δ = 0).
[a] Anchored as (1) not at all, (2) low extent, (3) medium extent, (4) high extent, and (5) very high extent.
[b] Anchored as (1) well below average; (2) below average; (3) average; (4) above average; and (5) well above average.

consistently identified in prior studies as critical aspects of quality initiatives (e.g., Ittner and Larcker 1995). Empirically, all seven items load on one factor (Table 2). Accordingly, we create a composite variable (*QUALSTRAT*) by averaging the seven item scores. *QUALSTRAT* ranges from 1.86 to 4.86 ($\mu = 3.55$, $\sigma = 0.63$), with an acceptable Cronbach α of 0.77.

Performance

Following the performance measurement categories used throughout this study, we measure performance of the manufacturing department along the four dimensions shown in Table 2: financial, operating, employee-oriented, and customer-oriented. All four performance categories are highly positively correlated ($p < 0.01$), and factor analysis holds all four items together on one factor with high loadings (Table 2). Hence, we construct an overall performance index (*ALLPERF*) by arithmetically averaging the individual item scores ($\mu = 3.67$, $\sigma = 0.64$, Cronbach $\alpha = 0.74$).

RESULTS

The Performance Effect of Measurement Diversity

Hypothesis 1 posits a positive association between firm performance and performance measurement diversity, regardless of strategy. Table 3, Panel A, shows positive and significant correlations between performance and the total number of all types of measures combined, as well as separately with the numbers of objective and subjective nonfinancial performance measures. Notably, performance is not significantly correlated with the number of financial measures. Our categorical diversity measure is also positively and significantly correlated with performance.

Panel B of Table 3 shows that we obtain the same results by regressing performance against the number of all measures combined (Regression (1)), and separately for each performance measurement type (Regression (2)) and the diversity measure (Regression (3)). All three regressions control for strategy, which is significant, but does not alter the inferences related to the performance measurement variables. In other words, the results in Panel B support the beneficial effects of performance measurement diversity regardless of, or after controlling for, strategy. The regressions also control for size (a broad proxy for other potential omitted correlated variables), which is not significant in any of the three regressions.

In total, these results support the existence of performance benefits from increasing performance measurement diversity (H1), particularly by expanding the use of objective and subjective nonfinancial measures. Moreover, the lack of a significant positive association between the number of financial performance measures and performance perhaps suggests that, due to their traditional preeminence, financial measures form a common floor for firms' performance measurement systems, upon which distinct advantages have to be constructed out of the less traditional nonfinancial measures.

The Performance Effect of Strategy-Measurement Alignment

Hypothesis 2 posits that firms pursuing more of a quality-based manufacturing strategy will rely more on objective and subjective nonfinancial performance measures. Table 4, Panel A, shows positive and significant correlations between quality strategy and the number of all types of measures, although the strongest correlations appear to be between quality strategy and nonfinancial and subjective measures, respectively. Our categorical diversity measure is also positively and significantly correlated with quality strategy. To provide a

TABLE 3
Association between Measurement Diversity and Performance
(n = 128)

Panel A: Correlation Test[a]

Number of financial measures (*#_FIN*)	0.06
Number of objective nonfinancial measures (*#_NONFIN*)	0.29***
Number of internal operating measures (*#_OPS*)	0.18**
Number of employee-oriented measures (*#_EMPL*)	0.27***
Number of customer-oriented measures (*#_CUST*)	0.23***
Number of subjective measures (*#_SUBJ*)	0.38***
Total number of measures (*#_TOTAL*)	0.30***
Categorical measure of performance measurement diversity (*C_DIVERSITY*)	0.36***

Panel B: Regression Test[b]

	(1)	(2)	(3)
Intercept	2.85 (0.25)***	2.47 (0.27)***	2.45 (0.24)***
#_FIN		−0.01 (0.03)	
#_NONFIN		0.03 (0.02)*	
#_SUBJ		0.13 (0.04)***	
#_TOTAL	0.02 (0.01)**		
C_DIVERSITY			0.23 (0.06)***
QUALSTRAT	0.23 (0.13)*	0.15 (0.08)**	0.34 (0.11)***
SIZE[c]	$-1.4 \cdot 10^{-5}$ ($1.2 \cdot 10^{-4}$)	$3.5 \cdot 10^{-5}$ ($1.2 \cdot 10^{-3}$)	$-5.9 \cdot 10^{-6}$ ($1.1 \cdot 10^{-4}$)
	F = 3.97***	F = 5.22***	F = 9.22***
	Adj. R^2 = 0.08	Adj. R^2 = 0.16	Adj. R^2 = 0.18

