

The Evolution Of A Balanced Scorecard

Mary A. Malina, Ph.D., University of Colorado Denver, USA

ABSTRACT

Both professionals and academics have long criticized the use of traditional financial performance measures and called for balance in performance measurement systems. In 1992, Kaplan and Norton introduced the Balanced Scorecard and it has been adopted widely around the world and offered as a superior combination of nonfinancial and financial measures of performance. This paper is the result of a 15-year field study of a Fortune 500 company's Balanced Scorecard. Both qualitative and quantitative data were collected to address the following research questions with respect to the Balanced Scorecard: 1) What has changed over time? 2) What has not changed over time? 3) Why has it endured? Changes highlighted are that the Balanced Scorecard was unaffected by a major change in organizational structure, a narrowing of focus and reduction in the scope over time, processes for changing the design were formalized, and that it has become engrained in the compensation system. Factors that have remained constant over time are the purpose of the Balanced Scorecard, its use for relative performance evaluation, and its use as a tool for best practice sharing. Two factors that appear to explain why it has endured are its use as a learning and communication tool and its ability to influence behavior. The paper concludes with a list of key success factors for building and sustaining a successful scorecard. This list might also be helpful to researchers seeking to investigate the design, use or impact of a Balanced Scorecard.

Keywords: Balanced Scorecard; Performance Measurement; Performance Management; Performance Measures

1. INTRODUCTION

Both professionals and academics have long criticized the use of traditional performance measures (e.g. *ex post* costs, profits, return on investment) as insufficient guides for decision making in today's business environment. Sole reliance on current, financial measures of performance does not arguably reflect the importance of current resource decisions for future financial performance. Though some firms recognized the importance of nonfinancial measures of performance years ago (e.g. General Electric in the 1950s), growing international competition and the rise of the total quality management (TQM) movement have widened the appeal of nonfinancial performance measures. Since the 1980s, authors have filled the professional and academic literature with recommendations to rely more on nonfinancial measures for both managing and evaluating organizations (e.g., Johnson and Kaplan, 1987; Dixon et al., 1990; Rappaport 1999). The call for balance in performance measurement systems was answered by Kaplan and Norton in 1992 with the introduction of the Balanced Scorecard (BSC). Since then, the BSC has been adopted widely around the world and has been offered as a superior combination of nonfinancial and financial measures of performance. Because the BSC explicitly focuses on links among business decisions and outcomes, it is intended to guide strategy development, implementation, and communication. Furthermore, a properly constructed BSC could provide reliable feedback for management control and performance evaluation.

During the summer of 1999, access was gained to study a newly-implemented BSC in a *Fortune 500* manufacturing firm. As of year-end 2011, the company employs over 40,000 employees world-wide and has over \$18 billion in sales. Competition is from similarly large domestic and international firms. The corporate office developed and introduced the quarterly distributor balanced scorecard (DBSC) to be used in evaluating and improving performance of its North American distribution network. The company sells its major equipment assemblies through OEM contracts and exclusive distributors, who also sell repair parts and maintenance services. Distributors provide maintenance contracts, engineering services and integrated products, and customize the

company products to cater to specific needs of end-users. Each distributor operates within an assigned geographic area and may not compete with each other.

When first introduced, the stated purposes of the DBSC were to 1) highlight areas within distributorships that need improvement to enhance customer relations, 2) provide an objective set of criteria, consistent with the company's new strategic initiatives, to guide and measure total distributor performance, 3) be used as the starting point for the three-year contract renewal process, and 4) be used for comparing and ranking distributorships and may be used for performance-based compensation. These stated purposes encompass tools for strategic change, comprehensive performance evaluation, relative performance evaluation, and compensation.

In a series of papers, Malina and colleagues studied the short- and mid-term impacts of the DBSC (Malina and Selto, 2001; Malina, Norreklit and Selto, 2007) and reasons for changes made to the structure of early versions of the DBSC (Malina and Selto, 2004). Given the continued interest in the BSC by organizations, academics and practitioners alike, the investigation of the DBSC continued with a return to the firm to check in on their DBSC effort. Whereas nearly all studies of a BSC are single or cross-sectional snapshots of BSC design or impact, this unique access provides an opportunity to study the evolution of a BSC over a 15-year period.

The objective of this paper is to address the following research questions with respect to the DBSC:

RQ1: What has changed over time?

RQ2: What has not changed over time?

RQ3: Why has the DBSC endured?

The data used in this paper was gathered during numerous interactions with the company over the years. In 1999, qualitative data was collected from interviews with three company employees, interviews with nine distributor owners, and internal documents. In addition, quantitative data was collected from company reports on distributors' performance. Follow-up in-person and phone interviews with several company employees occurred in 2000, 2005 and 2012. Most recently, in late 2012, phone interviews were again conducted with fifteen (of 16) current distributor owners.¹ Both quantitative and qualitative data were once again collected from company documents. During both sets of phone interviews with distributor owners, the interviewees were asked open-ended questions about the objectives of the DBSC, the design of the DBSC, the use of the DBSC, and suggestions for improving the DBSC. The length of each telephone interview was determined by how much the interviewee had to say, and all lasted between 45 and 75 minutes.

