revealing exercise, when working with groups of supply professionals, is to ask for a show of hands from those who say their organizations measure supplier performance. Those with hands up (usually a high percentage) are then asked to keep them raised if they are satisfied with their supplier measurement systems. The resulting rush of falling hands is quite an indictment of the state of supplier measurement today.

An unmistakable conclusion after researching academic and trade publications for information on supplier performance measurement is that not much is written on this topic. What is written about it reveals overwhelming agreement about its importance. Yet the development of effective measurement systems is still on the “to do” list for many organizations, particularly smaller ones. Even organizations that boast mature systems of supplier metrics should recognize that continuous improvement is an ongoing challenge—and that many such metrics systems have shortcomings.

This article identifies these shortcomings, and provides guidance about how to create an ideal supplier performance measurement and scorecard system.

A Primer on Supplier Scorecards

Let’s make sure we are on the same page as it relates to supplier performance measurement. It can be defined as the business process that includes the methods and systems used to collect and provide information in order to measure, rate, or rank suppliers on a continuous basis. Many companies use the term “scorecard” to describe the report that conveys performance information to suppliers.

The types of scorecards in use typically fall into one of three categories—categorical, weighted point, or cost-based. Categorical measurement systems require simple check-offs to items that describe a supplier’s performance across different categories. For relatively unimportant items, this may be an effective way to evaluate supplier performance. As it relates to supplier scorecards, most supply chain organizations use a weighted point system that includes a variety of performance categories, provides weights for each category, and defines the scales used for scoring within each category. The third type—cost-based systems—is used least. It attempts to quantify the total cost of doing business with a supplier over time. Some companies use a hybrid system comprising several of these approaches. Exhibit 1 summarizes the advantages and disadvantages of each system.

No standard measurement approach exists across industries, although supply chain organizations should strive internally for some consistency, particularly with respect to the technical aspects of their systems. Some organizations have also joined consortiums that share best measurement practices or attempt to follow standards that appear in the Supply Chain Operations Reference (SCOR) model. It does not make sense for every business unit or internal location to re-invent how they measure performance. The challenge today is to develop a measurement process and scorecard system that offers some flexibility to a company’s internal operations while maintaining company-wide consistency.
Most companies are less than satisfied with their systems in place to measure supplier performance (if they even have a system at all). But creation of an effective supplier scorecard—one that aligns directly with the outcomes sought from doing business with that supplier—is an eminently achievable goal. The key is to focus on a set of core characteristics.

**Measurement Systems that Fail**

Sherry Gordon, author of the book *Supplier Performance Management*, has stated that few purchasing and quality professionals are likely to answer “yes” when asked whether they are satisfied with their supplier assessment capabilities and results. Where supplier scorecards do exist, some are so ill-conceived that at times it might be better if they were not used at all. Far too often, measurement is an activity that fails to lead to improved results. Consider the following examples.

Several years ago, a consumer products company with $100 million in annual sales developed a scorecard to evaluate its suppliers, most of which were substantially larger than the company. It was bad enough that this scorecard was not pilot-tested and was less than professional in appearance. But the system failed when many larger suppliers challenged the accuracy of the company’s scores, particularly when the scores were lower than those received from the suppliers’ more sophisticated customers. Suffice it to say that this experience deterred the company from moving forward with its measurement objectives.

Procurement teams must take a hard look at their measurement processes long before suppliers can challenge the legitimacy of the metrics. The processes must not turn into the kind of exercise that one supplier’s executive described as “they present and we rebut.”

A second example highlights a variety of shortfalls that confront too many supplier measurement systems. Almost every supply chain organization has at least thought about developing a supplier scorecard system.
Those that are serious about the process have most likely committed serious time, budget, and resources toward development and maintenance of systems of measurement. One such company is a major logistics player. On the surface, this company’s system appears ideal. Do senior managers need a ranking of supplier performance sorted by commodity group? Do they want a listing of the company’s best or worst performing suppliers? This, and much more, is available at the push of a button.

However, during a training session at this logistics company, an instructor asked a buyer to name one of his best performing suppliers—what the company called an elite supplier. The intent was to use examples of real suppliers to demonstrate the data features of the system. Without hesitation, the buyer provided a supplier’s name. But from across the room, another participant responded by saying that the supplier just named was one of the worst suppliers that his operations group worked with every day.

How can one person cite a supplier as being worthy of preferred status while another, in the same supply chain organization, indicates that he would rather discontinue the relationship with that supplier? And what are the dangers of a system that awards high scores to poorly performing suppliers?

