

Tracking things like customer satisfaction and employee turnover can powerfully supplement traditional bookkeeping. Unfortunately, most companies botch the job.

Coming Up Short on Nonfinancial Performance Measurement

by Christopher D. Ittner and David F. Larcker

Tracking things like customer satisfaction and employee turnover can powerfully supplement traditional bookkeeping. Unfortunately, most companies botch the job.

Coming Up Short on Nonfinancial Performance Measurement

by Christopher D. Ittner and David F. Larcker

In the past decade, increasing numbers of companies have been measuring customer loyalty, employee satisfaction, and other performance areas that are not financial but that they believe ultimately affect profitability. Doing so can offer several benefits. Managers can get a glimpse of the business's progress well before a financial verdict is pronounced and the soundness of their investment allocations has become moot. Employees can receive better information on the specific actions needed to achieve strategic objectives. And investors can have a better sense of the company's overall performance, since nonfinancial indicators usually reflect realms of intangible value, such as R&D productivity, that accounting rules refuse to recognize as assets.

But the reality is that only a few companies realize these benefits. Why? Because they fail to identify, analyze, and act on the right nonfinancial measures. We conducted field research in more than 60 manufacturing and service companies and supplemented it with survey responses from 297 senior executives. To our

surprise, we discovered that most companies have made little attempt to identify areas of nonfinancial performance that might advance their chosen strategy. Nor have they demonstrated a cause-and-effect link between improvements in those nonfinancial areas and in cash flow, profit, or stock price.

Instead, many companies seem to have adopted boilerplate versions of nonfinancial measurement frameworks such as Kaplan and Norton's Balanced Scorecard, Accenture's Performance Prism, or Skandia's Intellectual Capital Navigator. And yet the frameworks' own inventors rightly insist that every company needs to dig deep to discover and track the activities that truly affect the frameworks' broad domains (domains such as "financial," "customers," "internal business processes," and "innovation and learning," in the case of the Balanced Scorecard).¹ But businesses often fail to establish such links partly out of laziness or thoughtlessness. As a result, self-serving managers are able to choose—and manipulate—measures solely for the purpose of making

themselves look good and earning nice bonuses.

How mindless or mendacious can managers be? Here are some examples:

- One of the world's top information-service providers began evaluating managers' performance according to how many patents the company filed each year. Whether it might have made more sense to license someone else's technology, whether the patents were ever put to work, or whether they ever earned back their cost was not considered. The reason for tracking patent awards? A more successful competitor owned a larger number.

- A large retail bank decided to base bonuses on customer satisfaction scores. But the polling company hired by the bank surveyed only those customers who physically entered bricks-and-mortar branches. So one branch manager who had received poor satisfaction scores in the past coaxed customers to visit and then put smiles on their faces by offering them free food and drinks on the premises.

- Managers of an automobile components manufacturer reached the firm's quality targets by reclassifying as acceptable certain flaws that once would have caused a part to be rejected.

When such things happen, a company's financial and nonfinancial performance diverge—an ironic outcome, since the original reason for tracking nonfinancial performance was to fill out the picture provided by traditional financial accounting. And yet, is it so surprising that nonfinancial measures would be equally, if not more, susceptible to manipulation as financial accounting? At least traditional accounting has rules that govern it.

In fact, the misuse of nonfinancial measures may be even more damaging because of the significant opportunity costs incurred. As the exhibit "The Difference It Makes" shows, the companies in our study that adopted nonfinancial measures and then established a causal link between those measures and financial outcomes produced significantly higher returns on assets and returns on equity over a five-year period than those that did not.

In the following pages, we discuss our research findings, which reveal a number of common mistakes companies make when trying to measure nonfinancial performance. We then highlight a number of practices that, in our view, will allow companies to realize the

genuine promise of nonfinancial performance measures.