See the Appendix for the list of performance measures and their breakdown by category (financial, internal operating, employee-oriented, customer-oriented, and subjective).
See Table 1 for the definition and descriptive statistics of these various performance measure variables.
See Table 2 for the definition and descriptive statistics of the quality strategy (*QUALSTRAT*) and performance (*ALLPERF*) variables.
[a] Panel A reports the correlations between performance (*ALLPERF*) and each performance measure variable. Significance levels are indicated as *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$ (two-tailed).
[b] Panel B reports OLS regressions with performance (*ALLPERF*) as the dependent variable. The reported statistics include the coefficient estimates with standard errors in parentheses. Significance levels are indicated as *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$ (two-tailed).
[c] *SIZE* is measured by the log of the number of employees in the manufacturing department. The results are unaffected when size is measured by either firm sales or employment.

sense of the difference in measurement use between firms that place low versus high emphasis on quality in manufacturing, Panel B shows that firms that place more emphasis on quality in manufacturing do use more of both objective nonfinancial measures (especially ones relating to internal operations and employees) and subjective measures. The inferences based on the categorical diversity measure are qualitatively similar. These results support H2. An interesting aspect of the findings in Table 4, however, is that firms that place greater emphasis on quality in manufacturing do not use correspondingly fewer financial measures; on the contrary, they use more. In other words, they use more of all three types of measures, but especially expand the number of objective and subjective nonfinancial measures.

TABLE 4
Association between Measurement Diversity and Quality Strategy
(n = 128)

Panel A: Correlation Test[a]

Number of financial measures (*#_FIN*)	0.22**
Number of objective nonfinancial measures (*#_NONFIN*)	0.44***
Number of internal operating measures (*#_OPS*)	0.31***
Number of employee-oriented measures (*#_EMPL*)	0.47***
Number of customer-oriented measures (*#_CUST*)	0.25***
Number of subjective measures (*#_SUBJ*)	0.28***
Categorical measure of performance measurement diversity (*C_DIVERSITY*)	0.19*

Panel B: Mean Differences in Performance Measurement Use between Firms with Low versus High Emphasis on Quality in Manufacturing Strategy[b]

	QUALSTRAT		
	Low	**High**	**Difference**
Number of financial measures (*#_FIN*)	5.77	6.78	$p = 0.05$
Number of objective nonfinancial measures (*#_NONFIN*)	12.55	15.66	$p < 0.01$
Number of internal operating measures (*#_OPS*)	5.88	6.88	$p = 0.01$
Number of employee-oriented measures (*#_EMPL*)	3.24	4.65	$p < 0.01$
Number of customer-oriented measures (*#_CUST*)	3.50	4.06	$p = 0.13$
Number of subjective measures (*#_SUBJ*)	4.25	4.92	$p = 0.04$
Categorical measure of performance measurement diversity (*C_DIVERSITY*)	2.92	3.45	$p = 0.02$

See the Appendix for the list of performance measures and their breakdown by category (financial, internal operating, employee-oriented, customer-oriented, and subjective).
See Table 1 for the definition and descriptive statistics of these various performance measure variables.
See Table 2 for the definition and descriptive statistics of the quality strategy variable (*QUALSTRAT*).
[a] Panel A reports the correlations between quality strategy (*QUALSTRAT*) and each performance measure variable. Significance levels are indicated as *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$ (two-tailed).
[b] Panel B reports the mean values, and the significance of their difference (two-tailed p-values), across two types of quality strategy (low versus high emphasis on quality in manufacturing), which is the quality strategy variable from Table 2 (*QUALSTRAT*) dichotomized at the median.