2. RQ1: WHAT HAS CHANGED OVER TIME?

2.1. Distributor Ownership Structure

Historically, distributors were independently owned but almost exclusively by individuals with employment experience in the company. When the DBSC was introduced, 31 distributors served North America, and all were independently owned. In the early 2000s, the company began implementing a structural change to reduce the number of distributors and to create joint ventures with distributors. The new structure was introduced for two reasons. First, the company wanted to increase the size of distributors by investing more capital in order to compete with competitors' larger distributors. Second, the company wanted to diversify its business to include related services and thereby to reduce the volatility of earnings. Sole reliance on the sales of its signature products caused analyst skepticism regarding the company's long-term financial success. The new structure reduced the number of distributors to 16; all but one of which are now joint ventures averaging 50/50 ownership between distributor and company. Although the company usually chooses not to own more than 50 percent in order to retain the owner's local knowledge, customer focus and entrepreneurial spirit, distributors without extensive experience may start at a lower level of ownership.

¹ Distribution ownership turnover is so frequent that two of 1999 interviewees were still running their distributorships in 2012 and, therefore, were interviewed again.

According to a company executive, this change in the distribution structure has been one of the largest changes at the company in recent history. Transaction cost economics-based literature suggests this expansion of the boundary of the firm by internalizing services could lead to more bureaucratic controls (e.g. Williamson, 1981). Transactions between the company and its distributors that had been market-mediated would now be controlled by bureaucratic means. Perhaps surprisingly, the design and use of the DBSC appears to be unaffected by this historic change. The company and its distributors already used the DBSC extensively, and more controls might have been deemed unnecessary. It is possible, however, that the joint ventures’ new governance structure for the distributorships was a substitute for more controls (e.g., Cohen, Dey, and Lys, 2012). Each joint venture has a Board of Directors comprised of two or three high-level company executives and two distributor employees. The Board oversees strategic development as well as distributor owner and leadership team compensation.

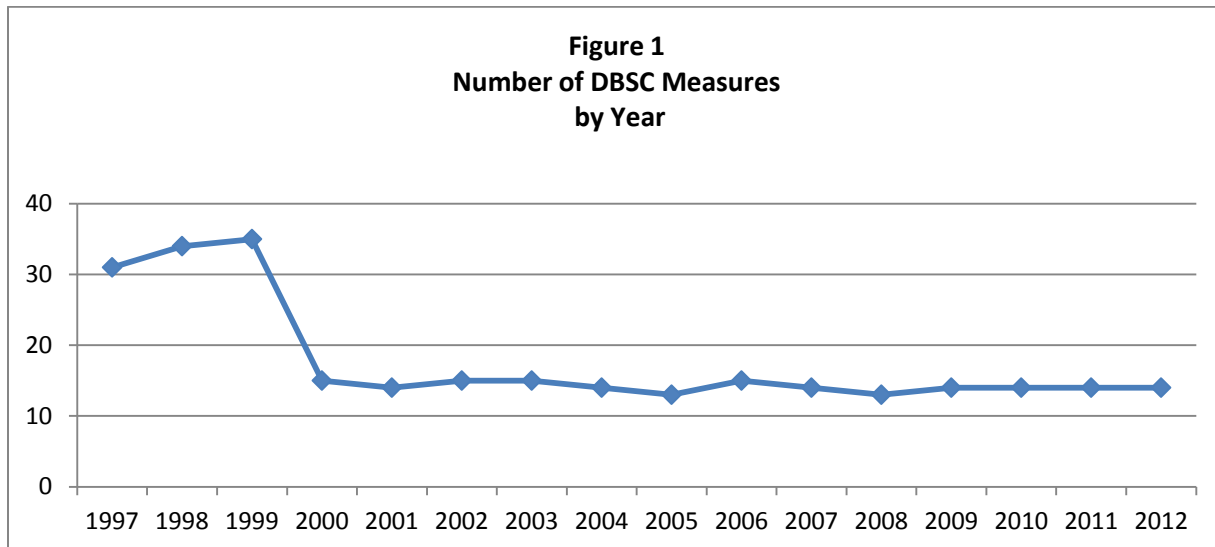
2.2. DBSC Design

Early versions of the DBSC during strategic implementation were very broad, containing up to 35 measures. Fifteen years later, while the distribution strategy has remained unchanged with respect to its focus on customer relations, the company has reduced the scope and integration of the DBSC. Measures have been added or dropped almost yearly. Figure 1 shows the changes in the number of measures over time. After a few years of use, the number of measures dropped by over half. Perhaps DBSC users were suffering from information overload, but Malina and Selto (2004) describe other reasons related to accuracy, objectivity, and cost of administration (including resolution of disputes). For example, distributors commented on measures that were subsequently dropped from the DBSC:

- *“How we measure [product X] market share is strictly information we would generate ourselves”* [Distributor, 1999]
- *“From feedback from other distributors, they are just taking a stab at [the employee training measure]. We actually compile the numbers, but others are getting green scores for just a guess. We’re yellow or red, and it’s a real number. The cost of the time isn’t worth it. But, it’s the right idea and the right thing to do.”* [Distributor, 1999]²