These differences of opinion led to some conclusions that almost everyone in attendance could agree upon. The consensus was that although the scorecard system was supported by an extensive database that allowed all kinds of rigorous analyses, the data to support the system was still collected and keyed in manually. Furthermore, many scorecard items required subjective judgments. On top of this, most buyers had responsibility for inputting data quarterly for about 25 suppliers, a heavy burden on top of their “normal” workload. Many in attendance also agreed that the data for the scorecards was keyed in just before, and sometimes after, the quarterly cutoff, meaning that the emphasis was hardly on the quality of the data. Attendees also acknowledged that supplier scores were used as an indicator of a buyer’s job performance.

The group also agreed that their suppliers were held to the same criteria and weights, even though not all suppliers were equally important to the company’s success. Participants further agreed that internal customers or stakeholders had no way to be part of the measurement process. There was also some confusion about what kind of organization qualified as a supplier since some suppliers provided material from multiple locations. Finally, no clear agreement emerged that the measurement process was contributing to higher performance.

What are some lessons here? Clearly, an effective scorecard system requires much more than a sophisticated database that can present data in many ways. While that capability is important, technical capabilities do not guarantee system success. And scorecards should not ignore the voices of internal customers. Managers at manufacturing plants, warehouses, distribution centers, and logistics hubs are often perfectly positioned to evaluate suppliers’ day-to-day performance.

Another lesson is that scorecards often place a serious work burden on the individuals responsible for maintaining them, which often results in scorecards that are late or completed at the last minute—which raises concerns about data integrity. Is a reliance on subjective and last-minute evaluations affecting the integrity of the scores?

A further learning is that scorecard systems can result in too much averaging of data for suppliers that provide goods from more than one location. If a supplier provides goods from 15 locations around the world, does this call for one scorecard or 15? If this supplier pursued ISO 9000:2000 certification, the certification would apply to individual sites, not to the entire company. Furthermore, the number of suppliers and the number of shipping
points are often very different figures.

A final lesson is that scorecard systems can drive the wrong behavior. The results will be skewed—and not fit for their intended purpose—if a buyer’s annual performance evaluation is based partly on the performance of her suppliers. Worse: Her performance is often being determined by scorecards that she is responsible for completing. The conflict of interest is obvious. While most everyone at her company may agree that supplier measurement can be a good thing, it is also evident that the system in place is far from ideal.

**Characteristics of an Ideal System**
Research and work with hundreds of supply chain organizations has provided a unique opportunity to identify what comprises an ideal supplier measurement system. With that in mind, the following characteristics (summarized in Exhibit 2) will go a long way toward defining that system.

**The Measurement System Allows Scoring Flexibility**
Perhaps the most obvious shortcoming of most scorecards is they treat each supplier the same way. If segmentation occurs at all, it might simply be between suppliers of material and suppliers of services. Why apply equivalent scorecard measures when few would argue that all suppliers are created equal? It is alarming to see how prevalent the “one size fits all” approach is at even some of the most well-recognized supply chain organizations.

Better systems will allow adjustments to the performance categories and their weights to reflect the realities of different supply requirements. The best scorecards align directly with the outcomes sought from doing business with a particular supplier.

In one good example, an automotive original equipment manufacturer (OEM) has changed the way it evaluates suppliers by involving more employees in the process and giving them the power to adjust the weights used to evaluate suppliers.² The OEM now relies on 240 internal “boards,” one for each of its product segments, with at least four employees on each board to determine annually the weights of the various performance categories against which suppliers are evaluated. Each board consists of specialists in cost, technology, quality, and logistics who are responsible for posting supplier data monthly on a global supplier portal. Suppliers are even able to see the names and performance of their competitors, although the product boards have the authority to withhold names within their product groups if they so choose.

**Internal Customers Evaluate Supplier Performance**
In today’s information age, internal customers should be able to submit comments and ratings about a supplier’s performance directly into a scorecard system. These individuals are usually in the best position to evaluate a supplier’s operational performance.

A good example of involving internal participants occurs at ADT Security Services, a business unit of Tyco. Fully 30 percent of a supplier’s performance score relates to something called “account management.” This reflects how well a supplier works with ADT and responds to requests and concerns. Buyers actively solicit input from engineering, product management, marketing, sales, and product support before assigning a score, reflecting an extensive level of cross-functional input across the company.³

Procurement teams should consider allowing suppliers to enter a Web-based portal or extranet to view any free-form comments or scores submitted by internal customers. This supports the efficient and open exchange of information, something that is widely practiced with other supply chain applications (think about sharing demand forecasts, for example). Most supply chain experts would agree that information-sharing across the supply chain is a good thing. So why should sharing of supplier performance data be any different?