Hypothesis 3 posits that firms that pursue more of a quality-based manufacturing strategy *and* make greater use of objective and subjective nonfinancial measure will have higher performance. Table 5 presents the results of two regressions testing this prediction. Of pivotal importance in this table are the interaction terms obtained by crossing quality-based manufacturing strategy with (1) the number of objective and subjective nonfinancial measures, respectively, and (2) the categorical diversity measure. These models also include the main effects, as well as a control variable for size.

The results in Model [1] indicate that firms that place greater emphasis on quality in their manufacturing strategies and make greater use of subjective measures have higher performance. However, there is no similar effect from the interaction between quality strategy and the use of objective nonfinancial measures. Model [2] shows a significant effect on performance from the interaction of quality strategy and the categorical diversity measure, which was accorded a higher numerical value for firms using subjective rather than objective nonfinancial measures. In total, the results in Table 5 support H3 for subjective measures, but not for objective nonfinancial measures.

TABLE 5
Performance Effects of the Alignment between Quality Strategy and Performance Measurement
(n = 128)

	(1)	(2)
Intercept	2.97 (0.69)***	1.68 (0.35)***
QUALSTRAT[a]	0.04 (0.19)	0.37 (0.09)***
#_NONFIN[b]	0.03 (0.03)	
#_SUBJ[b]	0.05 (0.11)	
C_DIVERSITY[c]		0.03 (0.14)
QUALSTRAT · #_NONFIN	−0.01 (0.01)	
QUALSTRAT · #_SUBJ	0.08 (0.04)*	
QUALSTRAT · C_DIVERSITY		0.06 (0.03)*
SIZE[d]	$-2.8^{\cdot 10^{-5}}$ ($1.1^{\cdot 10^{-4}}$)	$2.1^{\cdot 10^{-6}}$ ($1.2^{\cdot 10^{-4}}$)
	F = 4.62***	F = 9.72***
	Adj. R^2 = 0.17	Adj. R^2 = 0.25

OLS regressions with overall performance (*ALLPERF*) as the dependent variable (see Table 2). The reported statistics include the coefficient estimates with standard errors in parentheses. Significance levels are indicated as *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$ (two-tailed).
[a] *QUALSTRAT* as defined in Table 2.
[b] Number of nonfinancial (*#_NONFIN*) and subjective (*#_SUBJ*) measures, as reported in Table 1.
[c] Categorical measure of performance measurement diversity. See Table 1 for definition and descriptive statistics.
[d] *SIZE* is measured by the log of the number of employees in the manufacturing department. The results are unaffected when size is measured by either firm sales or employment.

The mixed findings for H3 motivated us to undertake an additional test, shown in Table 6. Following Ittner, Larcker, and Randall's (2003) approach, we derive benchmarks of performance measurement system and strategy alignment by separately regressing (1) the overall number of measures, (2) the number of financial measures, (3) the number of objective nonfinancial measures, (4) the number of subjective measures, and (5) the categorical diversity measure on the quality strategy measure. This allows us to obtain a residual from each regression for each firm. A positive (negative) residual in each case indicates the extent to which a firm uses *too many* (*too few*) measures, as compared to firms following a similar quality-based manufacturing strategy. The implication of the contingency view is that both types of *mis*alignment between the performance measurement system and strategy will be negatively associated with performance. For ease of interpretation, and in line with Ittner, Larcker, and Randall (2003), we take the absolute values of the negative residuals so that both positive and negative measurement *mis*alignment would produce a negative coefficient for performance.

Table 6 shows that using fewer measures (Models (1) through (4)) or having a lower categorical diversity measure score (Model (5)) than firms following similar strategies is negatively and significantly associated with performance (except in Model (3) where $p = 0.14$). These results are consistent with contingency theory (and the findings of Said et al. [2003] discussed earlier). However, there is no similar effect due to more extensive measurement. In fact, the sign of the positive measurement residuals is positive in all models (opposite to the direction predicted), and significantly so in Models (4) and (5). This last

TABLE 6
Association between Performance and the Alignment of Measurement Practices to Quality Strategy
(n = 128)