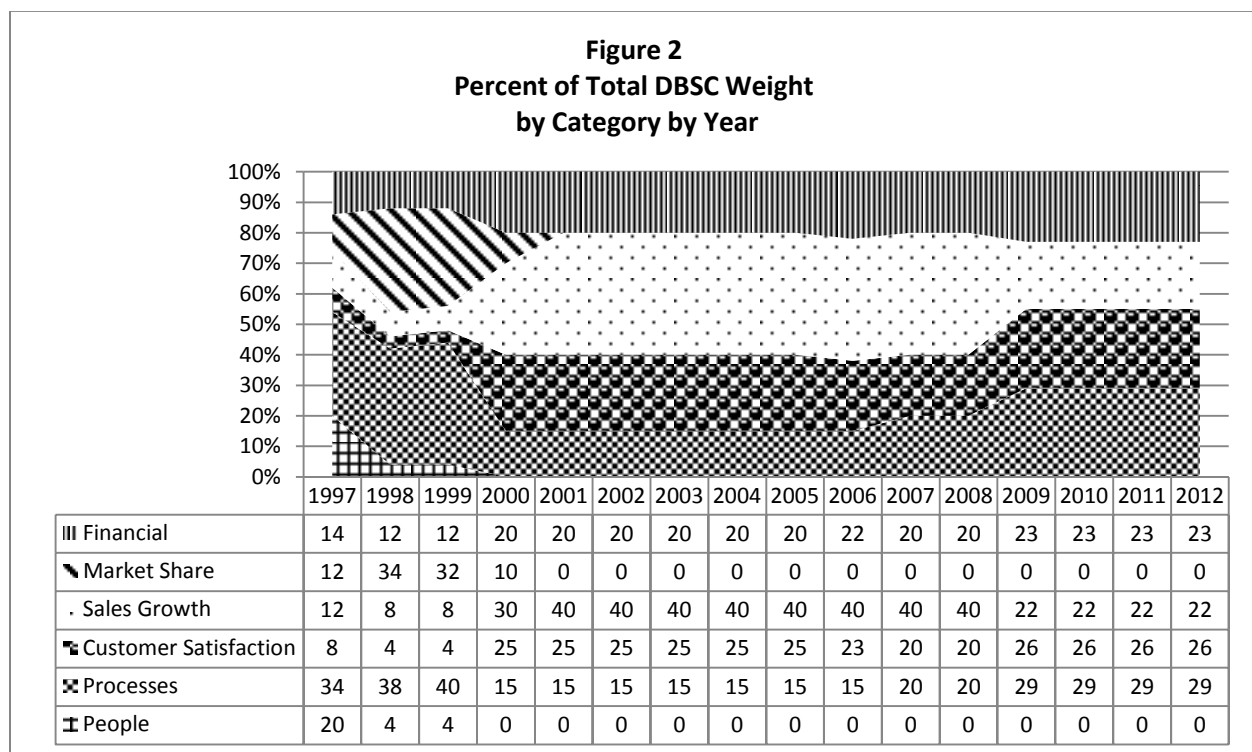
For the past twelve years, the number of DBSC measures has settled between 13 and 15. Now, distributors agree with the reduction in the number of measures over time:

- *“I think it’s down to a good number of metrics.”*[Distributor, 2012]



² The company labels and reports performance on each measure of the DBSC as red for “fails to meet criteria for acceptability,” yellow for “meets criteria for acceptability,” or green for “exceeds criteria for acceptability.”

The design features of the DBSC include a weighting placed on each measure, totaling 100 points. Figure 2 shows the weight and percentage of total weight for each DBSC category.^{3,4} Weight placed on financial measures increased in the early years and has settled at 20-23% over the past 12 years. Early on, market share measures comprised up to one-third of the weight of the DBSC. Within five years, the market share measures were eliminated from the DBSC and their weight was absorbed by sales growth measures. In line with the customer satisfaction strategy, the weight placed on customer satisfaction measures has grown over the life of the DBSC peaking at 26% in the most recent version. The weight placed on internal processes measures is U-shaped, comprising over a third of total weight in early versions, falling to 12% by 2005, then building to 29% for the past four years. The initial DBSC placed 20% of total weight on people measures, which quickly diminished, disappearing in just three years of use. In the past four years the weightings have settled almost equally into the four categories (financial, sales growth, customer satisfaction, and internal processes). Over the years, the weight placed on subjective measures (e.g. people and market share) has been reduced. These changes over time indicate the company’s willingness to sacrifice balance for objectivity.



2.3. Process for Implementing DBSC Design Changes

Initial interviews in 1999 revealed distributor discontent with the process (or lack thereof) by which measures were added to the DBSC. For example:

- “They [the company] added new measures that they didn’t tell us about and then they were red, it’s not a subtle sign that we need to look at that area.” [Distributor, 1999]

Company executives acknowledged the opportunity for improvement:

- “Probably should be communicating better when we change measures.” [Company Employee, 2005]

³ While Figure 2 reflects generic category terminology, the DBSC categories were customized for the company and terminology changed almost yearly.

⁴ The weights placed on measures from 2001-2005 are not available. The 2000 weights were used in the figure.

Today, formal procedures are in place for adding a measure and determining benchmarks for performance. To be added to the DBSC, a measure must have a minimum history of one year of calculated results or have a one-year trial period on the scorecard with zero points assigned so as not to unexpectedly impact the overall score.

Benchmarks for setting the green/yellow/red cutoffs are based on peer comparisons. The starting point for setting the red-yellow-green cutoffs is to calculate the mean and standard deviation for all distributors using the previous three years of data. Then, the green cutoff is set at +1.0 standard deviation and red cutoff at -1.0 standard deviation. Finally, the company qualitatively assesses the cutoffs and modifies them only if they are “unreasonable.” It appears that the company heard the complaints of the distributors and formalized the processes for changing the DBSC, focusing on the objectivity and accuracy of measures while maintaining some subjectivity in setting cutoffs.

2.4. Compensation

As highlighted in the introduction, one of the purposes of the initial DBSC was for compensation. Distributors have two sources of at-risk compensation; triennial contract renewal and annual variable compensation. Before the introduction of the DBSC, the contract renewal process was seen as unprofessional and not data-driven:

- *“Every improvement we can make helps [the distributors] know that they aren’t being judged on personality. I have to admit we’ve done that in the past.”* [Company Employee, 1999]

When introduced, the DBSC was used as the starting point for contract renewal. The DBSC was the objective centerpiece of the contract review. Company management reviewed the DBSC, and distributors had to address any areas in which they fell short. In the early stages of its use, both the company and the distributors focused on the DBSC just prior to the contract renewal timeframe.