MISTAKE ONE:

Not Linking Measures to Strategy

Whether the goal of a performance measurement system is to help direct the allocation of resources, to assess and communicate progress toward strategic objectives, or to evaluate managerial performance, a major challenge for companies is determining which of the hundreds, if not thousands, of nonfinancial measures to track.

Many companies believe that they have solved this problem by adopting a framework like the Balanced Scorecard, mistaking it for an off-the-shelf checklist or procedure that is universally applicable and completely comprehensive. But using such a framework by itself won't help identify which performance areas—and which drivers—make the greatest contribution to the company's financial outcomes. In a number of companies we studied, middle managers sarcastically referred to the Balanced Scorecard as the "four bucket" or "smorgasbord" approach because top management ordered them to come up with something for each of the scorecard's four perspectives, regardless of their business unit's strategy or objectives.

More successful companies have attacked this problem by choosing their performance measures on the basis of causal models, also called value driver maps, which lay out the plausible cause-and-effect relationships that may exist between the chosen drivers of strategic success and outcomes. The exhibit "Which Measures Matter" shows how one very successful fast food chain diagrammed its drivers of strategic success. The diagram demonstrates how better employee selection and staffing should lead to higher employee satisfaction and thus improve employee performance. The latter in turn should increase customer satisfaction and thus purchase frequency, customer retention, and referrals, ultimately leading to sustained sales growth and increased shareholder value. This model became the basis for selecting performance measures directly tied to the goals of the strategic plan, which was to become the premier generator of free cash flow in the fast foods sector and lead stock-price performance in that industry.

Despite the apparent logic and good sense

Christopher D. Ittner is a professor of accounting at the Wharton School of the University of Pennsylvania, in Philadelphia. He is coauthor of *Linking Quality to Profits: Quality-Based Cost Management* (ASQC Press, 1994). **David F. Larcker** is the Ernst & Young Professor of Accounting at Wharton.

of making such connections, fewer than 30% of the companies we surveyed have developed causal models, which show what areas are expected to improve as the result of commitments to particular courses of action, and then show how those improvements should affect long-term economic performance.

**MISTAKE TWO:
Not Validating the Links**

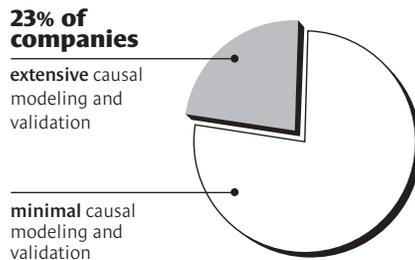
Even those companies that create causal models rarely go on to prove that actual improvements in nonfinancial performance measures affect future financial results. Of the companies we surveyed, only 21% did so. In far too

many cases, management simply relied on its preconceptions about what was important to customers, employees, suppliers, investors, or other stakeholders rather than verifying whether those assumptions had any basis in fact. Overlooked were questions like, Do experienced employees make fewer errors, and, if so, should we do whatever we can to reduce turnover? (Not before testing the hypothesis and determining which employees matter most.) Does accelerating product-development time lead to increased market share? (Not if our new products are only minutely different from our earlier models, or we have merely reverse-engineered those of our competitors.) If companies don't investigate whether there is a plausible causal relationship between actions and outcomes, they condemn themselves to measuring aspects of performance that don't matter very much.

