

WHY BALANCED SCORECARDS FAIL

Arthur M. Schneiderman

The Balanced Scorecard concept has spread throughout the worldwide business and consulting communities at lightening speed, even by today's fast paced standards. Its approach has instant appeal to a CEO. On one sheet of paper he can not only capture the key financial goals of his organization, but for the first time the most important non-financial drivers for their achievement. No longer will the operational side of his business be disconnected from the financial measures that stockholders use to judge his performance. He can be confident that if the non-financial measures, the independent variables of his

ARTHUR M. SCHNEIDERMAN is an independent consultant on process management. From 1986 to 1993 he was vice president of quality and productivity improvement at Analog Devices, Inc., where he facilitated the Quality Steering Council and chaired the Total Quality Management Implementation Council. Schneiderman has served on the Malcolm Baldrige National Quality Award as senior examiner and the Conference Board's US Quality Council II.

business are met, eventually the dependant financial performance will follow. In fact, he can rightly argue to his boss, the Board of Directors, that by achieving the non-financial goals, he is doing all that is humanly possible to advance the owners' interests. Any deviation from planned financial performance, particularly down-

ward, can only be attributed to exogenous factors over which he has no control and therefore cannot be held accountable.

What CEO would fail to be committed to the creation and management of this sheet of paper, this balanced scorecard? And what line manager would not welcome the agreed upon set of tangible operational goals? Given that lack of top management commitment has repeatedly been identified as the single most important factor in explaining the failure of organizational change initiatives such as TQM, Re-engineering and Activity Based Costing, is not the battle over and success assured?

Yet strip away the declarations of victory by those who make their living from them and you will find that the vast majority of so-called Balanced Scorecards fail over time to meet the expectations of their creators. After a few short years of use, they will join the other fads in the corporate scrap heap. Why should a tool that shows so much

EXECUTIVE SUMMARY

- *The Balanced Scorecard concept has intrinsic executive appeal.*
- *To be successful the Balanced Score-card must be viewed as the tip of the improvement iceberg.*
- *Less visible, but equally essential, are processes to assure that the scorecard contains the right things and that support systems are in place to maximize the chances of them being done right.*
- *External factors or impatience may overpower the long-term positive financial consequences of significant non-financial improvements.*
- *Tenacity and faith may be the most important CEO attributes for successful Balanced Scorecard implementation.*

promise have such an ignoble end?

I developed the first balance scorecard in 1987 while Vice President of Quality and Productivity at Analog Devices, Inc.^{1,2,3}. Although others have had involvement with more scorecard implementations, I base my views on many years of continuous experience in a single organization as the balanced scorecard “process owner.” In fact, the balanced scorecard is in its twelfth year as a valued part of Analog’s planning and review processes⁴. I firmly believe that a good scorecard can be the single most important management tool in Western organizations. To quote Tom Malone, President of Milliken and Company: “If you’re not keeping score, you’re only practicing.”

I offer the following view as to why most balanced scorecards fail:

1. The independent (i.e. non-financial) variables on the scorecard are incorrectly identified as the primary drivers of future stakeholder satisfaction.
2. The metrics are poorly defined.
3. Improvement goals are negotiated rather than based on stakeholder requirements, fundamental process limits, and improvement process capabilities.
4. There is no deployment system that breaks high level goals down to the sub-process level where actual improvement activities reside.
5. A state of the art improvement system is not used.
6. There is not and can not be a quantitative linkage between non-financial and expected financial results.

Let’s look at each of these more closely.

*A financial measure
has much greater
organizational weight
than its new
non-financial
sibling*

DETERMINING THE INDEPENDENT VARIABLES

Determining what went on Analog’s first scorecards was easy. Everyone could hear the voice (or was it the shout) of the customer loud and clear: “Where’s my order”? With a 20% yield in manufacturing, the cost driver was obvious. A long manufacturing cycle time (4 times what it could be) compounded the problem of recovering from a “yield bust.” Chaos on the manufacturing floor meant that non-revenue generating engineering lots, critical to the new product development process, were repeatedly bumped to the back of the queue. This significantly lengthened time-to-market. Show the resulting scorecard to any employee at any level and they would say “yup, those are the right things for us to be working on.” Show it to a customer and they’d say the same.