The DBSC continues to be used in the three-year contract renewal process, but more extensively. The company uses three years of DBSC information to assess the long-term strengths and weaknesses of the distributor. Teleconference reviews with the Board of Directors are performed after years one and two and a face-to-face review in year three of the contract. Mid-years, they look at action items that were identified from the DBSC. If there is potential for non-renewal, the distributor is aware of it one to two years in advance and there are clear expectations of changes that need to be made in the DBSC for renewal.

- *“The way they work with me is every year we go through an annual business review and every quarter they review the yellow and red items. They ask for follow-up action plans. All that is very objective and they don’t beat me over the head with it.”* [Distributor, 2012]

As experience was gained, the DBSC has become the foundation of the compensation system. It is not only used for contract renewal, but most distributors have cascaded the DBSC throughout their organizations. Distributor owners, managers and employees are compensated based on either DBSC measures or the leading indicators of a DBSC measure for lower-level employees.

- *“To get off the bottom on the DBSC...we did two things. We aligned everybody’s compensation schemes to reflect those metrics that they could influence. The other thing we did was we developed branch scorecards; those were not in existence twelve years ago.”* [Distributor, 2012]

Interestingly, the company did not dictate this diffusion of the DBSC.

3. RQ2: WHAT HAS NOT CHANGED OVER TIME?

3.1. Purpose of the DBSC

The overarching purpose of the DBSC as a comprehensive performance measurement tool has remained unchanged over its 15-year life.

- *“The DBSC is trying to give us a broader business set of measures of success than the more traditional financial and market share. It wraps a set of things together that make sense for managing the business.”* [Distributor, 1999]
- *“If we were running our business right, doing the right things, we would score well on the DBSC.”* [Distributor, 2012]

Unwavering over the years is the DBSC as a method of aligning distributor goals with company goals, which is the central purpose of a management control system (Anthony and Govindarajan, 2003).

- *“The company designed it to put priority on things they think are important.”* [Distributor, 1999]
- *“The distributors all accept it that they have to be graded. I report to somebody. I’m supposed to be looking after the company’s best interests. The scorecard was a way of creating that.”* [Distributor, 2012]

3.2. Relative Performance Evaluation and Best Practice Sharing

A consistent use of the DBSC has been for relative performance evaluation. According to company documents from 1999, the DBSC was used for comparing and ranking distributorships. In the early years of the DBSC, each distributor received its own report and its overall relative numeral ranking (e.g., 7th out of 31). Additionally to promote competition, the names of distributors achieving top ratings on individual measures were posted on the company’s intranet for all distributors to see.

- *“Anytime you publish a report and there are X 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.”* [Distributor, 1999]
- *“[Gathering] the information and sharing it back to us, saying other distributors are X. I can look at it and see how I am doing. Why am I different? I can use it as a lever to try to improve.”* [Distributor, 1999]

According to the company, posting the relative performance information also creates an avenue for best practice sharing. A natural tension exists between relative performance evaluation and sharing of best practices. For a distributor, there is minimal, if any, advantage to sharing information with another distributor that may, as a result, outperform them. However, most distributors do in fact help others to improve:

- *“I’ve got some good relationships in my peer group and other distributors and I’ll call up and say, ‘well, how did you do this?’ and ‘how did you do that?’ and they’ll help me out.”* [Distributor, 2012]
- *“You would think that since they rank us that we would want each other to do poorly so we would rank higher. That’s not really the case. We’re all good friends.”* [Distributor, 2012]

As DBSC use evolved over time, trust between distributors and company management has grown, and so has the amount of information shared across distributors. Today, the scores achieved on each measure for every distributor are published for all distributors to see.

4. RQ3: WHY HAS THE DBSC ENDURED?

4.1. Learning and Communication Tool

Over the course of this longitudinal study, it became clear that the company and distributors are engaged in a continual process of feedback, analysis, and adapting. Even a quick glance at RQ1 and RQ2 shows there were far more features of the DBSC that changed than remained the same.

- *“Over time, it’s been all over the map, or it evolved. A lot of things have changed over the period of time...the number of distributors, the ownership philosophy, the measures to focus on. I think the fact that it seems to be more collaborative in terms of what’s on there. And, hopefully, rather than a punitive tool, it’s more of how do we improve the business.”* [Distributor, 2012]

Initially, changes tended to be viewed negatively. However, the formalization of the process for making changes mitigated those reactions. Distributors have gained confidence in the DBSC over the years.

- *“That continuity over many, many years gives it credibility and longevity. It’s not a here-today-gone-tomorrow type of measure and the measures don’t change a lot year to year. That makes it a lot more effective.” [Distributor, 2012]*

Distributor confidence in the DBSC is reflected in its integration into their employees’ variable compensation.

A vital piece in the learning process is communication. As mentioned previously, in the early years of the DBSC, the company imposed DBSC measures and benchmarks without seeking input from the distributors. Although communication channels were in place, it was an imposed and very top-down approach to performance measurement.

- *“I’ve gotten no response to my concerns, and I’m P.O. ’d at them on this subject.” [Distributor, 1999]*

The result was distributors feeling ignored and trivialized because of their non-involvement. Today, information flows both ways but with a clear understanding that the DBSC is a company tool, not a distributor tool.