	(1)	(2)	(3)	(4)	(5)
Intercept	3.66 (0.09)***	3.65 (0.08)***	3.66 (0.08)***	3.64 (0.09)***	3.72 (0.09)***
Positive *#_TOTAL* residual[a]	0.23 (0.21)				
Negative *#_TOTAL* residual[a]	−0.31 (0.19)*				
Positive *#_FIN* residual[a]		0.19 (0.22)			
Negative *#_FIN* residual[a]		−0.33 (0.18)*			
Positive *#_NONFIN* residual[a]			0.24 (0.20)		
Negative *#_NONFIN* residual[a]			−0.26 (0.18)		
Positive *#_SUBJ* residual[a]				0.40 (0.17)**	
Negative *#_SUBJ* residual[a]				−0.32 (0.19)*	
Positive *C_DIVERSITY* residual[b]					0.28 (0.17)*
Negative *C_DIVERSITY* residual[b]					−0.47 (0.20)**
	F = 3.64**	F = 3.24**	F = 2.75*	F = 7.03***	F = 7.25***
	Adj. R^2 = 0.04	Adj. R^2 = 0.04	Adj. R^2 = 0.03	Adj. R^2 = 0.09	Adj. R^2 = 0.09

OLS regressions with overall performance (*ALLPERF*) as the dependent variable (see Table 2). The reported statistics include the coefficient estimates with standard errors in parentheses. Significance levels are indicated as *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$ (two-tailed).

[a] We obtain the measurement residual variables associated with the number of measures from (1) regressing the total number of measures (*#_TOTAL*) on *QUALSTRAT*; (2) regressing the number of financial measures (*#_FIN*) on *QUALSTRAT*; (3) regressing the number of objective nonfinancial measures (*#_NONFIN*) on *QUALSTRAT*; and (4) regressing the number of subjective measures (*#_SUBJ*) on *QUALSTRAT*. Negative measurement residuals are converted to absolute values.

[b] We obtain the measurement residual variables for the categorical diversity measure from regressing *C_DIVERSITY* on *QUALSTRAT*. Negative measurement residuals are converted to absolute values.

result is consistent with Ittner, Larcker, and Randall (2003), but contrary to Said et al. (2003) and contingency theory-based expectations.

DISCUSSION

We find that firms with more extensive performance measurement systems, especially ones that include objective and subjective nonfinancial measures, have higher performance. This finding holds regardless of the firm's manufacturing strategy and, as such, provides general support for the assertion that increasing the number of performance measures, per se, benefits performance.

But we also find evidence that partially supports the performance alignment view, namely, that performance measurement has to fit strategy, and that the strategy-measurement "fit" affects performance. Specifically, we find that firms pursuing a quality-based manufacturing strategy make more extensive use of both objective and subjective nonfinancial measures. In turn, we find a positive relationship between the strategy-measure pairing and firm performance when quality-based manufacturing strategies are combined with extensive use of subjective measures. However, there is no similar positive performance effect from more extensive use of objective nonfinancial measures.

The latter results are consistent with Ittner and Larcker (1995) who found that quality programs are associated with greater use of nontraditional (i.e., nonfinancial) measures and reward systems, but combining nontraditional measures with extensive quality programs does not improve performance. However, by differentiating between objective and subjective nonfinancial measures—thereby going beyond Ittner and Larcker (1995)—we find that performance is higher when the performance measures used in conjunction with a quality-based manufacturing strategy are of the subjective type.

Finally, in a different test of the contingency view that performance depends on the match between performance measurement and strategy, we find that among firms with similar quality-based strategies, those with *less* extensive performance measurement systems have lower performance, whereas those with *more* extensive performance measurement systems do not. In the case of subjective performance measures, firms that use them more extensively than firms with similar quality-based strategies actually have significantly higher performance.

Collectively, then, our findings provide stronger support for the performance measurement diversity than contingency/alignment view. First, using more objective and subjective nonfinancial measures appears to enhance performance, even in firms with relatively low emphasis on quality in manufacturing. Second, considering the match between performance measurement and strategy, our results suggest that using fewer measures than firms with similar quality-based manufacturing strategies hurts performance, whereas using more does not.