But, that was more than a decade ago. Today, nearly every surviving organization has made dramatic improvements in those then obvious areas. Now, the vital few are much less visible. One suggestion is to simply add more non-financial measures, but that will only result in a loss of organizational focus and a dilution of effort. A practical rule of thumb is the juggler’s limit of 7 to 10. In fact, my guidelines restrict the scorecard to a single 8½ x 11 sheet of paper, 18 pica or larger font size and a ratio of non-financial to

financial metrics of 6:1. This numerical imbalance is based on the fact that initially, a financial measure has much greater organizational weight than it’s new non-financial sibling.

The difficulty in identifying scorecard metrics is compounded by the emerging requirements of non-owner stakeholders: employees, customers, suppliers, communities and even future generations. More and more organizations are adding social responsibility as a stakeholder requirement by including discretionary environmental initiatives, diversity and employee wellbeing in their list of strategic objectives. Unless these requirements are explicitly considered, a balanced scorecard can be at their expense.

The most important implementation imperative for a successful scorecard is the enrollment of the entire organization in its achievement. Duncan MacDougal, a former Boston University professor, observed that all processes in an organization can be thought of as being connected by virtual “slack ropes.” Although any given process can initially be improved in isolation, eventually the slack comes out of the rope connecting it to some other process, requiring that process’s concurrent improvement. High performance organizations have no slack ropes, creating the need for total participation in achievement of significant goals. To paraphrase that old saying, an organization is no stronger than its weakest process.

The stretch objectives that are inherent in a good scorecard can not be achieved by doing things in the usual way. As Rita Mae Brown said, “the definition of insanity is doing the same thing over and over again expecting different results” or as Jim Bakken, former VP of

Quality at Ford Motor Company paraphrased it, "doing what you did will get you what you got." Organizational change is subsumed in the Balanced Scorecard, and organizations only change when employees share ownership for both the goals and means.

Given this complexity, how can an organization construct a scorecard that truly balances all of the stakeholders' sometimes conflicting desires? The only approach that I have found successful is to adapt the methodology and tools used in Quality Function Deployment⁵ (QFD). This involves three phases:

1. Establish prioritized (numerically weighted) stakeholder requirements based on strategy adjusted need for improvement.
2. Quantitatively rank the processes in terms of their aggregate impact on these requirements.
3. Create appropriate metrics for the processes at the top of the list.

I have found that the group activity associated with this approach not only leads to team consensus, but also produces a compelling and logical "story" which is invaluable in communicating the scorecard's rationale to the rest of the organization. Management consensus and a good story are often keys to getting buy-in from the remainder of the organization.

DEFINING GOOD METRICS

While financial metrics have undergone more than a century of development and refinement, non-financial metrics are relatively new to the scene. Little wonder that there are no standards and that current practice yields definitions that often have serious, even fatal

*The Metrics
may improve, but
all too often,
the underlying
processes don't.*

flaws. Yes, the metrics may improve, but all too often, the underlying processes don't. I've written in the past on the requirements for good metrics⁶ and applied the theory to the order fulfillment process. In summary, metrics can be classified as results (measures seen by the process customer) or process (internal measures that cause the results) metrics. Results metrics are most useful as a management tool and are usually what appear on a scorecard. Process metrics are most useful to improvement teams since they focus attention on the places within the process where improvements will have the greatest impact.

Good metrics are the following:

1. A reliable proxy for stakeholder satisfaction.
2. Weakness or defect oriented (have an ideal value of zero) and continuous valued.
3. Simple and easy to understand.
4. Have well documented, unambiguous, consistent, appropriately smoothed, and metrologically sound operational definitions,
5. Timely and accessible to those who can best use them,
6. Linked to an underlying data system that facilitates the identification of root causes of gaps in scorecard results, and
7. Have a formal process for their continuous review and refinement.