**Scorecards are Reviewed and Acknowledged by Suppliers’ Top Managers**
Key executives at each supplier should receive electronic copies of the scorecards. Perhaps most importantly, the party sending the scorecard should track acknowledgments that the scorecards were received and reviewed, along with any responses to specific queries.

Forwarding scorecards directly to executive managers supports at least two purposes. First, these executives will have access to information that their own personnel may not willingly share. More than one executive has been caught off guard because he or she was unaware of issues that affected

---

**EXHIBIT 2**

**Characteristics of the Ideal Supplier Scorecard System**

- The measurement system allows scoring flexibility.
- Internal customers evaluate supplier performance.
- Scorecards are reviewed and acknowledged by suppliers’ top managers.
- Suppliers with more than one location receive multiple scorecards.
- Scorecards include cost-based measures whenever possible.
- Scorecards are updated in real time.
- The measurement system separates the critical few from the marginal many.
- Measurement database allows user flexibility in retrieving and displaying data.
- The measurement system provides early-warning performance alerts.
- Suppliers can view and compare their performance online.
- The measurement system is benchmarked against best-practice companies.
customers. Second: Information will likely reach the individuals who can effect meaningful change when it is required.

The distribution list should include more than one executive. For smaller suppliers, it is worthwhile to place the chief executive, president or managing director on the list. It makes a lot of sense to provide vital “feedback” information to those who are ultimately accountable for performance results.

**Suppliers with More than One Location Receive Multiple Scorecards**

As noted earlier, there can be a tendency to count a supplier as a single entity, yet many suppliers provide material from multiple locations. To aggregate different locations into a single scorecard can be misleading. It also makes it harder to assign scores to specific locations.

A possible solution is to evaluate each supplier’s shipping locations across a basic set of operational metrics (such as cost, quality, and delivery) while the supplier as a corporate entity is evaluated by a set of higher-level metrics. Examples of such metrics include assessments of supplier innovation, responsiveness, and willingness to invest in the buyer-seller relationship.

One major OEM has addressed the issue of multiple supplier locations by assigning a five-digit code to each of its suppliers. Each location for that supplier receives a suffix to identify it as a unique location that will receive a scorecard. For example, a supplier with three shipping locations will have a corporate code (say, 45633). Its three shipping locations are then designated as 45633A, 45633B, and 45633C. This approach keeps the unique locations grouped within the supplier’s code, which helps when conducting any analyses.

**Scorecards include Cost-Based Measures Whenever Possible**

Most scorecards include price as a performance category simply because price is easy to measure. Unfortunately, price never reflects the total cost of doing business. To compensate for any disconnect between price and total cost, progressive supply chain organizations calculate metrics that reflect more than unit price.

An example of a total cost metric is the supplier performance index (SPI). The SPI assumes that any quality or performance infraction committed by a supplier increases the total cost of doing business with that supplier. If a supply chain manager can track these infractions and assign a cost to them, the calculated index can then be used in a scorecard to supplement a price metric. The SPI or even the adjusted price can be included in the scorecard rather than simply the price paid (although price can still be included as a scorecard metric).

**Scorecards are Updated in Real Time**

Too many scorecards still resemble a batch updating system that features periodic input of data submitted manually each month or each quarter. In a perfect world, anyone who is granted access to a scorecard system should be able to view supplier performance levels in real time. Whenever a transaction occurs, whether it involves the results of a quality audit at a receiving dock or an accounts payable transaction, data records should flow seamlessly into the scorecard database with real-time updating of supplier performance. Of all the attributes of an ideal measured system described in this article, this is the one that is rarely implemented.

For real-time updating to work, the scorecard system must be linked to other supply chain constituencies, including accounts payable, quality control, and transportation. Theoretically, any system that stresses objective rather than subjective assessment, particularly in a real-time environment, should receive serious consideration. It’s safe to conclude that most supply chain systems are moving toward real-time data visibility. Some purchasing organizations are beginning to rely on suppliers to self-report and submit their performance to the scorecard system on a frequent basis. A few leading companies are even beginning to solicit performance data from or about second-tier suppliers.

