

# *Easy Innovation and the IRON CAGE*

Best Practice, Benchmarking,  
Ranking, and the Management  
of Organizational Creativity



H. George Frederickson

*An occasional paper of the  
Kettering Foundation*

---

June 2003

**Program Officer:**

**John Dedrick**

Editor:

Ilse Tebbetts

Copy Editor:

Betty Frecker

Design:

Long's Graphic Design, Inc.

Publisher:

Robert E. Daley

This publication is one of a series of occasional papers published by the Kettering Foundation, 200 Commons Road, Dayton, Ohio 45459-2799. The Kettering Foundation is a nonprofit operating foundation, chartered in 1927, that does research to learn how democracy can work as it should. The foundation does not make grants but welcomes partnerships with other institutions or individuals who are actively working on problems of communities, politics, and education.

The interpretations and conclusions contained in this paper, unless expressly stated to the contrary, represent the views of the author or authors and not necessarily those of the foundation, its trustees, or officers.

**Acknowledgments**

The author would like to thank John Nalbandian, Joseph Freeman, John Bryson, Lawrence O'Toole, and William Gormley. They read and critiqued earlier versions of this essay and it is much the better for it.

Copyright© 2003 by the Kettering Foundation

# EASY INNOVATION AND THE IRON CAGE: BEST PRACTICE, BENCHMARKING, RANKING, AND THE MANAGEMENT OF ORGANIZATIONAL CREATIVITY

H. George Frederickson

*The search for the best is the enemy of the good.*

— Voltaire

The modern belief that corporate innovation can be managed begins with the Xerox Narrative, a master narrative, told and retold in boardrooms across America. It is the tale of how, beginning in the mid-1980s, Xerox saved itself by using what came to be known as best practices and benchmarking. The narrative tells how a group of high-level Xerox “change agents” created the Central Logistics and Assets Management (CLAM) unit, which systematically surveyed Xerox’s competitors, identified the best of their work processes, labeled them “best practices,” and used them as benchmarks against which their own products and processes were measured (Camp, 1995). Xerox stock rose from a low of \$7.80 per share in the early 1990s, to \$68.00 by January 1999. The company won the Malcolm Baldrige National Quality Award in recognition of the logic of best practice and benchmarking. The Xerox Narrative became the predicate for all subsequent treatments of managed innovation. By 1994, it could be said that:

The simple fact is that benchmarking has been driven substantially by the Malcolm Baldrige National Quality Award, in which the need for benchmarking

is a stated requirement in the application criteria. While the benchmarking requirement is specifically called for in Section 2.2 of the Baldrige award, the need for benchmarking is mentioned more than 200 times in the 1994 criteria. Directly or indirectly, benchmarking affects up to 50 percent of the award's scoring and, therefore, is seen as a critical quality tool (Camp, 1995).

The Xerox Narrative and the logic of best practice and benchmarking as institutionalized in Baldrige Award criteria have formed a hegemony reinforced by subsequent Baldrige winners — Ford, Alcoa, AT&T, DuPont, IBM, Johnson & Johnson, Kodak, and Texas Instruments. By the mid-1990s, the Xerox Narrative and the devices of best practice and benchmarking had become part of the belief system of management consultants, conferences, symposia, and training courses.

In the corporate world, the widely shared belief that innovation, like all else, is manageable has become part of the general business ethos. Benchmarking, best practice, and the pressure of rankings and prizes inform that ethos and its attendant narratives and myths. And, as always, that general business ethos has migrated to the public and nonprofit sector. This essay is an exploration of the colonization of the public and nonprofit sectors by the modern business ethos. It describes how attempts to manage innovation have combined with organizational rankings and other expressions of status, to enter an iron cage, which reduces rather than enhances the prospects for institutional innovation.

The Xerox Narrative seemed to confirm Moore's Law, the observation that the density of silicon chips closely follows a curve, showing that the information storable on a given amount of silicon has roughly doubled every year since the technology was invented. This relation, first expressed in 1964 by semiconductor engineer Gordon Moore (who cofounded Intel), held until the late 1970s, at which point the doubling period slowed to 18 months, which still holds. Moore's Law is apparently self-fulfilling. The implication is that, if one rests on one's laurels, somebody else, somewhere else, will build a better chip. The more important implications of Moore's Law are, first, the assumption that corporate innovation is orderly and patterned, and second, that it can be managed.