Metrics need to be defined and maintained in a tops-down and bottoms-up process that combines the detailed knowledge of the process executors with the big picture perspective of the executive. This need for joint ownership of metrics definition is often overlooked with the result that the metrics are either unactionable or disconnected from business objectives.

SETTING SCORECARD GOALS

Unlike its sports counterpart, a balanced scorecard needs to have specific goals and timeframes. Unfortunately, most scorecard goals are negotiated, but, as I have previously observed:

"Therein lies the basic flaw in current goal setting: specific goals should be set based on knowledge of the means that will be used to achieve them. Yet the means are rarely known at the time goals are set. The usual result is that if the goal is too low, we will underachieve relative to our potential. If the goal is too high, we will underperform relative to others' expectations. What's really needed to set rational goals is a means of predicting what is achievable if some sort of standard means for improvement were used⁷."

In that 1988 article, I went on to propose the half-life method. Based on an analysis of nearly 100 improvement efforts, I observed that the resulting metrics improved at a constant rate, expressed in months to achieve a 50% defect reduction. The observed half-life depended on the organizational and technical complexity of the process and ranged from 1 to 22 months. Rather than negotiating scorecard goals, they should be based on knowledge of the required corrective actions, or absent that knowledge the capabilities of the improvement

process as captured in an empirical model such as the half-life method.

SCORECARD DEPLOYMENT

We all know that our financial systems consolidate data generated at the transaction level. For example, individual sales are aggregated to the product level, then to the product line level, and on until a total corporate sales number is calculated. This process can be reversed providing the means to explain changes in total sales. Non-financial measures should in principle follow the same model.

Unfortunately, while sales are denominated in consistent units of currency, most non-financial measures have incomparable units. Combining often involves mixing apples and oranges. However, the value of deploying scorecards from the top to the bottom of the organization is particularly beneficial in providing alignment of improvement activities. Without this alignment, significant process improvements throughout the organization fail to generate bottom line results.

My view is that scorecard deployment needs to be a major activity in the management of balanced scorecards. Wherever possible and sensible, scorecard goals should be disaggregated and deployed downward in the organization so that each employee understands their piece of the big picture and can share in the knowledge of their contribution to the organization's overall success. Where this is not possible, fuzzy linkages between scorecards can be made. There is great value in even subjective agreement that if all of the goals of subordinate scorecards are achieved, than a higher level goal will also be achieved, almost

EXHIBIT 1

Non-Financial Scorecard as Published in Sloan Management Review in 1988

Analog Devices: Quality Improvement Program Goals			
Measurement	1987	Half-Life (in months)	1992
External			
On-Time Delivery	85%	9	>99.8%
Outgoing Defect Level	500 ppm	9	<10 ppm
Lead Time	10 weeks	9	<3 weeks
Internal			
Manufacturing Cycle Time	15 weeks	9	4-5 weeks
Process Defect Level	5000 ppm	6	<10 ppm
Yield	20%	9	>50%
Time to Market	36 months	24	6 months

with certainty. This approach is a centerpiece of Hoshin Kanri⁸, a planning and management system widely used in Japan.

One metric that does transcend processes is the ratio of improvement half-life to normative target half-life. I consider this to be the prime measure of organizational learning⁹.

STATE-OF-THE-ART IMPROVEMENT PROCESS

Nearly half a century ago, the Japanese codified a superior process improvement methodology called the "7-Step Method"¹⁰. This approach embodies the scientific methodology at a level that can be employed anywhere and by anyone. Even before that, similar methods such as Kepner-Tregoe were in wide use in the West. Yet, I am amazed by the number of well known organizations that I've visited that still rely on trial and error as their official improvement methodology. They do not call it that, but diagnosis reveals the lack of a scientific approach. Usually missing are essentials such as root cause analysis, verification of improvement, documentation of

changes, and reflection on the improvement process itself.