- *“They’re introducing a new thing on the scorecard around dealer support, and I guess some team created it, but many distributors didn’t participate in it. So as you can imagine, there was a bit of vibe or lack of buy-in. I have a meeting tomorrow where we’re going to talk through some of them. They’re giving us some opportunity to weigh in on it.” [Distributor, 2012]*
- *“They listen to us and we listen to them and we found each kind of modification drives behavior.” [Distributor, 2012]*

The company seems to have struck a balance between a top-down and bottom-up approach to performance measurement.

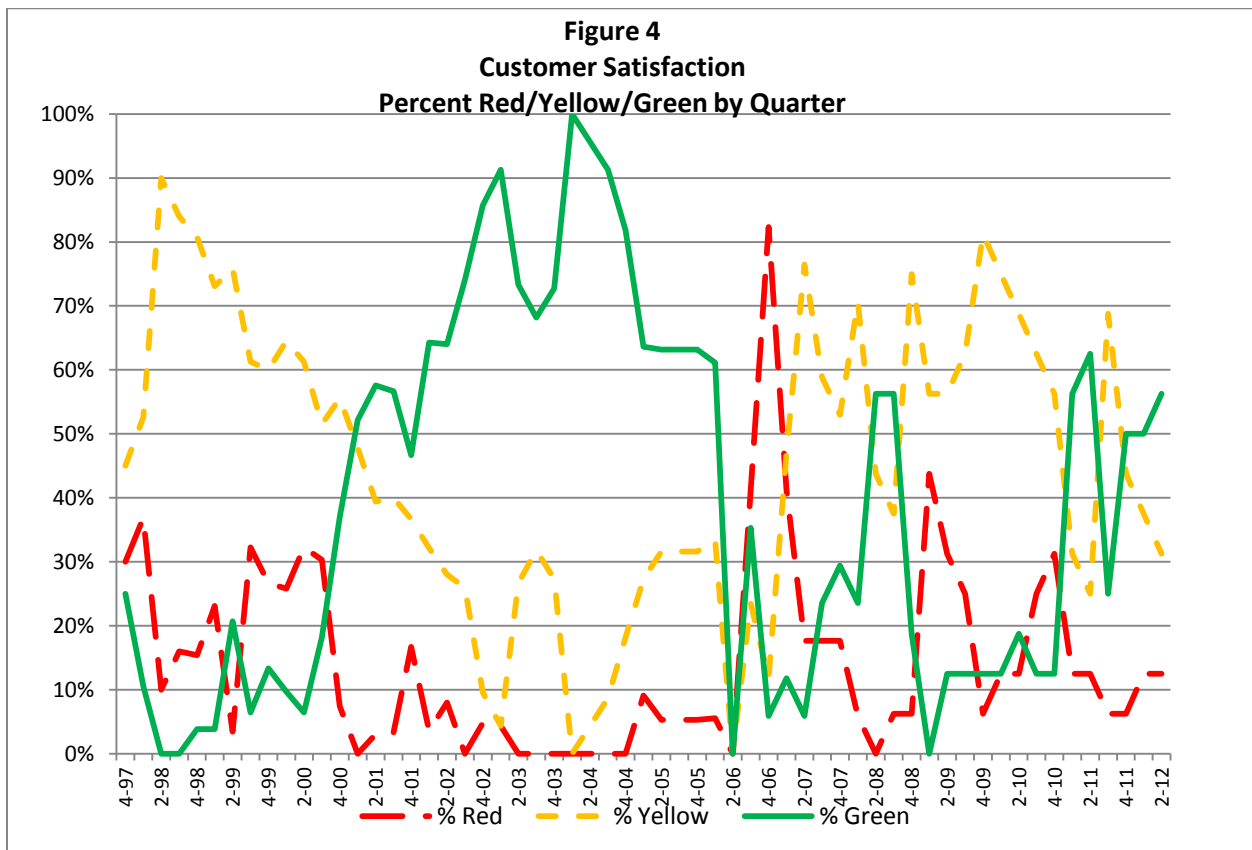
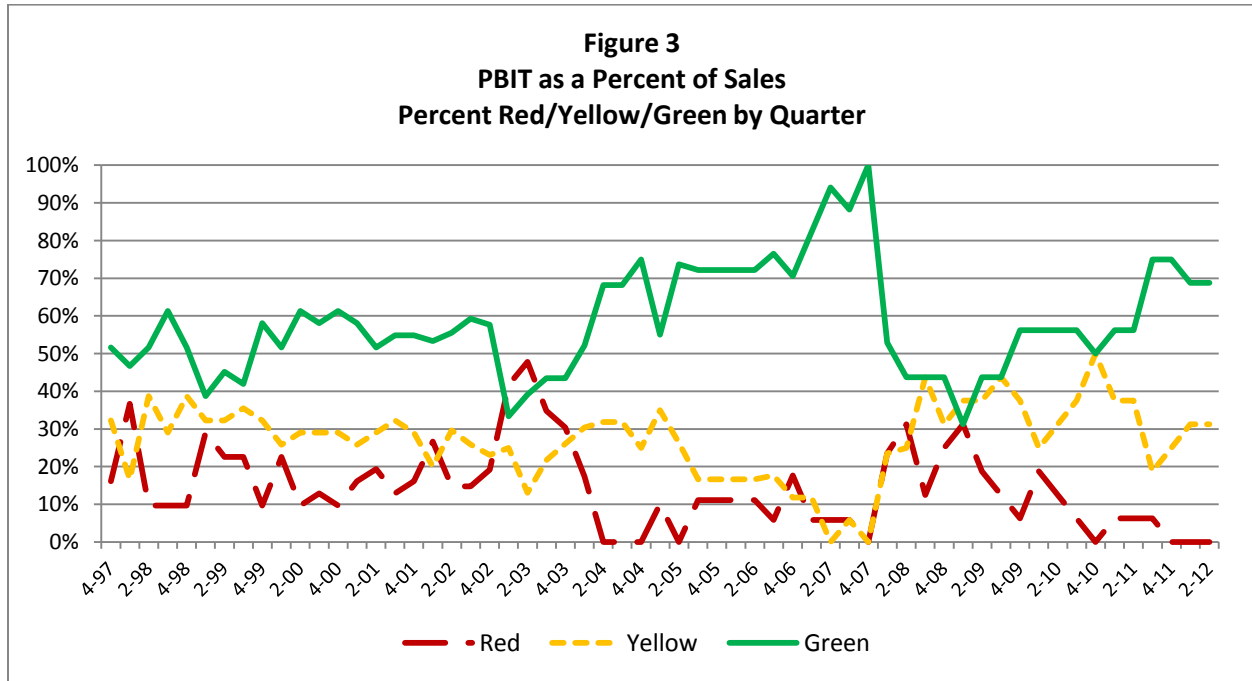
4.2. Influencing Behavior

