



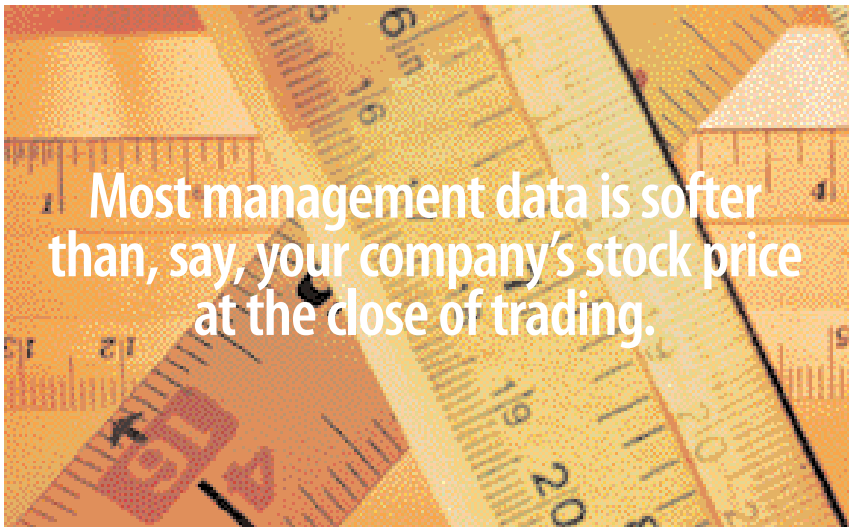
Management's Great Addiction

**It's time we recognized
that we just can't
measure everything.**

By Susan Webber

Corporate America is obsessed with numbers. Analyst meetings focus on earnings expectations, revenue growth, and margins rather than business fundamentals. PowerPoint presentations look naked if they lack charts and graphs to buttress their three-point message. Lofty corporate mission statements are often trumped by pressure to “hit the targets.” Job applicants are advised to stress tangible achievements and, above all, to quantify them. And the ultimate sign of a trend past its sell-by date: a January 2006 *Business Week* cover story, “Math Will Rock Your World.”

This love affair with figures increasingly looks like an addiction. Numbers serve to analyze, justify, and communicate. But they are, fundamentally, abstractions. When



Most management data is softer than, say, your company's stock price at the close of trading.

numbers begin to assume a reality of their own, independent of the reality they are meant to represent, it's time for a reality check. Some are already frustrated with the trend: In a recent McKinsey survey of more than a thousand public-company directors, most said they wanted to hear less about financial results and more about things not so readily quantified, such as strategy, risks, leadership development, organizational issues, and markets.

Metrics presuppose that situations are orderly, predictable, and rational. When that tenet collides with situations that are chaotic, nonlinear, and subject to the force of personalities, that faith—the belief in the sanctity of numbers—often trumps seemingly irrefutable facts. At that point, the addiction begins to have real-world consequences.

Business managers must recognize the limitations of metrics. Mind you, I'm not arguing that metrics and measurement are inherently bad things. To note just one example, a well-structured performance-measurement system is essential to the well-being of large enterprises. But quantitative measures can be and frequently are used naïvely. It's all too easy to abdicate judgment to the output of a model or scorecard. And even when we recognize that certain measurements are in-

complete, we often reflexively strive to make the model more elaborate, rather than exploring other approaches that might yield more insight.

In an ideal world, a team of experts could draw up a roadmap for the proper use of metrics—what should be measured and when, what are the best ways to measure, how should they be interpreted. But whether quantitative measures are beneficial, irrelevant, or counterproductive depends less on the measurements themselves than on how the people who use them integrate the numbers into decision processes. Quick and dirty calculations can be helpful if users understand their limitations; precise models can be dangerous if given undue credence.

Why We Love Numbers

Math-based techniques have, of course, led to important advances in business. Modern commerce would be impossible without them. But history demonstrates that, as with any good thing taken to excess, metrics are prone to overuse.

Frederick Taylor, whose 1911 *Principles of Scientific Management* introduced time-and-motion studies and job descriptions, accelerated the development of large-scale manufacturing enterprises.

Henry Ford's search for efficiency not only created process engineering and, with it, the assembly line—his drive to achieve mass production and economies of scale fostered administrative innovations such as logistics planning, standard operating procedures, and functional administration design.

But Taylorism also devalued workers, in effect treating them as machines. Some scholars believe that this dehumanization of the workplace strengthened and radicalized the union movement, a development that plagued industry for the next fifty years. Similarly, some companies took the Fordist pursuit of scale economies to the point where they lost strategic flexibility. For example: Coca-Cola built up a sizeable inventory of its distinctive six-ounce bottles in the 1930s and was unable to respond when a struggling, twice-bankrupt Pepsi, using recycled beer bottles, was able to sell a ten-ounce drink for a nickel.