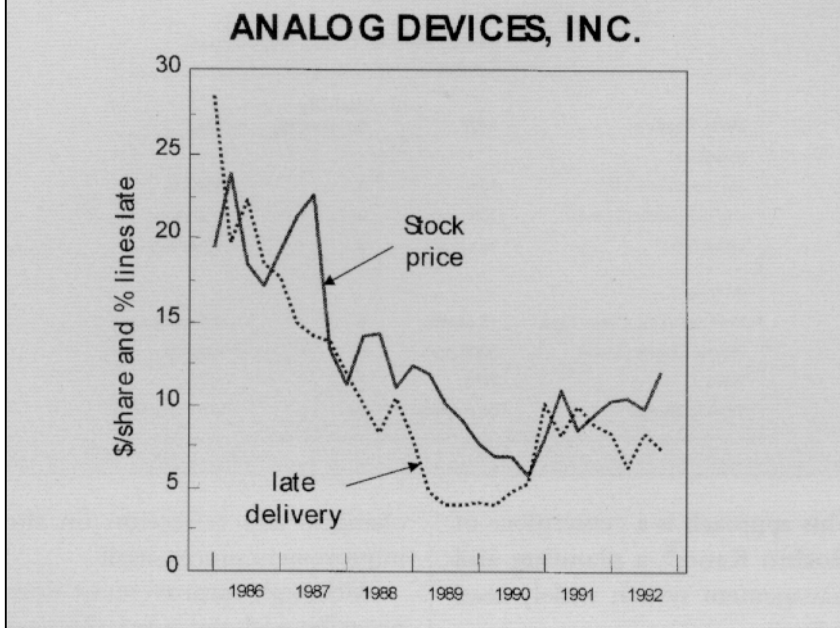
Although improvement does occur by trial-and-error, the rates of improvement are less than 10% of what they might be. This is compounded by executives' natural tendency to expect improvement at a rate 10 times what it rationally could be. This combination provides ample fuel for frustration.

LINKING BOTH SIDES OF THE SCORECARD

I started this article with the premise that both sides of the scorecard are linked by a metaphorical equation. The non-financial measures represent the independent variables, the prospective or leading indicators of change. The financial measures are the dependant variables and are the retrospective, lagging indicators. Some organizations are tempted to make this linkage quantifiable. They ask their improvement team leaders to "quantify the opportunity," that is to dollarize the likely bottom line impact of their proposed effort. Fortunately, these same organizations run this system open loop so there is no ex post facto accountability for the forecasted financial results.

EXHIBIT 2

Are the Financial and Non-Financial Sides of the Balanced Scorecard Linked?



I have learned the painful lesson that an organization is not just the sum of its parts. Being a complex and organic creature, much of its nature lies in the interaction of these parts with each other and with the external environment. But not only does organizational and technological complexity confound the equation. We are beginning to learn about the applicability of chaos theory to business systems. In chaos theory, very small, even minute decisions have an unexpected yet profound and lasting effect.

I believe that management needs to take on faith or fuzzy logic the linkage between the financial and non-financial sides of the scorecard. We do the non-financial things because it is the collective wisdom of the organization that they will improve our chances of success.

This "leap of faith" can be evidenced in a number of ways. For example, in 1988, Ray Stata, then CEO of Analog Devices, included its 5-year non-financial scorecard

(1987-1992) in his article in the Sloan Management Review (see Exhibit 1). In doing so, he not only published what was heretofore-proprietary information (delivery, yields, defect levels and cycle times), but also publicly committed to specific future improvements. In addition, Analog came within days of committing to the publication of the entire balanced scorecard as a regular part of its Annual Report. Concern that "Wall Street" was not ready for it led to the last minute cancellation of this potential innovation.

Nor does the passage of time necessarily justify the balanced scorecard. Achievement of the non-financial goals can not assure absolute business success. The external environment often dominates over internal improvements. Take for example Analog's case. Exhibit 2 shows its principal non-financial delivery performance metric and the concurrent stock price. There appears to be a good correlation

($R^2=0.64$) between percent shipments late and stock price over the seven year period. Unfortunately, the correlation with delivery performance is negative. In other words, as delivery performance gets better, stock price drops!

