
The Five Traps of Performance Measurement

by Andrew Likierman

In an episode of *Frasier*, the television sitcom that follows the fortunes of a Seattle-based psychoanalyst, the eponymous hero's brother gloomily summarizes a task ahead: "Difficult and boring—my favorite combination." If this is your reaction to the challenge of improving the measurement of your organization's performance, you are not alone. In my experience, most senior executives find it an onerous if not threatening task. Thus they leave it to people who may not be natural judges of performance but are fluent in the language of spreadsheets. The inevitable result is a mass of numbers and comparisons that provide little insight into a company's performance and may even lead to decisions that hurt it. That's a big problem in the current recession, because the margin for error is virtually nonexistent.

So how should executives take ownership of performance assessment? They need to find measures, qualitative as well as quantitative, that look past this year's budget and previous results to determine how the company will fare against its competitors in the future. They need to move beyond a few simple, easy-to-game metrics and embrace an array of more sophisticated ones. And they need to keep people on their toes and make sure that today's measures are not about yesterday's business model.

In the following pages I present what I've found to be the five most common traps in measuring performance and illustrate how some organizations have managed to avoid them. My prescriptions aren't exhaustive, but they'll provide a good start. In any event, they can help you steal a march on rivals who are caught in the same old traps.

Trap 1: Measuring Against Yourself

The papers for the next regular performance assessment are on your desk, their thicket of numbers awaiting you. What are those numbers? Most likely, comparisons of current results with a plan or a budget. If that's the case, you're at grave risk of falling into the first trap of performance measurement: looking only at your own company. You may be doing better than the plan, but are you beating the competition? And what if the estimates you're seeing were manipulated?

To measure how well you're doing, you need information about the benchmarks that matter most—the ones outside the organization. They will help you define competitive priorities and connect executive compensation to relative rather than absolute performance—meaning you'll reward senior executives for doing better than everyone else.

The trouble is that comparisons with your competitors can't easily be made in real time—which is precisely why so many companies fall back on measurements against the previous year's plans and budgets. You have to be creative about how you find the relevant data or some proxy for them.

One way is to ask your customers. Enterprise, the car-rental company, uses the Enterprise Service Quality Index, which measures customers' repeat purchase intentions. Each branch of the company telephones a random sample of customers and asks whether they will use Enterprise again. When the index goes up, the company is gaining market share; when it falls, customers are taking their business elsewhere. The branches post results within two weeks, put them next to profitability numbers on monthly financial statements, and factor them into criteria for promotion (thus aligning sales goals and incentives).

Of course you have to make sure you don't annoy your customers as you gather data. Think about how restaurant managers seek feedback about the quality of their service: Most often they interrupt diners' conversations to ask if everything is OK; sometimes they deliver a questionnaire with the bill. Either approach can be irritating. Danny Meyer, the founder of New York's Union Square Hospitality Group, gets the information unobtrusively, through simple observation. If people dining together in one of his restaurants are looking at one another, the service is probably working. If they're all looking around the room, they may be wowed by the architecture, but it's far more likely that the service is slow.

Another way to get data is to go to professionals outside your company. When Marc Efron, the vice president of talent management for Avon Products, was trying to determine whether his company was doing a good job of finding and developing managers, he came up with the idea of creating a network of talent management professionals. Started in 2007, the New Talent Management Network has more than 1,200 members, for whom it conducts original research and provides a library of resources and best practices.

Trap 2: Looking Backward

Along with budget figures, your performance assessment package almost certainly includes comparisons between this year and last. If so, watch out for the second trap, which is to focus

on the past. Beating last year's numbers is not the point; a performance measurement system needs to tell you whether the decisions you're making now are going to help you in the coming months.

Look for measures that lead rather than lag the profits in your business. The U.S. health insurer Humana, recognizing that its most expensive patients are the really sick ones (a few years back the company found that the sickest 10% accounted for 80% of its costs), offers customers incentives for early screening. If it can get more customers into early or even preemptive treatment than other companies can, it will outperform rivals in the future.

The quality of managerial decision making is another leading indicator of success. Boards must assess top executives' wisdom and willingness to listen. Qualitative, subjective judgments based on independent directors' own experience with an executive are usually more revealing than a formal analysis of the executive's track record (an unreliable predictor of success, especially for a CEO) or his or her division's financial performance. (See "Evaluating the CEO," by Stephen P. Kaufman, HBR October 2008.)