Innovation and creativity are now highly organized and managed. Modern communication and computing firms are heavily invested in the research required to develop faster, smaller, and higher-capacity technology. Pharmaceutical companies have large and well-funded research and development units on the hunt for new drugs and cures. The U.S. Department of Defense heavily invests in research and development, seeking breakthroughs in weapons systems. In these and many other examples, it is now generally assumed that innovation is something that can be managed.

But history has not been kind to the master narrative. At this writing, Xerox was in an 18-month free-fall. Its stock had fallen from \$68 to \$10, with no bottom in sight. The competition — Canon and Ricoh — are gobbling up market share. Earnings are flat and the company has avoided losses only by deep downsizing — nearly 10,000 employees or about 10 percent of its workforce. If there were an Unbaldrige Award for the year 2003, Xerox would be a contender. It appears that Baldrige Awards are not revoked. Yet, even in the face of these reversals, the Xerox Narrative rolls on.

A company that uses benchmarking, identifies good work processes and customer services provided by other companies in the same or somewhat similar line of business and, to put it kindly, borrows those practices. The logic of best practices and benchmarking assumes a competitive market. Indeed, one of the better-known books on the subject, Robert Camp's *Business Process Benchmarking*, moves rhetoric beyond market competition to warfare, heading each chapter with a quote from Sun Tzu, as in Chapter 2: What to Benchmark — “He who knows when he can fight and when he cannot will be victorious,” and Chapter 3: Whom to Benchmark — “If you know your enemy and know yourself, your victory will not stand in doubt.”

As presented in the accounts that flow from the master narrative, the benchmarking process is as follows:

- Decide what to benchmark by identifying the key internal business work processes with the largest opportunity for improvement.
- Determine which other companies use superior work processes and investigate them.

- Determine the performance gap between superior work processes identified elsewhere.
- Build acceptance for the need to improve work processes so as to close the gap.
- Set the benchmarks of performance that will close the gap, including, when possible, precise measurement of performance. Implement the benchmark and performance plan.
- Update benchmarks to stay current and competitive.

Although it is dressed up in a new language, benchmarking is little more than the practice of hard-nosed business competition, tighter internal controls, and a keener knowledge of a changing world.

We should applaud benchmarking and other improved business practices because they attempt to improve the products we buy, the services we receive, and the value of our stock portfolios. There may even be some reason to think that setting up the Baldrige Award office in the U.S. Department of Commerce is useful. The award is given annually to businesses, government agencies (in the federal government it is called the President's Award for Quality) and nonprofit organizations. From 1988 to 1997, publicly traded Baldrige Award winners outperformed the market by 2.7 to 1. Since then, however, they have not done as well and, at this writing, were trading at about the same as the S&P 500. Baldrige Award applicants are evaluated on a 1,000-point scale that is loaded up with benchmarking, strategic planning logic, and other techniques for presuming to manage innovation. There are 110 points for leadership; 80 points for strategic planning; 80 points for customer and market focus; 80 points for information and analysis; 100 points for work systems, training, and employee satisfaction; 100 points for work process management; and 450 points for business results — the bottom line (Blazy, 1998). These criteria simply assume that leadership, customer focus, good systems of information, analysis and training, and well-developed strategic planning are clearly associated with corporate innovation and creativity. Because such criteria fairly describe well-managed companies, it is assumed the same criteria also describe innovative companies.

Innovation comes from *novus*, or new, and is understood to be the introduction of something new, a novelty, a changed way of doing things. The questions are: Does managed innova-

tion work? Does it cause or even encourage originality? Does it cause or even encourage innovation? The answer to these questions is mostly, No. The exuberant CEO and boardroom embrace of benchmarking as the key to the corporate management of innovation runs almost directly contrary to what we actually know about how innovation really works in complex organizations, the subject to which we now turn.