Other techniques were simply applied too broadly. During the 1970s heyday of econometrics, most large companies had an in-house econometrist and a penchant for modeling problems, when simple back-of-the-envelope calculations often would have sufficed.

We have a cultural bias in favor of science and mathematics. We see numbers as “hard” outputs: objective, reliable, repeatable, verifiable. But most management data is softer than, say, your company's stock price at the close of trading. Even if we understand those limitations intellectually, we somehow lose that perspective when we wrestle with figures.

In fact, we have a romanticized view of not only management information but of science itself; we attribute to science a rigor and degree of accuracy that would give any scientist pause. This cognitive bias has been repeatedly discussed

SUSAN WEBBER is founder of Aurora Advisors, a New York-based management-consulting firm. Her last article was “The Incredible Shrinking Corporation” in the Nov/Dec 2005 issue.

in scientific literature. In 1961, philosopher of science Michael Scriven called “inaccuracy” the key attribute of physical laws because “its almost universal presence is a kind of unadmitted shocking fact like the Emperor’s nakedness, and needs to be pointed out if we are to get a true picture of the role of laws.” Mind you, this article refers to fields we regard as scientifically mature, such as Newtonian physics. Philosopher Nancy Cartwright’s 1983 book *How the Laws of Physics Lie* takes Scriven’s arguments further.

In reality, matters that laypeople may assume are settled and even obvious—such as what constitutes the nature of proof—are open questions. In addition, statistical inferences, which are the type of analysis most commonly used in business, are not conclusive. First, under even the best of circumstances, measurements are not perfectly accurate. Second, the sample chosen for study may not truly represent the population as a whole. Third, correlation is not causation: There may be other factors at work, and the ones we have focused on may be secondary or even incidental—recall how ulcers were once believed to be caused by diet and stress?

Since scientific findings are less solid than many of us would like to believe, it’s prudent to regard management “findings” with a healthy dose of skepticism.

What We Get Wrong

Since metrics provide a window for viewing our world, how can we recognize when the glass is clouded? Watch for:

Focusing on numbers rather than behaviors. In the management-information game, “how much” is easy to capture, when “how” can be more illuminating. Too often, companies unwittingly mimic the drunk looking for his lost keys under the streetlight, because that is where he can see well, rather than where he lost them.

Take R&D spending. It is an article of faith that companies should increase their R&D budgets if they want more new products. Yet a 2006 Booz Allen survey of the top thousand U.S. corporations, measured by their R&D spending, found “no discernible statistical relationship between R&D spending levels and nearly all measures of business success, including sales growth, gross profit, operating profit, enterprise profit, market capitalization or total shareholder return.” MIT researcher Michael Schrage, an expert on innovation and modeling, argues in a recent *Financial Times* article that R&D spending is similarly unrelated to innovation. He cites examples such as Illinois Tool Works and Reckitt Benckiser, an Anglo-Dutch cleaning-products company, highly innovative organizations that each spend only 1 percent of sales on R&D (versus a European figure of 3.3 percent and a U.S. average of 4.5 percent). Schrage also cites Apple Computer, whose R&D spending of 5.9 percent of revenues considerably lags the industry norm of 7.6 percent. Conversely, GM has spent more on R&D over the last quarter-century than any company on earth, and its flirtation with bankruptcy shows how little this outlay has produced.

Framing the problem incorrectly. Sometimes mistakes are glaringly obvious, once they are pointed out. A financial-services company regularly surveyed its network partners on their satisfaction, measured by ratings on various attributes of the service, and then would try to improve the low scores. No one bothered to ascertain which aspects of service were important to these partners. As a result, considerable effort was spent improving low ratings in categories that had no impact on the relationship.

Or, since businessmen often liken competitive struggles to combat, consider a military example: the Vietnam War. Two key measures used to measure progress were the

“body count,” meaning enemy deaths, and “hamlets under GVN [government of Viet Nam] control.”

The United States saw the problem in conventional-warfare terms, of gaining territory and thinning the enemy’s ranks. But this viewpoint turned out to be woefully misguided. First, those reporting the body-count totals often exaggerated, sometimes by a considerable margin. Second, we didn’t know whom we were killing: Were they really VC, or local sympathizers, or just civilians caught in the crossfire? The more we killed non-combatants, the more we alienated the population and facilitated VC recruitment. Third, and probably most important, we misunderstood the fundamental nature of the war. The North Vietnamese saw it as a war of liberation, to eject yet another colonial power. U.S. decision-makers hugely underestimated the Vietnamese will. For instance, Rand experts who had dealt with prisoner-interrogation material from World War II, Korea, and Eastern Europe had never seen interviews like the ones of VC, and concluded that unlike other opponents, they could not be coerced. Thus, the body counts were relevant only to measure the progress toward exterminating the entire population—if that qualified as “progress.”