Our findings, however, should be interpreted with the following caveats in mind. First, our method and tests implicitly assume that the firms in our sample are in equilibrium. But, firms in our sample may have been in the process of modifying their strategies and performance measurement systems, and in the transition process there may be firms that achieve better matches (hence, performance) than others (Moores and Yuen 2001). This may explain, at least partly, the mixed support for the contingency view of performance measurement in our, but also prior, work (Ittner, Larcker, and Randall 2003; Said et al. 2003). Cross-sectional, contemporaneous data cannot totally rule out this possibility. Future research that uses longitudinal data and/or is able to capture identifiable changes in firms' strategies and measurement systems would help to address this issue.

Second, this study used a survey to facilitate access to a relatively broad sample and data that are not publicly available. But our data limit the scope and depth of exploration that is feasible. For example, with these data, we cannot address *why* subjectivity is associated with higher performance in firms that emphasize quality in manufacturing, which we feel is an important avenue for future research. Despite these limitations, however, we contribute to the understanding of the role of performance measurement by making a finer distinction within and between performance measurement types, and relating these distinctions to discussions of performance measurement diversity and alignment.

APPENDIX
Survey Section on Performance Measures

With respect to the current performance measurement system in place, please check all the measures that are used by upper management to gauge your department's performance.

(*Please only check (or write in) those measures that are reported, analyzed, and discussed on a regular basis for the purpose of performance measurement and evaluation.*)

1. FINANCIAL PERFORMANCE MEASURES FOR YOUR DEPARTMENT (Check **all** that apply.)

- ☐ Deployment of assets (e.g., ROI) in your department
- ☐ *Total* gross margin or contribution margin of your department
- ☐ *Unit* gross margin or contribution margin (per individual product or product category)
- ☐ *Total* manufacturing cost budget
- ☐ *Unit* manufacturing cost (per individual product or product category)
- ☐ Manufacturing cost budget "line-items," such as:
  - ☐ Labor cost variances
  - ☐ Material cost variances
  - ☐ Indirect cost (overhead) variances
  - ☐ Maintenance expenditures
- ☐ Dollar amount spent on manufacturing process improvements

Other (please list):

2.1. INTERNAL OPERATING PERFORMANCE MEASURES FOR YOUR DEPARTMENT (Check **all** that apply.)

- ☐ Production volume
- ☐ Labor productivity (e.g., hours used/hours available, overtime hours)
- ☐ Machine productivity (e.g., hours running/hours available, downtime on equipment)
- ☐ Material usage (e.g., material usage inefficiency, material waste)
- ☐ Setup efficiency (e.g., setup time, number of setups)
- ☐ Manufacturing cycle time (e.g., total process time)
- ☐ Inventory (e.g., inventory turnover)
- ☐ Product defects (e.g., number of errors, rework, scrap)
- ☐ New product introductions (e.g., total number, percentage of sales from new products)
- ☐ New product-design efficiency (e.g., time to develop new products, on-time schedule)

Other (please list):

2.2. EMPLOYEE-ORIENTED MEASURES FOR YOUR DEPARTMENT (Check **all** that apply.)

☐ Employee satisfaction (e.g., results of employee surveys, number of grievances filed)
☐ Employee skills (e.g., level of education, level of experience)
☐ Employee empowerment (e.g., number of suggestions submitted, % of employees on improvement teams)
☐ Safety (e.g., number of accidents, number of injuries)
☐ Employee training/education (e.g., number of hours or % of employees' time allocated for training)
☐ Employee loyalty/turnover (e.g., years in job, years with firm)
☐ Absenteeism
Other (please list):

2.3. CUSTOMER-ORIENTED PERFORMANCE MEASURES FOR YOUR DEPARTMENT (Check **all** that apply.)

☐ Market share
☐ Time to fill customer orders
☐ Delivery performance (e.g., on-time delivery, percent of correct delivery)
☐ Time to respond to customer problems
☐ Flexibility/responsiveness (i.e., ability to vary product characteristics)
☐ Customer satisfaction (e.g., results from customer surveys, number of customer complaints)
☐ Customer acquisition (e.g., number of new customers, percent sales from new customers)
☐ Customer retention/loyalty (e.g., number of repeat customers)
Other (please list):

3. SUBJECTIVE PERFORMANCE MEASURES FOR YOUR DEPARTMENT

The above categories dealt with the performance measures that are quantified and reported on a regular basis. However, performance evaluations by upper management also may include subjective assessments of various, not always clearly specified, aspects of performance. Which of the following factors do you believe upper management makes an assessment of when evaluating your performance? (Check **all** that apply.)