Management control is a process that managers use in order to influence the performance and behavior of the people forming an organization (e.g., Tannebaum, 1967; Collins, 1982; Fisher, 1995). Both qualitative and quantitative data collected during the study illustrate the DBSC’s impact on distributor behavior. Distributors discussed the impact of the DBSC on their behavior:

- *“I wasn’t an advocate [of a service time measure] at the start, but now I am. It tells us how quickly we figure out what’s wrong so we can make an intelligent statement to the customer, and so they can say ‘go ahead’ or not...it’s helping us, though it’s not easy to change the mentality, but it’s good.” [Distributor, 1999]*
- *“[The DBSC] does influence the priority of initiatives, the resources that go to initiatives and really highlights the areas we need to focus on improving.” [Distributor, 2012]*

Figures 3 and 4 present quantitative evidence of changes in distributor behavior. Figure 3 shows the improvement in the PBIT to sales ratio. The first ten years of the DBSC shows a clear trend that more distributors are green and fewer are red. By the end of 2007, all distributors achieved the green level for this measure. In 2008, the company raised the cutoff to achieve green on this measure by three percentage points. This change is reflected in the graph by a drop in the percentage of green distributors and an increase in red and yellow. Within a year or so, there is a gradual increase in green and decrease in red. Figure 4 displays the same information for the customer satisfaction measure. Clearly, customer satisfaction is a more volatile measure. Over the first nine years of the DBSC, there is a clear improvement trend, with a lessening percent of yellow and red distributors and a growing percent of green distributors. When almost all or all distributors reached the green threshold, the company raised the bar. For the first DBSC of 2006, the company changed the cutoff, making it more difficult to achieve green status. Although the trend is not as clear as prior to 2006, after 2006 the percentage of red distributors is falling and, in general, the

percentage of green distributors is growing. As the company makes it harder to achieve a green status, the competitive distributors are motivated to achieve the green level, resulting in improved performance.



5. DISCUSSION

Although it is tenuous to make generalizations based on one field study, major themes that come to the forefront over the years are; 1) the broad definition of balance, 2) the role of subjectivity, 3) the importance of cause and effect among measures, and 4) the determination of success. Over the course of this study, it became clear that there are numerous definitions or uses of the term ‘balance’. From an academic point of view, the main stream performance measurement literature views balance as the combination of financial and non-financial measures (e.g. Kaplan and Norton, 1992). This dichotomy encompasses the concepts of leading/lagging and short-term/long-term, with financial measures viewed as lagging and short-term and nonfinancial measures viewed as leading and more long-term. The original company designers of the initial DBSC were Harvard Business School graduates and embraced this view of balance. Distributors referenced this view of balance:

- *“[The DBSC is] trying to give us a broader business set of measures of success than the more traditional financial and market share.”* [Distributor, 1999]
- *“In many cases we tend to be more precise on the financial side than other sides...whether it’s people or environmental concerns or facility concerns as well as the financial side of the business.”* [Distributor, 1999]

Over the years, the definition of DBSC balance expanded beyond that initial dichotomy to include balance in terms of an even playing field:

- *“[The DBSC is] intended to be a way that the company can measure the performance of distributor network in a such a manner that it puts everyone on a level playing fields as far as measures.* [Company Employee, 2012]
- *“As I say, the DBSC has endured because the company feels that it is balanced measurement in trying to take out some of that territorial pluses and minuses that come with each distributor’s territory.”* [Distributor, 2012]

In addition, the balance between objectivity and subjectivity is a recurring theme throughout this longitudinal study. Subjectivity in performance measurement is a complex concept since there is a general preference for objectivity (Moers, 2005; Merchant, 2006) but high quality objective performance measurement may not be attainable. Managers may choose to take on the obstacles of objective performance measurement or they may turn to subjectivity as an alternative.

As mentioned in the discussion of RQ1, the company appears to have little tolerance for subjectivity in performance measures. Numerous subjective measures (e.g. people measures) were included in early DBSCs but as experience was gained, the subjective measures were dropped. Apparently this company, with its long history of conservative, bottom-line financial management, could not tolerate subjectivity in measurement (Malina and Selto, 2004). Subjectivity was gradually removed from the DBSC measures. Another area where there appears to be no role for subjectivity is in the ranking of distributors. The total score for each distributor is a mathematical calculation and that score is published for all distributors to see.