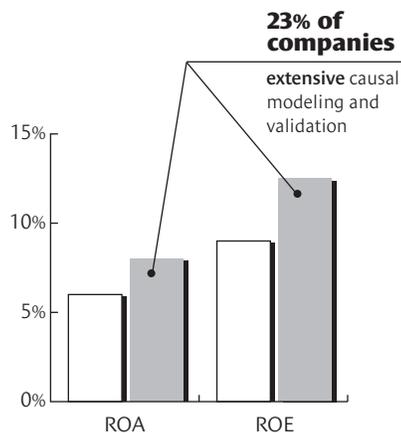
When we asked managers why they didn't try to establish these connections, they often responded that the links were self-evident: Of course improvements in customer loyalty, employee retention, new product introductions, or other common nonfinancial measures lead to higher profits and shareholder value. But unfortunately, our research indicates that such assumptions are often half-baked or wrong. Consider the fast food chain discussed earlier. Before creating its causal model, the company chose employee turnover as a key performance indicator, believing that high employee retention indicated a high level of satisfaction and motivation, which would in turn improve customer service and eventually boost profits. This set of assumptions led the chain to consider implementing a series of costly initiatives, such as cash bonuses and increased benefits at employees' one-year anniversaries, to reduce voluntary turnover. Subsequent analysis, however, found that the profitability of restaurants with identical turnover rates varied dramatically. That's because a 150% annual turnover rate at one restaurant could include turnover of cooks and cashiers as well as management and supervisory personnel, while that same 150% turnover rate in another restaurant could reflect 200% turnover among lower-level workers but only 30% turnover among supervisors. What distinguished profitability was the turnover among supervisors, not among lower-level workers. The company was not wrong in believing that turnover was important. But a

The Difference It Makes

In our survey of 157 companies, only 23% consistently built and verified causal models (diagrams laying out the cause-and-effect relationships between the chosen drivers of strategic success and outcomes).



Yet those 23%, on average, had 2.95% higher ROA and 5.14% higher ROE than companies that didn't use causal models.



Note: Industry-adjusted ROA (or ROE) is computed as five-year ROA (or ROE) minus the median five-year ROA (or ROE) for the company's industry.

Copyright © 2003 Harvard Business School Publishing Corporation. All rights reserved.

failure to investigate whose turnover really mattered nearly led to a substantial waste of resources.

In another case, an information service provider believed that it could improve its service offerings by creating alliances with vendors of technology products. The higher service levels, in turn, were expected to strengthen ties to

customers, who would then, theoretically, purchase more services. The company accordingly went to great lengths to forge alliances and rate its progress at doing so. Yet we could find no evidence that the alliances improved the company's chances of winning new work or having its contracts renewed.

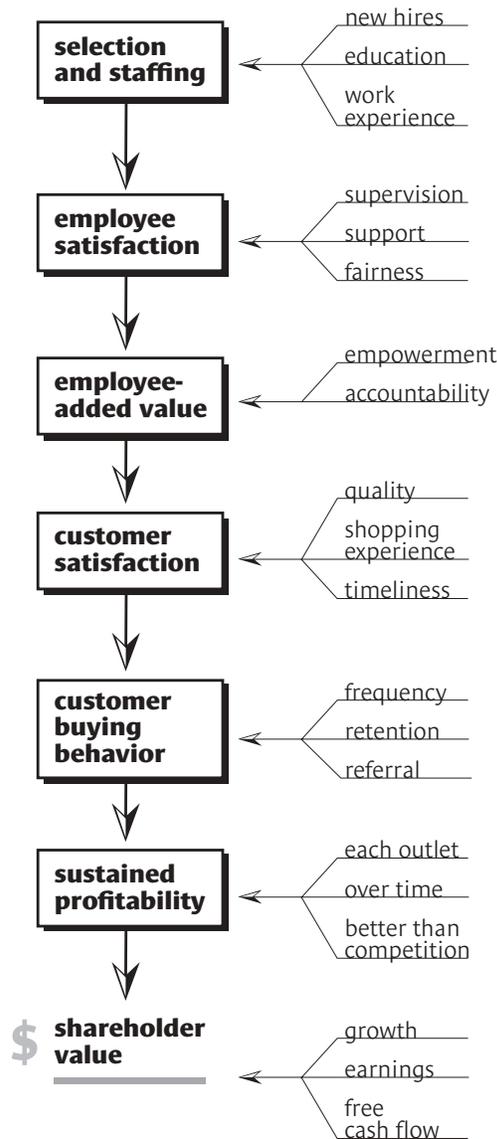
Businesses that do not scrupulously uncover the fundamental drivers of their units' performance face several potential problems. They often end up measuring too many things, trying to fill every perceived gap in the measurement system. The result is a wild profusion of peripheral, trivial, or irrelevant measures. Amid this excess, companies can't tell which measures provide information about progress toward the organization's ultimate objectives and which are noise. A leading home-finance company, for example, implemented an "executive dashboard" that eventually grew to encompass nearly 300 measures. The company's chief operating officer complained, "There's no way I can manage my business with this many measures. What I'd really like to know are the 20 measures that tell me how we are really doing."