We all know that correlation does not mean causality. But try explaining these data to someone who has been only reluctantly convinced that the non-financial scorecard metrics are a leading indicator of future financial success. To make matters worse, Analog's delivery performance worsened after 1993 and the stock price continued to rise to a 1998 split-adjusted peak of around \$100. There are two possible explanations: the semiconductor business cycle dominated, or the lag time between delivery and stock price was more than five years. Determining which requires analysis and assumptions that remain unconvincing. With data like these, it is tempting to move to the "relative" business success argument: things would have been much worse had we not achieved the non-financial goals. Again, an argument that is hard to prove to the skeptic.

Yet the reality is that the lag time between non-financial and financial performance can be much longer than we would initially expect. First, there's the time it takes for the customer to perceive the change and become convinced that it is permanent. Then there's the time for them to change their purchasing patterns, often lengthened by existing multi-year purchase contracts. It would not be surprising for the aggregate lag time in many situations to be in the range of 5-10 years! This requires

<p>organizational leadership skilled in what Admiral Hyman Rickover, the “Father of the Nuclear Navy” called “courageous impatience.” Even in monolithic Japan, it took over 25 years after the 1950 introduction of TQM for the world’s perception of Japanese product quality and subsequent purchasing patterns to change.</p> <p>In conclusion, it is somewhat ironic that the first balanced scorecard did address each of the above challenges. It did not do so perfectly, but annually the issues were reexamined and refined. Analog’s scorecard was the frosting on a very substantial cake.</p>	<p>Unfortunately, many subsequent scorecard attempts have focused on the frosting, not the underlying substance. It should be of no surprise that the wished for silver bullet mysteriously melts away before reaching it’s distant target.</p> <hr/> <p>¹Ray Stata, “Organizational Learning – The Key to Management Innovation, Sloan Management Review, Spring 1989 63-74</p> <p>²Robert S. Kaplan, Analog Devices: The Half-Life System (Boston, MA: Harvard Business School, 1989) Case #9-190-061</p> <p>³Robert S. Kaplan, “Companies as Laboratories,” in <i>The Relevance of a Decade</i>, Paula Barker Duffy (ed.) (Boston: Harvard Business School Press, 1994): 179- 182.</p> <p>⁴Robert Stasey, “The Evolution of ADI’s Scorecard” in <i>New Management Accounting, How Leading-Edge Companies Use Management Accounting to Improve</i></p>	<p>Performance, William F. Christopher (ed.) (Crisp Publications, Inc., 1998): 85-101.</p> <p>⁵Yoji Akao (ed.) <i>Quality Function Deployment</i>, Yoji Akao (ed.) Productivity Press, 1990</p> <p>⁶Arthur M. Schneiderman, “Metrics for the Order Fulfillment Process,” <i>Journal of Cost Management</i> (Part 1: (Summer 1996): 30-42.,Part 2: (Fall 1996): 6-17)</p> <p>⁷Arthur M. Schneiderman, “Setting Quality Goals,” <i>Quality Progress</i> (April 1988): 51-57</p> <p>⁸Yoji Akao (ed.) Hoshin Kanri, <i>Policy Deployment for Successful TQM</i> (Cambridge, MA: Productivity Press, Inc., 1991).</p> <p>⁹Arthur M. Schneiderman, “Measurement, the Bridge Between the Hard and Soft Sides,” <i>Journal of Strategic Performance Measurement</i>, April/May 1998 14-21.</p> <p>¹⁰ Arthur M. Schneiderman, “Are There Limits to Total Quality Management?,” <i>Strategy & Business</i>, Second Quarter 1998, 35-45 also reprinted in abridged form in <i>Measuring Business Excellence, the Quarterly Journal of Business Performance Management</i>, Second Quarter 1998, 4-9.</p>
--	--	---