It may sound trite, but how the company presents itself in official communications often signals the management style of top executives. In August 2006 the *Economist* reported that Arijit Chatterjee and Donald Hambrick, of Pennsylvania State University, had devised a narcissism index on which to rate 105 company bosses, based on the prominence of the CEO's photo in the annual report, his or her prominence in press releases, the frequency of the first person singular in interviews with the CEO, and his or her compensation relative to that of the firm's second-highest-paid executive.

Finally, you need to look not only at what you and others are doing but also at what you aren't doing. The managers of one European investment bank told me that they measure performance by the outcomes of deals they've turned down as well as by the outcomes of deals they've won. If the ones they've rejected turn out to be lemons, those rejections count as successes. This kind of analysis seems obvious once stated, but I've noticed a persistent bias in all of us to focus on what we do over what we don't do. Good management is about making choices, so a decision not to do something should be analyzed as closely as a decision to do something.

Trap 3: Putting Your Faith in Numbers

Good or bad, the metrics in your performance assessment package all come as numbers. The problem is that numbers-driven managers often end up producing reams of low-quality data. Think about how companies collect feedback on service from their customers. It's well known to statisticians that if you want evaluation forms to tell the real story, the anonymity of the respondents must be protected. Yet out of a desire to gather as much information as possible at points of contact, companies routinely ask customers to include personal data, and in many cases the employees who provided the service watch them fill out the forms. How surprised should you be if your employees hand in consistently favorable forms that they themselves collected? Bad assessments have a tendency to mysteriously disappear.

Numbers-driven companies also gravitate toward the most popular measures. If they're looking to compare themselves with other companies, they feel they should use whatever measures others use. The question of what measure is the right one gets lost. Take Frederick Reichheld's widely used Net Promoter Score, which measures the likelihood that customers will recommend a product or service. The NPS is a useful indicator only if recommendations play the dominant role in a purchase decision; as its critics point out, customers' propensity to switch in response to recommendations varies from industry to industry, so an NPS is probably more important to, say, a baby-food manufacturer than to an electricity supplier.

Similar issues arise about the much touted link between employee satisfaction and profitability. The Employee-Customer-Profit Chain pioneered by Sears suggests that more-satisfied employees produce more-satisfied customers, who in turn deliver higher profits. If that's true, the path is clear: Keep your employees content and watch those profits soar. But employees may be satisfied mainly because they like their colleagues (think lawyers) or because they're highly paid and deferred to (think investment bankers). Or they may actually enjoy what they do, but their customers value price above the quality of service (think budget airlines).

A particular bugbear of mine is the application of financial metrics to nonfinancial activities. Anxious to justify themselves rather than be outsourced, many service functions (such as IT, HR, and legal) try to devise a return on

investment number to help their cause. Indeed, ROI is often described as the holy grail of measurement—a revealing metaphor, with its implication of an almost certainly doomed search.

Suppose an HR manager undertakes to assign an ROI number to an executive training program. Typically, he or she would ask program participants to identify a benefit, assign a dollar value to it, and estimate the probability that the benefit came from the program. So a benefit that is worth \$70,000 and has a 50% probability of being linked to the program means a program benefit of \$35,000. If the program cost \$25,000, the net benefit is \$10,000—a 40% ROI.

Think about this for a minute. How on earth can the presumed causal link be justified? By a statement like "I learned a production algorithm at the program and then applied it"? Assessing any serious executive program requires a much more sophisticated and qualitative approach. First you have to specify ahead of time the needs of the program's stakeholders—participants, line managers, and sponsors—and make sure that the syllabus meets your organizational and talent management objectives. Once the program has ended, you have to look beyond immediate evaluations to at least six months after participants return to the workplace; their personal feedback should be incorporated in the next annual company performance review. At the soft drinks company Britvic, HR assesses its executive coaching program by tracking coachees for a year afterward, comparing their career trajectories with those of people who didn't get coached.

Trap 4: Gaming Your Metrics

In 2002 a leaked internal memo from associates at Clifford Chance, one of the world's largest law firms, contended that pressure to deliver billable hours had encouraged its lawyers to pad their numbers and created an incentive to allocate to senior associates work that could be done by less expensive junior associates.

Lawyers aren't the only ones: A number of prominent companies have been caught trying to manipulate their numbers. Since 2004 Royal Dutch Shell has paid \$470 million to settle lawsuits relating to its overstatement of reserves. Morgan Stanley was reportedly willing to lose €20 million on a securities trade for the Finnish government just before closing its books for 2004 in order to improve its position in the league table for global equity capital market rankings.