If companies can't prove basic causality, they certainly can't determine the relative importance of the measures they select. And not being able to weigh these measures makes it hard to allocate resources according to their most beneficial uses or to create meaningful incentive plans. For instance, does a dollar invested in product development yield higher returns than a dollar spent on customer retention?

In the absence of such knowledge, companies in our study came up with various solutions for assigning relative weights to different measures. One of the simplest solutions was to give each performance measure equal weight. As one executive at a consumer electronics manufacturer put it, "It's difficult to precisely assign weightings, so I just assume they are of equal importance." But perhaps even more often, managers base weightings purely on their assumptions about the measures' strategic importance. Or they stress the measures that have become most fashionable in the business press or among consultants. Or, particularly when bonuses are at stake, they place greater weight on measures whose targets they know they can hit.

Which Measures Matter

One of the companies in our study, a successful fast food chain, developed this causal model proposing the drivers of strategic success. Note that behind the drivers of performance lie the drivers of those drivers: For instance, timeliness and quality drive customer satisfaction, which in turn drives customer buying behavior, which in turn drives profits.



Copyright © 2003 Harvard Business School Publishing Corporation. All rights reserved.

It's not uncommon for business units within the same company to use different methodologies to measure the same thing.

MISTAKE THREE:

Not Setting the Right Performance Targets

Outstanding nonfinancial performance is not always beneficial. Indeed, it often produces diminishing or even negative economic returns—and again, most companies have no idea when they have achieved too much of a good thing.

We studied one company in an unregulated segment of the telecommunications industry in which customers' switching costs were minimal. To hold on to the customers it had, the company set its sights on achieving 100% satisfaction for every one of them. However, the company never attempted to discover whether a correlation actually existed between an individual customer's level of satisfaction and the revenues and profits that customer generated. We discovered, in fact, that the expected relationship did appear—but only up to a point. Customers who were 100% satisfied spent no more money than those who were only 80% satisfied. In short, getting to 100% required considerable investment, with little or no payback. Only by determining the level at which satisfaction ceases to contribute to revenue growth can a business know whether and how much to invest, at any given point, in trying to raise it.

Target setting is inherently difficult because it always takes awhile for improvements in a driver of corporate performance to produce improvements in the performance it's meant to affect. Sometimes, efforts to improve nonfinancial measures can even damage short-term returns. However, if a company can reasonably estimate when the nonfinancial performance improvements will pay off, and by how much, it can set lower interim financial goals, which can subsequently be adjusted upwards. Unfortunately, many companies don't make the effort, preferring to focus on initiatives that promise short-term financial results even though other initiatives may have higher long-term payoffs.

MISTAKE FOUR:

Measuring Incorrectly

Finally, even companies that build a valid causal model and track the right elements can fall down when determining how to measure them. At least 70% of companies, we found, employ metrics that lack statistical validity

and reliability. "Validity" refers to the extent to which a metric succeeds in capturing what it is supposed to capture, while "reliability" refers to the degree to which measurement techniques reveal actual performance changes and do not introduce errors of their own. For example, many companies attempt to assess extremely complex performance dimensions using surveys containing only one or a few questions. The questions may offer respondents only a small number of scale points (for instance, 1 = low, and 5 = high). Many companies then collapse these already simplified answers into crude binary scales (for example, customers are deemed satisfied if the score is 4 or 5, and dissatisfied if the score is 1 through 3). Although inexpensive to use and easy to understand, such simplistic surveys lack validity and reliability and impair companies' ability to discern superior performance or predict financial results.