**The Measurement System Separates the Critical Few from the Marginal Many**

Several leading consulting firms have recently criticized the fact that not all suppliers are being measured using scorecard systems. But should this really be a cause for concern? In an era when fewer suppliers are providing a greater share of total purchases, there has never been more need to separate the critical few from the marginal many. At Procter & Gamble, for example, 400 suppliers out of 90,000 worldwide receive 25 percent of the company’s $50 billion in annual purchases.

If a supply chain organization is adamant about measuring most of its suppliers, then the less critical suppliers should receive a basic scorecard—perhaps even one that is categorical. (See Exhibit 1). At some point, depending on the level of effort required to obtain scorecard data, the cost to measure a supplier could outweigh the value of measuring that supplier. When this is the case, a logical response is to not measure, measure less frequently, or simplify the type of scorecard used.
The Metrics Database Allows User Flexibility in Retrieving and Displaying Data
An effective system will not only generate the scorecard itself; it will enable data to be presented in a variety of reporting formats, along with easy generation of useful reports. Various on-demand reports can show side-by-side supplier rankings, demonstrate performance changes by category, and highlight the suppliers that improved or deteriorated in performance over a certain period. A database that allows the slicing and dicing of raw data is an essential element of an ideal scorecard system.

The Measurement System Provides Early-warning Performance Alerts
Most measurement systems are reactive in that they report what has happened, not what is likely to happen. As with a statistical process control system, an ideal measurement system would be able to “look ahead” to spot troublesome trends and non-random changes in a supplier’s performance before it becomes out of control. An ideal system would notify supply chain managers of potential problems before the impact of those problems is even realized. The system would have predictive capabilities.

Consider the possibility of generating early warnings when using advance shipping notices (ASNs). Any time an ASN reveals a possible late delivery after comparing expected transit times against a due date, a material planner would receive a warning of the potential delay. Or consider real-time GPS tracking systems that could reveal that supply chain delays are occurring, with a notification sent to the appropriate personnel. It is almost always better to be proactive.

Suppliers Can View and Compare their Performance Online
For many years, almost every supply chain organization refused to identify the scores and names of competing suppliers within a category or commodity group. Later, most organizations became more willing to show relative comparisons against competing suppliers identified by letters (but not names). The time has come to accept that scorecards present a good way to create healthy competition among suppliers. That means permitting and enabling them to access their scores online, complete with comparisons to other suppliers in the same or similar commodity groups.

Scorecard transparency is an idea whose time has come. Note that transparency does not violate any buyer-seller ethics, laws, or standards of confidentiality. It is analogous to looking at the standings of any sports league. Doesn’t every team know precisely where it stands in relation to competing teams? At the academic level, colleges and universities are routinely rated and ranked against one another. Somehow these institutions survive the ordeal. Suppliers will be no different.

The Measurement System is Benchmarked against Best-Practice Companies
Performance benchmarking involves comparing products, practices, processes, or strategies against key competitors or companies that are considered best-in-class. Benchmarking methodologies can involve working directly with other companies to compare scorecard practices, searching databases and the Internet to find information on performance measurement and working with professional contacts to obtain scorecard information. Some supply chain organizations belong to research consortiums that feature the sharing of best-practice information. While informal benchmarking can occur at any time, formal reviews of the scorecard system should occur at least annually. In an era when almost too much information is available, there is no excuse for not remaining current regarding the trends and technologies that relate to supplier performance measurement.

It’s a challenge for supply chain organizations today to step back and take an unbiased view of their supplier performance measurement systems. The objective should be to take a poor measurement system and make it better—or to transform a good system into an excellent one. A worthwhile exercise is to assemble an internal team to compare the current state of supplier measurement against an ideal future state. Any gaps that exist between the current and future states—and there could be many—will require a clear plan to bring an existing system closer to a preferred system.

Footnotes:
1 For a detailed PowerPoint presentation of Total Cost of Ownership systems, please contact: rjt2@lehigh.edu.
2 This example is adapted from Armstrong, J., “Chrysler Changes Scorecard,” Automotive News, Vol. 79, No. 1610 (November 8, 2004), 16.
3 Teague, Paul, “A Seat at Every Table,” Purchasing, Vol. 138, No. 9 (September 17, 2009), 36-46.
5 Teague, Paul, “P&G is King of Collaboration,” Purchasing, Vol. 137, No. 9 (September 11, 2009), 46.
Stay connected between issues.

- Email Newsletters
- Online
- Like us on Facebook
- Magazine
- Follow us on Twitter
- Customer Service
- RSS Feeds