☐ My long-term perspective on the business
☐ My ability to effectively acquire new skills/knowledge
☐ My willingness to share knowledge within the organization
☐ My cooperation with other departments within the organization
☐ Employee spirit/morale in my department
☐ My management style/leadership skills
☐ My loyalty toward the firm
Other (please list):

■ The above measures can be classified into three categories: (1) **financial performance measures** (from Section 1 above); (2) **objective nonfinancial performance measures** (Sections 2.1, 2.2, and 2.3); and (3) **subjective performance measures** (Section 3), which is the terminology we use in the questions that follow.

REFERENCES

Abernethy, M. A., and A. M. Lillis. 1995. The impact of manufacturing flexibility on management control system design. *Accounting, Organizations and Society* 20 (4): 241–258.

Baiman, S., and M. V. Rajan. 1995. The informational advantage of discretionary bonus schemes. *The Accounting Review* 70 (4): 557–579.

Baker, G. 1992. Incentive contracts and performance measurement. *Journal of Political Economy* 100 (3): 598–614.

———, R. Gibbons, and K. J. Murphy. 1994. Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics* 109: 1125–1156.

Balkcom, J. E., C. D. Ittner, and D. F. Larcker. 1997. Strategic performance measurement: Lessons learned and future directions. *Journal of Strategic Performance Measurement* 1 (2): 22–32.

Banker, R. D., G. Potter, and R. G. Schroeder. 1993. Reporting manufacturing performance measures to workers: An empirical study. *Journal of Management Accounting Research* 5: 33–55.

———, ———, and D. Srinivasan. 2000. An empirical investigation of an incentive plan that includes nonfinancial performance measures. *The Accounting Review* 75 (1): 65–92.

Bushman, R. M., R. J. Indjejikian, and A. Smith. 1996. CEO compensation: The role of individual performance evaluation. *Journal of Accounting and Economics* 21: 161–193.

Butler, A., S. Letza, and B. Neale. 1997. Linking the balanced scorecard to strategy. *Long Range Planning* 30 (2): 242–253.

Campbell, D., S. Datar, S. Kulp, and V. G. Narayanan. 2004. Using the balanced scorecard as a control system for monitoring and revising corporate strategy. Working paper, Harvard Business School.

Campbell, J. P. 1990. Modeling the performance prediction problem in industrial and organizational psychology. In *Handbook of Industrial and Organizational Psychology*, Vol. 1, 2nd edition, edited by M. D. Dunnette, and L. M. Hough, 687–732. Palo Alto, CA: Consulting Psychologists Press.

Chenhall, R. H. 1997. Reliance on manufacturing performance measures, total quality management and organizational performance. *Management Accounting Research* 8 (2): 187–206.

———, and K. Langfield-Smith. 1998. The relationship between strategic priorities, management techniques and management accounting: An empirical investigation using a systems approach. *Accounting, Organizations and Society* 23 (3): 243–264.

———. 2003. Management control system design within its organizational context: Findings from contingency-based research and directions for the future. *Accounting, Organizations and Society* 28 (2/3): 127–168.

Chow, C. W., K. M. Haddad, and J. E. Williamson. 1997. Applying the balanced scorecard to small companies. *Management Accounting* 79 (2): 21–27.

Clinton, B. D., and K. C. Hsu. 1997. JIT and the balanced scorecard: Linking manufacturing control to management control. *Management Accounting* 79 (3): 18–24.

Daniel, S., and W. Reitsperger. 1991. Linking quality strategy with management control systems: Empirical evidence from Japanese industry. *Accounting, Organizations and Society* 16 (7): 601–618.

Datar, S., S. Kulp, and R. Lambert. 2001. Balancing performance measures. *Journal of Accounting Research* 39 (1): 75–92.