Many companies also make the mistake of collecting data before deciding what they want to find out. By the time they have identified the level of analysis they want to undertake and the areas of performance they want to compare, the data have already been gathered and organized in a manner that renders the desired analyses impossible. For example, one management-consulting firm we studied tracked customer satisfaction at the individual client level, but employee performance at only the regional level, and operational performance at only the project level, making it impossible to determine how employee performance affected customers or project outcomes. Companies that don't or can't know in advance the correlations they want to explore would do well to assign a unique identifying tag or code to each one of the smallest units measured.

Measures can also lose validity and reliability when the methods for evaluating nonfinancial attributes are inconsistent across the company. We found that business units within the same company often used different methodologies to measure the same thing. One consulting firm we studied enlisted three different internal groups to measure corporate reputation. Each group used a different measurement technique, and each produced very different—indeed, contradictory—results. At another company, an appliance manufacturer, several factories mea-

sured total employee turnover, while others measured only voluntary turnover. Such inconsistencies make it hard for top management to assess overall progress or to compare one unit's performance with another's.

Sometimes the problem lies in the nature of the thing being measured. Most businesses have trouble discovering how they are doing at such elusive endeavors as developing leadership or maintaining supplier relations. Nearly half of all Balanced Scorecard users surveyed by Towers Perrin said they had difficulty quantifying qualitative results. One unfortunate response to these frustrations is to avoid measuring altogether the "hard to measure." In fact, a Conference Board study found that for 55% of the senior executives it surveyed, the leading obstacle to implementing strategic-performance measurement systems was an unwillingness to measure activities that posed this problem. And many of the companies that did try to track more qualitative measures ignored them when making decisions. When we asked managers why they chose to overlook them, the typical response was lack of trust in measures that were unproven and therefore subject to favoritism and bias. Although such wariness saves companies from relying on misleading results, it also denies them a comprehensive picture of their performance.

Doing It Right

At the root of these four mistakes is the failure to discover which nonfinancial factors have

the most powerful effects on long-term economic performance. The root of the solution, therefore, is to base decision making on a well-established series of links. By following the steps listed below, companies should be able to realize the full promise of nonfinancial performance measures.

Develop a causal model. The first step is to develop a causal model based on the hypotheses in the strategic plan. Unfortunately, however, many companies' strategic plans are more like mission or vision statements than road maps. In the absence of strategic clarity and concrete detail, managers are prone to disagree about which performance areas are critical to success, and that can make consensus about the causal model difficult to reach. If that's the case, it's best to test a couple of different causal models. Once its merits have been proven, the model finally chosen will be hard to argue with and will be the source of broad-based agreement about strategy.

Pull together the data. Most companies already track large numbers of nonfinancial measures in their day-to-day operations. So to avoid going to the trouble of collecting data that already exist, companies should take careful inventory of all their databases. This inventory should not limit itself to performance measurement systems but should extend to any information systems (such as purchasing, manufacturing control, and customer service) that may contain useful data on key performance drivers. One important byprod-

Our Research

This article is based on extensive field research into more than 60 manufacturing and service companies, where we interviewed senior and middle managers about their organizational strategies and performance measurement systems. Since the data obtained from these companies are proprietary, we have kept company names confidential.

In 14 of the companies, we investigated the extent to which nonfinancial performance measures that were chosen on the basis of their supposed ability to affect future financial performance actually had the expected relationship. Executives at each of these companies gave us detailed descriptions of the financial and nonfinancial measures they be-

lieved to be the key drivers of their organization's economic performance. We then statistically evaluated whether these drivers really affected their companies' performance.