You can't prevent people from gaming numbers, no matter how outstanding your organization. The moment you choose to manage by a metric, you invite your managers to manipulate it. Metrics are only proxies for performance. Someone who has learned how to optimize a metric without actually having to perform will often do just that. To create an effective performance measurement system, you have to work with that fact rather than resort to wishful thinking and denial.

It helps to diversify your metrics, because it's a lot harder to game several of them at once. Clifford Chance replaced its single metric of billable hours with seven criteria on which to base bonuses: respect and mentoring, quality of work, excellence in client service, integrity, contribution to the community, commitment to diversity, and contribution to the firm as an institution. Metrics should have varying sources (colleagues, bosses, customers) and time frames. Mehrdad Baghai and coauthors described in "Performance Measures: Calibrating for Growth" (*Journal of Business Strategy*, July–August 1999) how the Japanese telecommunications company SoftBank measured performance along three time horizons. Horizon 1 covered actions relevant to extending and defending core businesses, and metrics were based on current income and cash flow statements. Horizon 2 covered actions taken to build emerging businesses; metrics came from sales and marketing numbers. Horizon 3 covered creating opportunities for new businesses; success was measured through the attainment of preestablished milestones. Multiple levels like those make gaming far more complicated and far less likely to succeed.

You can also vary the boundaries of your measurement, by defining responsibility more narrowly or by broadening it. To reduce delays in gate-closing time, Southwest Airlines, which had traditionally applied a metric only to gate agents, extended it to include the whole ground team—ticketing staff, gate staff, and loaders—so that everyone had an incentive to cooperate.

Finally, you should loosen the link between meeting budgets and performance; far too many bonuses are awarded on that basis. Managers may either pad their budgets to make meeting them easier or pare them down too far to impress their bosses. Both practices can destroy value. Some companies get around the problem by giving managers leeway. The office supplier Staples, for example, lets them exceed their

budgets if they can demonstrate that doing so will lead to improved service for customers. When I was a CFO, I offered scope for budget revisions during the year, usually in months three and six. Another way of providing budget flexibility is to set ranges rather than specific numbers as targets.

Trap 5: Sticking to Your Numbers Too Long

As the saying goes, you manage what you measure. Unfortunately, performance assessment systems seldom evolve as fast as businesses do. Smaller and growing companies are especially likely to fall into this trap. In the earliest stages, performance is all about survival, cash resources, and growth. Comparisons are to last week, last month, and last year. But as the business matures, the focus has to move to profit and the comparisons to competitors.

It's easy to spot the need for change after things have gone wrong, but how can you evaluate your measures before they fail you? The answer is to be very precise about what you want to assess, be explicit about what metrics are assessing it, and make sure that everyone is clear about both.

In looking for a measure of customer satisfaction, the British law firm Addleshaw Booth (now Addleshaw Goddard) discovered from a survey that its clients valued responsiveness most, followed by proactiveness and commercial-mindedness. Most firms would interpret this finding to mean they needed to be as quick as possible. Addleshaw Booth's managers dug deeper into the data to understand more exactly what "responsiveness" meant. What they found was that they needed to differentiate between clients. "One size does not fit all," an employee told me. "Being responsive for some clients means coming back to them in two hours; for others, it's 10 minutes."

The point is that if you specify the indicator precisely and loudly, everyone can more easily see when it's not fit for the purpose. The credit-rating agencies have come under attack because they gave AAA ratings to so many borrowers who turned out to be bad risks. The agencies have argued in their own defense that lenders misunderstood what the ratings meant. The AAA rating, they claim, was awarded on the basis of borrowers' credit records, and it described the likelihood of default under normal market conditions; it did not factor in what might happen in the event of a massive shock to the finan-

cial system. Reasonable as this explanation may be, it is no consolation to those who thought they knew what the magic AAA represented.

...

Why do organizations that excel in so many other ways fall into these traps? Because the people managing performance frameworks are generally not experts in performance measurement. Finance managers are proficient at tracking expenses, monitoring risks, and raising capital, but they seldom have a grasp of how operating realities connect with performance. They are precisely the people who strive to reduce judgments to a single ROI number. The people who understand performance are line managers—who, of course, are crippled by conflicts of interest.

A really good assessment system must bring finance and line managers into some kind of meaningful dialogue that allows the company to benefit from both the relative independence of the former and the expertise of the latter. This sounds straightforward enough, but as anyone who's ever worked in a real business knows, actually doing it is a rather tall order. Then again, who says the CEO's job is supposed to be easy?

Andrew Likierman is the dean of London Business School, a nonexecutive director of Barclays Bank, and the chairman of the UK's National Audit Office.

Reprinted from *Harvard Business Review*, October 2009
([product #R0910L](#)).