## II

It is our good fortune that there are three richly documented studies of corporate innovation and creativity. Unlike the Xerox Narrative, these studies reject benchmarking and strategic planning as the keys to corporate innovation and find those keys hidden in other places. The first, *Breakthroughs!* by P. Rangamath Nayak and John P. Kettingham, is a study of 14 commercial innovations so significant and lasting as to constitute breakthroughs. The authors' studies included significant innovations in products, services, and industrial processes among them — the compact disc, the VCR, the ulcer drug Tagamet®, the CAT scan, overnight airmail package service (Federal Express) and, of all things, Club Med. They represent a broad spectrum of industries, and for this reason the findings are important to those interested in innovation and creativity — not only in the private sector, but also in the public and nonprofit sectors.

Nayak and Kettingham (1994) found that breakthroughs thrive mostly in rich soil, but grow also from barren soil, rocky soil, or no soil at all. Breakthroughs come from organizations that foster creativity as well as those with poor records of innovation. They come from creative teams joined by their management, ignored by their management, supported only belatedly by their management, misunderstood by their management, or castigated by their management.

Breakthroughs are not organizational creations, although they may be helped or hindered by organizations. Once an innovation is successful, it is eagerly claimed by the organization and often by management consultants. Innovations are more like works of art than works of commerce. Teams of people who accomplish breakthroughs behave more like

disciples emerging from the tutelage of a great visionary than like the graduates of a prestigious management curriculum.

The authors' findings contradicted their own preconceptions about what makes for truly exceptional businesses. They first found one common element among those businesses: "a core ideology — core values and a sense of purpose beyond just making money."

James C. Collins and Jerry L. Porras in *Built to Last: Successful Habits of Visionary Companies*, is a report on a 6-year study of 18 companies including IBM, Sony, Wal-Mart and Walt Disney.

What did these companies have in common? Charismatic and visionary leadership? Complex strategic planning processes? Brilliant or elegant mission statements? An overriding commitment to maximizing profits? None of the above.

The authors' findings contradicted their own preconceptions about what makes for truly exceptional businesses. They first found one common element among those businesses, "a core ideology — core values and a sense of purpose beyond just making money."

Collins and Porras are quick to point out that such core values are not the same from company to company, nor are they necessarily those we would regard as enlightened or humanistic — although sometimes they are. Nevertheless, great and enduring business institutions are organized around a "cultlike" devotion to a set of values that gives those companies their reason for being and command the dedication of their employees.

Adherence to core values does not mean that visionary companies resist change or are strategically conservative. Indeed, such companies use what the authors call "big hairy audacious goals" to stay energized and challenged.

The best moves in advancing those goals do not necessarily follow from systematic planning. Visionary companies accommodate lots of inefficiency in the pursuit of excellence. They encourage experimentation, seize opportunities even when they don't fit into the strategic plan, and take advantage of the accidental. What, in retrospect, looks like the result of pre-scient planning is sometimes the happy outcome of a policy of, as the authors put it, "Let's just try a lot of stuff and keep what works."

Clayton M. Christensen's recent *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail* studies innovation in leading and surviving companies. His first and most provocative finding is that customer responsiveness does not lead to innovation because customers seldom know or can even imagine what they do not now need. The Internet and e-commerce had its origins in the minds of information technology and computer folks who imagined Internet capability and only worried later about how it might be made commercially viable. While the power of the Internet is everywhere evident, its commercial applications are being sorted out — some firms survive and flourish; others do not.

Christensen calls their innovations a form of “disruptive technology.” Disruptive technology is often rejected by an institutional culture built on decades of profitability based on known and understood markets. Therefore, as a company manages better and better to respond to those markets it is less and less able to invest in risky and unproven disruptive technology. Christensen found successful organizations tend to be particularly intolerant of failure and risk born out of experimentation, which can lead to failure. Finally, Christensen gives the edge to smaller emergent firms because they are responding to inchoate emerging markets, something that seldom makes sense to established corporate leaders.

Despite their endowments in technology, brand names, manufacturing prowess, management experience, distribution muscle, and just plain cash, successful companies populated by good managers have a genuinely hard time doing what does not fit their model for how to make money. Because disruptive technologies rarely make sense during the years when investing in them is most important, conventional managerial wisdom at established firms constitutes an entry and mobility barrier that small-firm entrepreneurs and investors can bank on. It is powerful and pervasive.