Conversely, subjectivity has been gradually introduced into other areas of the DBSC. Initially, the cutoffs for red/yellow/green as well as overall assessment of DBSC performance were determined objectively. As mentioned in RQ1, the cutoffs are now modified if they are deemed ‘unreasonable’. Another area where the company appears to tolerate subjectivity is in the overall evaluation of the distributor. In the early years of DBSC use, a company executive stated:

- *“If a distributor is in the bottom quartile for 2-3 quarters in a row, then they are on probation.”* [Company Employee, 1999]

Distributors are no longer on probation for repeatedly underperforming on individual measures or overall DBSC scores.

- *“I think [the DBSC is] objective first, and then subjective second. In other words, if a distributor is in the bottom quartile for three or four quarters, there’s a discussion that goes on at some point, to the extent there’s some logic to the position and why the distributor’s there, then that’s the subjective piece.”* [Distributor, 2012]

Subjectivity can be used to recalibrate rewards to effort regardless of failure to achieve targets (Gibbs et al., 2004) As the company recognizes, factors specific to the distributor are taken under consideration in determining contract renewal.

- *“The DBSC is not an absolute gauge. If the distributor is red for 12 quarters it doesn’t mean they’re in jeopardy. It could be factors specific to the distributor. It’s not just based on color.”* [Company Employee, 2012]

It appears the company uses subjectivity in assessing performance to neutralize the impact of uncontrollable events, while still objectively measuring distributor response to those events. It also appears that the company has, over time, reached a balance between objectivity and subjectivity.

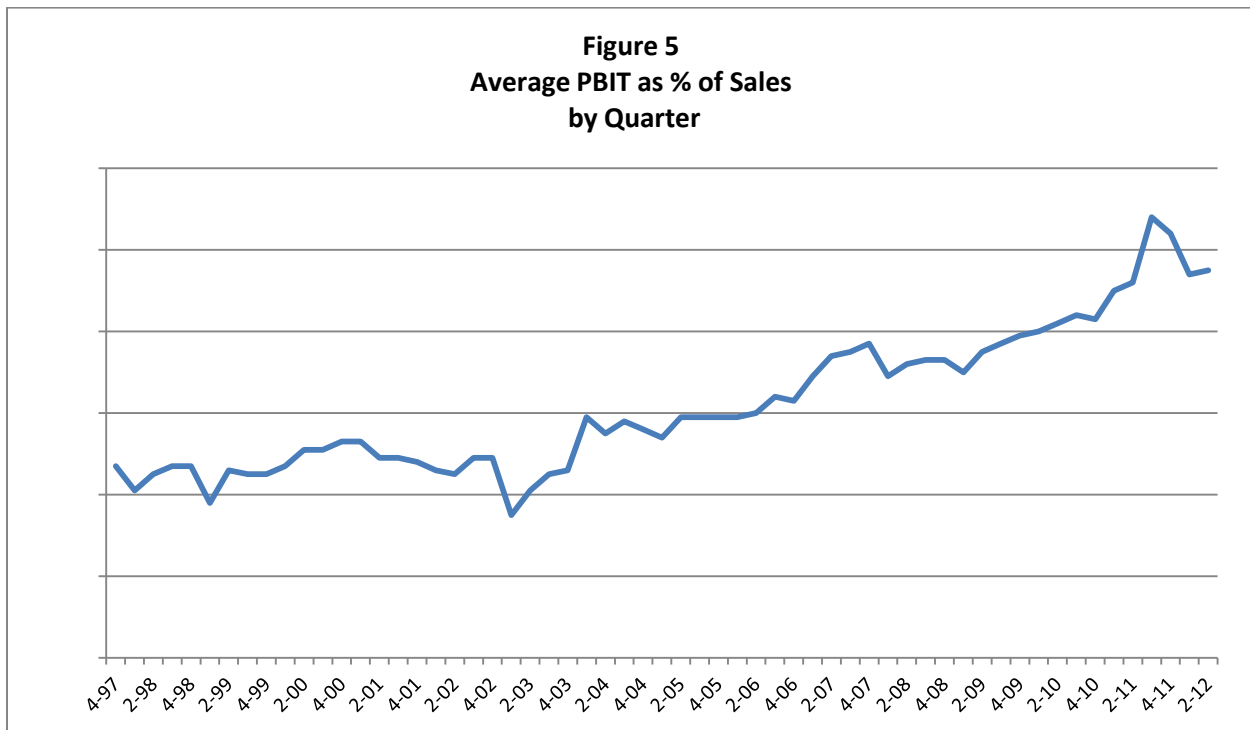
When the present study began in 1999, academics were stressing the importance of causality in BSC effectiveness (e.g. Kaplan and Norton, 1996; Epstein et al., 2000). In fact, causality was a dominant theme in early investigations of the DBSC. Distributors also recognized the theoretical connection among the DBSC measures:

- *“The theory of how the measures tie together is good. There are logical correlations.”* [Distributor, 2012]

Surprisingly, the company and distributors had never conducted statistical analysis to test the relationships among financial and non-financial DBSC measures. Recently, the company tested for correlation among measures and determined that too much noise exists in the data to find even significant correlations.⁵ For example, customer satisfaction complaints tend to be about one-third based on product quality or warranty issues that have nothing to do with the distributor. The company believes that it ‘makes sense’ that the measures are correlated, but other things get in the way of finding it, including the instability of the DBSC makeup.

What makes the DBSC, or any other BSC for that matter, a success? Distributor perceptions of fairness and legitimacy appear to create motivation and support conformity. Figure 5 highlights the steady increase in average PBIT as a percent of sales that may be providing positive financial feedback. Changes in distributor behavior coupled with positive financial feedback may provide sufficient proof of effectiveness for continued use of the DBSC.

⁵ This is consistent with Malina, Norreklit and Selto (2007) who rigorously refuted causality in the DBSC.