Dillman, D. 1978. *Mail and Telephone Surveys: The Total Design Method*. New York, NY: Wiley.

———. 1999. *Mail and Internet Surveys: The Tailored Design Method*. New York, NY: Wiley.

Dixon, J. R., A. J. Nanni, and T. E. Vollmann. 1990. *The New Performance Challenge: Measuring Operations for World-Class Companies*. Homewood, IL: Dow Jones-Irwin.

Edvinsson, L., and M. Malone. 1997. *Intellectual Capital: Realizing Your Company's True Value by Finding its Hidden Brainpower*. New York, NY: Harper Business.

Epstein, M., and J. F. Manzoni. 1997. The balanced scorecard and "tableau-de-bord." *Management Accounting* 79 (August): 28–36.

Feltham, G. A., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal-agent relations. *The Accounting Review* 69 (3): 429–453.

Fisher, J. 1992. Use of nonfinancial performance measures. *Journal of Cost Management* 6 (1): 31–38.

———. 1995. Contingency-based research on management control systems: Categorization by level of complexity. *Journal of Accounting Literature* 14: 24–53.

Fulk, J., A. P. Brief, and S. H. Barr. 1985. Trust-in-supervisor and perceived fairness and accuracy of performance evaluations. *Journal of Business Research* 13 (August): 301–313.

Gibbs, M., K. A. Merchant, W. A. Van der Stede, and M. E. Vargus. 2004. Determinants and effects of subjectivity in incentives. *The Accounting Review* 79 (2): 409–436.

Ghosh, D., and R. F. Lusch. 2000. Outcome effect, controllability and performance evaluation of managers: Some evidence from multi-outlet businesses. *Accounting, Organizations and Society* 25 (4/5): 411–425.

Hackman, J. R., and R. Wageman. 1995. Total quality management: Empirical, conceptual, and practical issues. *Administrative Science Quarterly* 40 (2): 309–342.

Hawkins, S., and R. Hastie. 1990. Hindsight-biased judgments of past events after the outcomes are known. *Psychological Bulletin* 107 (May): 311–327.

Hemmer, T. 1996. On the design and choice of ''modern'' management accounting measures. *Journal of Management Accounting Research* 8: 87–116.

Heneman, R. L. 1986. The relationship between supervisory ratings and results-oriented measures of performance: A meta analysis. *Personnel Psychology* 39 (4): 811–826.

Hoffecker, J., and C. Goldenberg. 1994. Using the balanced scorecard to develop company-wide performance measures. *Journal of Cost Management* 8 (3): 5–17.

Holmstrom, B. 1979. Moral hazard and observability. *Bell Journal of Economics* 10: 74–91.

———, and P. Milgrom. 1991. Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization* 7: 24–52.

Hopwood, A. G. 1974. *Accounting and Human Behavior*. London, U.K.: Haymarket Publishing.

Hoque, Z., and W. James. 2000. Linking balanced scorecard measures to size and market factors: Impact on organizational performance. *Journal of Management Accounting Research* 12: 1–17.

Institute of Management Accountants (IMA). 1996. *Are Corporate America's Financial Measures Outdated?* Montvale, NJ: IMA.

Ittner, C. D., and D. F. Larcker. 1995. Total quality management and the choice of information and reward systems. *Journal of Accounting Research* 33 (Supplement): 1–34.

———, and ———. 1997. Quality strategy, strategic control systems, and organizational performance. *Accounting, Organizations and Society* 22 (3/4): 293–314.

———, and ———, and M. V. Rajan. 1997. The choice of performance measures in annual bonus contracts. *The Accounting Review* 72 (2): 231–255.

———, and ———. 1998a. Innovations in performance measurement: Trends and research implications. *Journal of Management Accounting Research* 10: 205–38.

———, and ———. 1998b. Are nonfinancial measures leading indicators of financial performance: An analysis of customer satisfaction. *Journal of Accounting Research* 36 (Supplement): 1–35.

———, and ———. and M. Meyer. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review* 78 (3): 725–758.

———, and ———. and T. Randall. 2003. Performance implications of strategic performance measures in financial services firms. *Accounting, Organizations and Society* 28 (7/8): 715–741.