If possible, the “hamlets under GVN control” stats were even more dubious. These figures were reported by the Vietnamese government, which obviously wanted to maintain U.S. sponsorship. Yet the U.S. government took these reports at face value, ignoring objections of experienced American operatives. These measures allowed the command structure to assert that the North Vietnamese effort was on the verge of collapse, until the Tet Offensive of 1968 dramatically demonstrated otherwise.

Overlooking perverse incentives and feedback loops. In 1975, management expert Steven Kerr wrote the classic “On the Folly of

Rewarding A, While Hoping for B.” The article describes a range of “fouled up” incentive systems in athletics, academia, medicine, the military, and business.

Kerr (who currently oversees leadership development at Goldman Sachs) lists the causes of these misguided incentives, two of which are particularly germane. The first is “fascination with an ‘objective’ criterion.” While managers prefer simple, quantifiable standards, those standards tend to work only in areas in which the activities are highly predictable, and break down elsewhere. Second is “overemphasis on highly visible behaviors,” which tends to encourage individual action at the expense of activities such as creativity and team-building, which are difficult to observe.

For instance, there has been a great deal of teeth-gnashing about the corporate fixation with quarterly earnings targets, which are objective, to the detriment of long-term competitiveness, which is harder to assess. A 2005 *McKinsey Quarterly* article, “Building the Healthy Corporation,” describes how some companies have responded by developing scorecards for “performance and health.” It recommends general measures in five areas, with particular emphasis on metrics. Although the article sets forth a coherent, wide-ranging program, it is inadequate to the task. A problem that may have started with metrics is not necessarily solvable through metrics, or even metrics plus exhortation.

The McKinsey article gives CEOs the hope that if they retool their systems and encourage silo-ized managers to play together, they can redirect their managers’ actions. And normally, that would be a good assumption. But in this case, the “health” program has two major hurdles to overcome. First, most of the health measures will be seen as soft, even if they are quantified (What does it mean to raise customer satisfaction from 3.2 to 3.5?

What is that worth?), and as discussed earlier, “soft” measures are generally taken less seriously than “hard” ones, like costs. Second, placating Wall Street has become an overarching objective in corporate America, reinforced daily in the business press. Even though the McKinsey piece advocates educating analysts to the payoff to be gained by paying attention to these health metrics, that sort of persuasion is an uphill battle. Similarly, the article also recommends cultivating new investors, ones more long-term-oriented. Short of going private, it’s hard to see how to put that into effect.

The fact is: No middle-level manager is going to do things differently unless he gets an unambiguous signal from the top, like Costco’s explicit rejection of analyst calls to extract more short-term profit. Until top management demonstrates that it will not slavishly bow to the dictates of the financial community, its efforts to persuade the ranks otherwise are likely to fail.

Another danger is that of feedback loops. Measurement systems are often self-referential, and participants can, sometimes innocently, game the system. The seemingly unending rise in CEO pay shows how this process operates. You know the drill: A compensation survey determines what “comparable” CEOs earn, and the CEO’s package is set in reference to this universe. But just as the children at Lake Wobegon are all above average, the CEO is understandably reluctant to be paid in the bottom half, and his board isn’t about to argue. So there is a mechanism in place to keep moving the averages upward, despite the weak to nonexistent correlation between CEO pay and performance.

Why does this persist? One culprit is the social dynamic among the directors and compensation consultants. But a contributing factor is that CEO pay is “market” pay, and “market” data is seen as objective,

even virtuous. But this notion quickly breaks down under scrutiny. In a real market, such a price rise would fuel a search for substitutes, which in this case would be dark-horse candidates, such as business-unit managers who might have delivered great performance but be wanting in polish. (Liz Claiborne chairman and CEO Paul Charron, who had neither CEO nor apparel experience before he took the helm, is the exception that proves the rule.) Similarly, few boards are willing to consider whether their CEO could really get a comparable package anywhere else.

Misreading the data. Even when you have the right metrics, you may not interpret them correctly. And since this happens to the best and the brightest, it can certainly happen to you.