We supplemented our field research with two surveys of performance measurement practices. The first survey, in which a market research firm obtained data from 157 chief financial officers and other senior executives in a broad range of industries, was conducted jointly by the Wharton School and PricewaterhouseCoopers. The aim of the survey was to determine the types of performance measures these executives used, to what extent they tried to assess the relation-

ship between nonfinancial measures and future economic performance, and to what extent they developed and adopted business models that established causal links between the company's financial and nonfinancial measures.

A second survey of 140 senior executives in the financial services industry (banks, savings and loans, insurance companies, and diversified financial institutions) was conducted jointly by Wharton and Ernst & Young. This survey focused on executives' perceptions of the key nonfinancial drivers of long-term economic performance, as well as the use and adequacy of related nonfinancial performance metrics for those drivers.

Many companies' strategic plans are more like mission or vision statements than road maps.

uct of this step is that it begins the process of refining vague or ambiguous definitions and of developing consistent measures for the organization as a whole.

It may be, however, that a company lacks the data it needs even to formulate a causal model. If that's the case, executives might want to focus first on a performance area believed ultimately to advance the company's strategy and positively affect corporate financial performance (employee satisfaction, say). Next, it might take a small number of actions believed to improve performance within that area (such as more training). The final step would be to precisely and consistently measure the effects of those actions. Did more training actually increase employee satisfaction?

One problem we repeatedly encountered in this step was data "fiefdoms." An automobile manufacturer we studied wanted to determine whether its manufacturing defects were generating too many warranty claims, in which case it would need to change its factory inspections. But the marketing people refused to share their findings with the operations people, making such detective work impossible. Ultimately, a senior executive had to step in.

Turn data into information. There are many statistical methods for testing the causal model. Most companies have experience using correlation analyses and multiple regressions in their market research and quality improvement efforts. A good example of such statistical techniques is an approach used at Sears, which sought to develop a causal model and scorecard focused on three domains: employee relations ("compelling place to work"), customer satisfaction and loyalty ("compelling place to shop"), and results for shareholders ("compelling place to invest"). Like many companies, the retailer had already tracked hundreds of suspected drivers of performance within these domains. Because the data on them came from a large cross section of stores, the company was able to use regression analysis to identify the handful of activities that actually were driving performance and therefore belonged in the causal model.²

In addition to these familiar statistical tools, a slew of other techniques, many developed by marketers, can be used to validate the assumed relationships in the causal model. Qualitative analyses such as focus groups and one-

on-one interviews can test management's hunches about what's important to customers, employees, suppliers, investors, and other stakeholders. For instance, a major industrial gas supplier decided that a primary driver of customer retention was customer satisfaction with the supplier's billing system. Accordingly, the supplier began soliciting bids for a new, improved system. However, interviews with individual customers revealed that the billing process was not a major issue. Far more important was technical assistance. On the strength of this finding, the supplier dropped its plans for the new billing system and directed its capital instead to hiring new technicians and retraining existing ones.

Continually refine the model. Causal modeling, if used at all, is often used only once. But reassessment of results should be ongoing and regular. A new competitive environment can weaken or neutralize the effectiveness of formerly key activities, and the company's strategic response can marginalize once important performance areas.

Even in stable environments, ongoing analysis allows companies to continually refine their performance measures and deepen their understanding of the underlying drivers of economic performance. For example, a company may believe correctly that low employee absenteeism is a key driver of financial performance, but its managers will still need to know whether employees fail to turn up because they are unhappy with their pay, with their working conditions, or for some other reason.

At one information technology company in our study, a cross-functional team conducts analyses of integrated operational, accounting, and customer data every quarter and develops hypotheses about the relationships between particular company efforts and outcomes. For example, what types of customers is the company most likely to lose if operational metrics fall below a certain threshold? Does higher customer satisfaction on some attributes (such as assistance in problem solving or flexibility in meeting changing demands) really lead to higher customer profitability? The hypotheses and associated test results are then presented to senior management. In virtually every meeting, these presentations spark new questions about the underlying drivers of value, which are examined again in the next quarter's data analysis.