Kaplan, R. S., and D. P. Norton. 1992. The balanced scorecard: Measures that drive performance. *Harvard Business Review* 70 (1): 71–79.

———, and ———. 1993. Putting the balanced scorecard to work. *Harvard Business Review* 71 (5): 134–142.

———, and ———. 1996a. Using the balanced scorecard as a strategic management system. *Harvard Business Review* 74 (1): 75–85.

———, and ———. 1996b. *The Balanced Scorecard: Translating Strategy into Action*. Boston, MA: Harvard Business School Press.

———, and ———. 1996c. Linking the balanced scorecard to strategy. *California Management Review* 39 (1): 53–79.

———, and ———. 2000. Having trouble with your strategy: Then map it. *Harvard Business Review*, 78 (5): 167–176.

———, and ———. 2001. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, MA: Harvard Business School Press.

Lambert, R. A. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32 (1/3): 3–87.

Langfield-Smith, K. 1997. Management control systems and strategy: A critical review. *Accounting, Organizations and Society* 22 (2): 207–232.

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

Lingle, J., and W. Schiemann. 1996. From balanced scorecard to measurement. *Management Review* (March): 56–61.

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

———, and ———. 2002. A note on the judgment effects of the balanced scorecard's information organization. *Accounting, Organizations and Society* 27 (6): 531–540.

MacLeod, W. B., and D. Parent. 1999. Job characteristics and the form of compensation. *Research in Labor Economics* 18: 177–242.

Maisel, L. S. 1992. Performance measurement: The balanced scorecard approach. *Journal of Cost Management* 6 (2): 47–52.

Milgrom, P. 1988. Employment contracts, influence activities, and efficient organization design. *Journal of Political Economy* 96 (February): 42–60.

Moers, F. 2005. Discretion and bias in performance evaluation: The impact of diversity and subjectivity. *Accounting, Organizations and Society* 30 (1): 67–80.

Moores, K., and S. Yuen. 2001. Management accounting systems and organizational configuration: A life-cycle perspective. *Accounting, Organizations and Society* 26 (4/5): 351–389.

Murphy, K. J., and P. Oyer. 2004. Discretion in executive incentive contracts: Theory and evidence. Working paper, University of Southern California and Stanford University.

Neely, A. 1999. The performance measurement revolution: Why now and what next? *International Journal of Operations and Production Management* 19 (2): 205–228.

Perera, S., G. Harrison, and M. Poole. 1997. Customer-focused manufacturing strategy and the use of operation-based nonfinancial performance measures: A research note. *Accounting, Organizations and Society* 22 (6): 557–572.

Prendergast, C. 1993. A theory of "yes men." *American Economic Review* 83 (September): 757–770.

———, and R. Topel. 1993. Discretion and bias in performance evaluation. *European Economic Review* 37: 355–365.

———, and ———. 1996. Favoritism in organizations. *Journal of Political Economy* 104: 958–978.

Provost, L., and S. Leddick. 1993. How to take multiple measures to get a complete picture of organizational performance. *National Productivity Review* (Autumn): 477–490.

Ruddle, K., and D. Feeny. 2000. *Transforming the Organization: New Approaches to Management, Measurement, and Leadership.* Oxford, U.K.: Templeton College.

Said, A. A., H. R. HassabElnaby, and B. Wier. 2003. An empirical investigation of the performance consequences of nonfinancial measures. *Journal of Management Accounting Research* 15: 193–223.

Saraph, J. V., P. G. Benson, and R. G. Schroeder. 1989. An instrument for measuring the critical factors of quality management. *Decision Sciences* 20: 810–829.

Scott, T. W., and P. Tiessen. 1999. Performance measurement and managerial teams. *Accounting, Organizations and Society* 24 (3): 263–286.

Stivers, B. P., T. J. Covin, N. G. Hall, and S. W. Smalt. 1998. How nonfinancial measures are used. *Management Accounting* 79 (February): 44–49.

Van der Stede, W. A., S. M. Young, and X. C. Chen. 2005. Assessing the quality of evidence in empirical management accounting research: The case of survey studies. *Accounting, Organizations and Society* 30 (7/8): 655–684.