You’ll recall the Long-Term Capital Management debacle in 1998: A high-flying hedge fund with the industry’s finest analytical talent, including two Nobel laureates, had to be bailed out by a consortium of banks to prevent wide-scale disruption of the financial markets. Although there are different views of why the firm collapsed (the big reason is that it began to trade in markets in which it had limited experience), some argue that it was a “perfect storm” that perhaps no one could have anticipated. A more jaundiced and persuasive view is that LTCM’s models assumed a normal, bell-curve distribution of events, when in fact markets often exhibit both an asymmetrical distribution and “fat tails” (meaning that events “far” from the mean in a statistical sense have a greater likelihood of happening than assumed by a normal distribution). Given the huge bets LTCM was taking in unfamiliar waters, it would have seemed a reasonable precaution to stress-test its models using other distributions of events.

A 2006 Malcolm Gladwell article discusses how various organizational and social issues have re-

mained unsolved because the remedies assumed a normal distribution, when in fact the problem had a “power law” distribution (think of it as 80/20 on steroids). For example, the Los Angeles Police Department studied complaints about the use of excessive force. They had expected to see the complaints broadly distributed across the entire LAPD, suggesting that more training and better procedures were the answer. Much to its surprise, the LAPD found instead that the complaints were concentrated among a very few officers. The solution was to fire them or, at the very least, get them off the street.

Another factor is cognitive bias. The field of behavioral finance has analyzed the many ways that people fail to deal rationally with numbers. One phenomenon is anchoring. Individuals’ estimates are influenced by random suggestion. In one famous experiment, a roulette wheel generated an illustrative, and clearly arbitrary, value when participants were asked to estimate the percentage of U.N. countries that are in Africa. High numbers on the wheel elicited considerably higher guesses.

Or consider a more relevant illustration: acquisitions. Inevitably, the discussion of the pending deal revolves around the projections. The financial model assumes a reality of its own. Anything that isn’t incorporated in the model is implicitly assumed away.

Buyers continue to use this approach despite evidence that it produces bad outcomes. Depending on which study you choose, every analysis of mergers says that most fail (typical estimates are in the 60 to 75 percent range), and the buyer overpaying is the most frequent cause. Yet the “anchor” of the forecasts is powerful, and woefully difficult to dislodge.

Making Math Work

The most important change a senior executive can make is a shift

in mindset, to regard figures as a useful input rather than gospel. It helps to recognize when these quantitative measures are most valid (e.g., when applied to discrete processes that can be measured objectively, such as transaction processing) and when they are more tenuous. To make sure you maintain and reinforce a healthy skepticism:

Perform retrospective reviews. Postmortems are standard practice in sport and in medicine, but they are virtually nonexistent in business. A noteworthy exception: A major financial institution is analyzing the results of its bonus and promotion process to see if they in fact reward the behaviors that senior management believes they are rewarding.

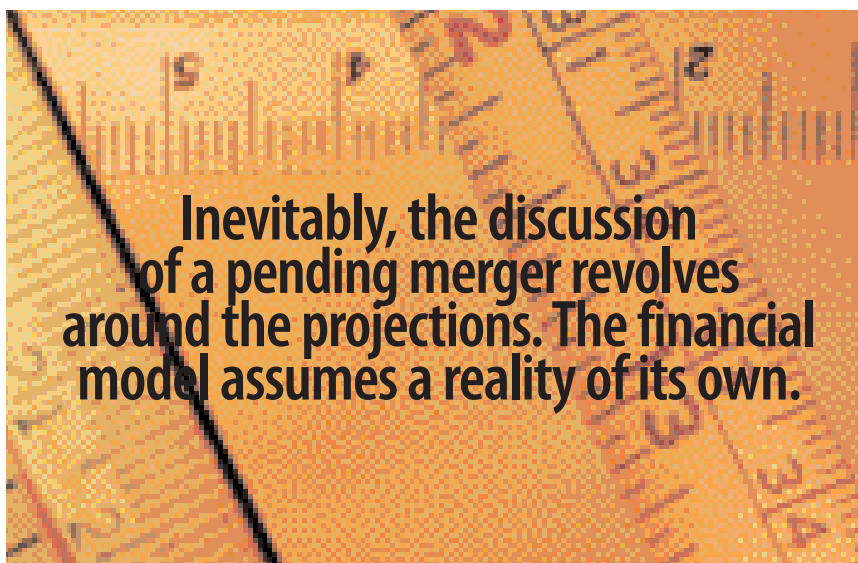
Most companies would learn a great deal if they looked at, for instance, all their capital-investment decisions (both the projects approved and the ones turned down) over a given four-year period—from, say, 1999 to 2002—to see which decisions were good, which were not so astute, and what, if anything, could have been done to improve the decision process.