*Y-axis labels were suppressed at the company’s request.

This longitudinal study provided the opportunity to track the evolution of the DBSC over its 15-year life. The focus of this DBSC study was threefold; what has changed over time, what has remained the same, and why has it endured. Changes highlighted are (1) the DBSC was unaffected by a major change in organizational structure, (2) a narrowing of focus and reduction in the scope of the DBSC over time, (3) processes for changing the design of the DBSC were formalized, and (4) that the DBSC has become engrained in the compensation system. Factors that have remained constant over time are (1) the purpose of the DBSC, (2) its use for relative performance evaluation, and (3) its use as a tool for best practice sharing. Two factors that appear to explain why the DBSC has endured are (1) its use as a learning and communication tool and (2) its ability to influence behavior.

The evolution of the DBSC suggests a tentative list of key success factors for building and sustaining a successful scorecard:

1. Make the development process collaborative among the key stakeholders and establish clear processes for making changes to the scorecard. This creates buy-in, makes expectations explicit, and eliminates surprises.
2. Communication is critical. Whether top-down, bottom-up, or among evaluatees, communication builds trust in the scorecard process and fosters legitimacy and information sharing.
3. Create a balance of objectivity and subjectivity in scorecard design and use. Objective performance measurement is the unachievable “holy grail” and subjectivity in measurement and evaluation of performance helps overcome its inherent deficiencies.
4. Embrace the evergreen nature of the scorecard. Stability in its purpose and use, coupled with flexibility in its design may be critical to success.
5. Delay using the scorecard extensively for compensation. As experience with the scorecard is gained, trust is established and scorecard integration into the compensation system may be seamless.
6. Periodically review the scorecard for its effectiveness in promoting desired behaviors and financial results. When necessary, make changes to its design or use.

These key success factors might also be helpful to researchers seeking to investigate the design, use or impact of a scorecard in one or many organizations. For example, a survey with the aim of correlating scorecard variables with performance might include questions about each of the six items listed. Otherwise one would ignore what seem to be important stage-of-development factors that might impact performance.

AUTHOR INFORMATION

Mary A. Malina is an Assistant Professor of Accounting at the University of Colorado Denver. Her research interests include management control systems, performance measurement, and subjectivity in performance evaluation. Her work has been published in several journals including *Contemporary Accounting Research*, *Journal of Management Accounting Research*, and *Group & Organization Management*. Mary A. Malina, Ph.D., University of Colorado Denver, P.O. Box 173364, Campus Box 165, Denver, CO 80217-3364. E-mail: mary.malina@ucdenver.edu

REFERENCES

1. Anthony, R., and V. Govindarajan. (2003). *Management Control Systems*. Boston: McGraw-Hill/Irwin.
2. Cohen, D., A. Dey, and T. Lys. (2012). Corporate Governance Reform and Executive Incentives: Implications for Investments and Risk-taking. *Contemporary Accounting Research* (forthcoming).
3. Collins, F. (1982). Managerial Accounting Systems and Organizational Control: A Role Perspective. *Accounting, Organizations and Society*, 7(3) 107-121.
4. Dixon, J., A. Nanni, and T. Vollman. (1990). *The New Performance Challenge: Measuring Manufacturing for World Class Competition*. Homewood, IL: Dow Jones-Irwin.
5. Epstein, M., P. Kumar, and R. Westbrook. (2000). The Drivers of Customer and Corporate Profitability: Modeling, Measuring, and Managing the Causal Relationships. *Advances in Management Accounting*, 9(1) 43-72.
6. Fisher, J. (1995). Contingency-Based Research on Management Control Systems: Categorization by Level of Complexity. *Journal of Accounting Literature*, 14 24-53.
7. Gibbs, M., K. Merchant, W. Van der Stede and M. Vargus. (2004). Determinants and Effects of Subjectivity in Incentives. *The Accounting Review*, 79(2) 409-436.
8. Guilding, C. (1999). Competitor-Focused Accounting: An Exploratory Note. *Accounting, Organizations, and Society*, 24(7) 583-595.
9. Johnson, H. and R. Kaplan. (1987). *Relevance Lost: The Rise and Fall of Management Accounting*. Boston, MA: Harvard Business School Press.
10. Kaplan, R., and D. Norton. (1992). The Balanced Scorecard – Measures That Drive Performance. *Harvard Business Review*, 70(1) 71-79.
11. Kaplan, R., and D. Norton. (1996). *The Balanced Scorecard: Translating Strategy Into Action*. Boston, MA: Harvard Business School Press.
12. Malina, M. and F. Selto. (2001). Controlling and Communicating Strategy: An Empirical Test of the Effectiveness of the Balanced Scorecard. *Journal of Management Accounting Research*, 13 47-90.
13. Malina, M. and F. Selto. (2004). Choice and Change of Measures in Performance Measurement Models. *Management Accounting Research*, 15(4) 441-69.
14. Malina, M., H. Nørreklit, and F. Selto. (2007). Relations Among Measures, Result Control and Performance Measurement Models. *Contemporary Accounting Research*, 24(3) 935-982.
15. Merchant, K. (2006). Measuring General Managers' Performances. *Accounting, Auditing & Accountability Journal*, 19(6) 893-917.
16. Moers, F. (2005). Discretion and Bias in Performance Evaluation: The Impact of Diversity and Subjectivity. *Accounting, Organizations, and Society*, 30(1) 67-80.
17. Rappaport, A. (1999). New Thinking on How to Link Executive Pay to Performance. *Harvard Business Review*, 77(2) 91-101.
18. Tannebaum, A. (1967). *Control in Organizations*. New York: McGraw Hill, Inc.