In short, the refinement process should be never ending. Beneath the proven drivers of performance lie the drivers of those drivers. Since a business can't ever know whether it's gone deep enough, the effort to uncover these drivers must never cease.

Base actions on findings. Ultimately, the conclusions drawn from data analyses must be used in decision making if nonfinancial performance measures are to improve financial results. And clearly, companies should act on the conclusions that appear to promise the greatest financial reward. For example, a major finance company found that, in ascending order of importance, employee satisfaction, quality (the number of processing mistakes), and customer satisfaction were the fundamental drivers of financial performance. Consequently, the company began requiring managers to base their recommendations for allocating capital according to the drivers' relative importance. It also required them to explain how success in these three realms would be measured and to estimate the financial payback in these three areas.

Assess outcomes. The final step in the performance measurement process is determining whether the action plans and the investments that support them actually produced the desired results. In our research, very few companies did "postaudits" that could confirm whether investments actually paid off. Even if the postaudit showed negative financial outcomes, it would have the positive effect of suggesting revisions to the causal model, and it might expose managers' data-

gathering errors and manipulation efforts.

• • •

The original purpose of nonfinancial performance measures was to fill out the picture provided by traditional financial accounting. Instead, such measures seem to have become a shabby substitute for financial performance. Our study shows that they will offer little guidance unless the process for choosing and analyzing them comes to rely less on generic performance measurement frameworks and managerial guesswork and more on sophisticated quantitative and qualitative inquiries into the factors actually contributing to economic results. Otherwise, having proliferated in prosperous times, such measures are likely to be abandoned in lean ones, along with the managers who are charged with tracking—and justifying—them. 

1. For more on the Balanced Scorecard, see Robert S. Kaplan and David P. Norton, "The Balanced Scorecard—Measures That Drive Performance," HBR January–February 1992; for more on the Intellectual Capital Navigator, see Leif Edvinsson and Michael S. Malone, *Intellectual Capital: Realizing Your Company's True Value by Finding Its Hidden Brainpower* (HarperBusiness, 1997).

2. See Anthony J. Rucci, Steven P. Kim, and Richard T. Quinn, "The Employee-Customer-Profit Chain at Sears," HBR January–February 1998.

Reprint [R0311F](#); *Harvard Business Review* OnPoint [5380](#)

To order, see the next page

or call 800-988-0886 or 617-783-7500

or go to www.hbr.org



Harvard Business Review OnPoint articles enhance the full-text article with a summary of its key points and a selection of its company examples to help you quickly absorb and apply the concepts. *Harvard Business Review* OnPoint collections include three OnPoint articles and an overview comparing the various perspectives on a specific topic.

Further Reading

This article is also available in an enhanced *Harvard Business Review* OnPoint article, (Product no. [5380](#)), which includes a summary of its key points and company examples to help you put the ideas to work. The OnPoint article also includes the following suggestions for further reading:

[Why Hard-Nosed Executives Should Care About Management Theory](#)

Clayton M. Christensen and
Michael E. Raynor
Harvard Business Review
September 2003
Product no. 4937

[The Employee-Customer-Profit Chain at Sears](#)

Anthony J. Rucci, Steven P. Kirn, and
Richard T. Quinn
Harvard Business Review
January–February 1998
Product no. 3537

[Focusing Your Organization on Strategy— with the Balanced Scorecard](#)

Harvard Business Review OnPoint Collection
October 2000
Product no. 5157

To Order

For reprints, *Harvard Business Review* OnPoint orders, and subscriptions to *Harvard Business Review*:
Call 800-988-0886 or 617-783-7500.
Go to www.hbr.org

For customized and quantity orders of reprints and *Harvard Business Review* OnPoint products:
Call Frank Tamoshunas at
617-783-7626,
or e-mail him at
ftamoshunas@hbsp.harvard.edu