Question the logic. Too often, managers are reluctant to ask how certain analyses were derived for fear of appearing ignorant. Yet it is important to inquire, particularly

for one-off studies, what analytic methods were used and what assumptions were implicit in the use of that methodology. For example, a regression analysis assumes a linear relationship between the variables. What if the relationship is actually a step function? Trying to fit a regression to the data would produce misleading results. Similarly, “real options” have become popular as a way of valuing investment opportunities. But option valuation is a tricky business. According to Black-Scholes, some variables, such as the length of the option and the implied volatility, have a significant impact on the option price. It is critical to understand the rationale for the use of the chosen values and run sensitivity analyses around those assumptions.

If you don’t have the appetite for this line of inquiry, look outside and engage someone with advanced math skills, such as a doctoral candidate in applied mathematics, to serve as a house skeptic.

Probe the data. Again, all data is not created equal. Generally speaking, the most reliable information is that about physical activity. Financial and accounting data is less solid, and metrics on consumer behavior are even trickier. It is notoriously difficult to ascertain whether and why consumers will buy a product, since their responses are very





American corporations have for some time been engaged in what can well be described as cynical behavior.

much influenced by the test environment. Anyone who has done survey research, for example, will confirm that results can be skewed significantly by how a given question is phrased. Similarly, researchers have found that when consumers are given a taste test and asked to rate, say, which salsa they like best, their answers are completely different if they are asked to rate the various attributes (spiciness, texture, etc.) and then say which they prefer. In a bizarre analogy to the Heisenberg uncertainty principle, the act of making the consumer explain why he likes a product shifts his choice.

And probably the most slippery data of all is personnel assessments, where most large companies force subjective and, despite their efforts, not very comparable information into tidy grids and rankings. (Doubtters should read Patrick D. Larkey and Jonathan P. Caulkins' 1992 paper "All Above Average and Other Unintended Consequences of Performance Evaluation Systems," a provocative and well-documented indictment.)

The remedy, particularly for important decisions, is to understand the factual underpinnings and be willing to invest in additional research. For example, Coca-Cola considered New Coke a shoo-in because it scored so well on sip tests compared to other colas—including

traditional Coke. But consumers don't buy soda in sips, they buy entire cans, and sip tests favor sweeter drinks. New Coke bombed because—among other reasons—many target customers found it to be cloying. Had Coca-Cola used multiple approaches to vet New Coke's consumer appeal, they might well have surfaced its shortcomings.

Be alert to new information. A persistent mistake is attachment to an old perception of a situation (another manifestation of anchoring). Although it's fashionable to blame lumbering corporate-reporting systems that filter out bad news, this tendency to dismiss new data is a well-documented *individual* behavior. Recall Thomas Kuhn's *The Structure of Scientific Revolutions*: Scientists who grew up with an old paradigm simply could not accept a new model. A whole generation had to die out before a new theory, no matter how well proven, became widely accepted.

How can you overcome this cognitive inertia? By demonstrating keen interest in new developments—and fostering that attitude in others. The tried-and-true approach of talking directly to customers is invaluable. It also helps to ask front-line staff often about trends and developments they see, and encourage them to pass them along. You will get a lot of noise along with some choice nuggets, but in this

case, the discipline of cultivating awareness and mental flexibility is as important as any bits of intelligence you glean.

Consult your gut. A recent study published in the journal *Science* found that for complex decisions (defined in this study as involving twelve variables, versus four for the "simple" decision), unconscious decision processes yielded much better results than trying to "reason it out." Our rational mind can comprehend a limited amount of data, while our unconscious processes, honed over thousands of years of evolution, are better at dealing with complicated situations. In these cases, studying the data and sleeping on it produces demonstrably better choices.

An over-reliance on metrics can lead to "knowing the price of everything and the value of nothing." Take heed: That's how Oscar Wilde defined a cynic, and cynicism is not viewed favorably in most organizations.

Yet American corporations have for some time been engaged in what can well be described as cynical behavior: taking aggressive accounting measures, engaging in short-term expediencies to improve results, too often displaying little concern for the impact of their actions on employees and communities. Now, it is no doubt a stretch to blame these actions on the use of numbers. But the two do seem to go hand in hand.

Management is the art of making decisions in the face of uncertainty. Statistics and analysis can help us understand the nature of that uncertainty and dimension of the risks we are taking, but they can also provide false comfort and engender undue confidence. Perhaps the biggest obstacle to corporate America giving up the metrics habit is that it will require executives to acknowledge their limitations. But the benefits—however difficult to quantify—will be worth it. ♦