Easily, the best study of innovation in the public and non-profit sector is Paul C. Light's *Sustaining Innovation: Creating Nonprofit and Government Organizations that Innovate Naturally*. He defines innovation in public and nonprofit settings as “an act that challenges the prevailing wisdom as it creates public values.” After a 5-year search, 26 Minnesota

institutions were culled from a much longer list based on a history of sustained innovation. Each institution was studied in depth, including site visits and interviews. The sample included 8 governmental and 18 nonprofit institutions ranging from large state agencies, to tiny nonprofits, to county governments, from new agencies to well-established ones, and from the well-endowed to the grant dependent.

Many of these organizations got away with a staggering amount of confusion with regard to chains of command. Although there are preferred ways to organize internal structure, absolute clarity about just who is the boss is not one of the prescriptions for successful innovation. In fact, these groups achieved relatively low returns from structural reforms. While the conventional wisdom is that organizations should reorganize to flatten their hierarchies, rewire their organization charts and generally make things fit, most of the organizations in the study did quite well without such restructuring.

Leaders of the 26 organizations were not naïve about human nature. Most had stories of employees gone wrong and ideas lost to jealousy and bitterness. But they chose to see the positive potential in each employee rather than the negative. They designed for innovation, even as they made sure their organizations had the systems to limit damage. Paul Light calls it a rigorous optimism. They created that optimism by adopting practices that helped them achieve innovative leadership.

Innovating organizations need not be perfect. An innovating organization can do many things imperfectly and still succeed. It can get by with too many job classifications or too few; it can survive with a state-of-the-art computer system or none at all. What it cannot survive is poor financial management systems and a lack of concern for outcomes. The former is essential both for preventing financial disaster and for creating room for investment, and the latter is an absolute requirement for changing from compliance-based to performance-based accountability.

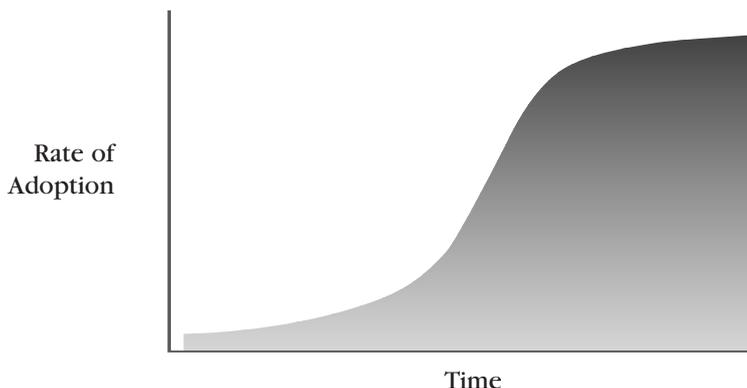
Light found that there are no gimmicks or shortcuts to innovation. Instead, real innovative institutions get there by practicing good old-fashioned democratic organization and management. His conclusions, like those of the previously cited researchers, make no mention of best practices, bench-

marking, or other parts of the logic of managed innovation. Continuing success in innovation is based on four core values that sound like a combined Sunday school lesson and management consultant's manual. Light found that the core values shared by surviving innovative institutions were trust in employees, clients, and partners; honesty about the organization's mission and about who makes decisions; rigor in internal management and outcome evaluation; and, faith in the purposes of the organization and its possibilities. According to Light, innovation is more likely a result of good democratic administration than a result of attempts to directly manage it. (See also Frederickson and Johnston, 1999.)

### III

The empirically reliable knowledge of institutional innovation and creativity is complemented by a robust understanding of the diffusion of innovation. Best practice and benchmarking are not so much the management of innovation as they are the management of the diffusion of innovation.

Doubtless, the ultimate study is Everett M. Rogers' *Diffusion of Innovations*, published in 1962, 1971, and 1983. Innovation diffusion exhibits a common pattern — an S-curve (as seen in the diagram below). At first the adoption of innovation is slow, with experimentation, trial and error, and the challenges of being the guinea pig. Once a few others successfully adopt an innovation, there tends to be a steep climb in adoption, followed by a leveling off. When an innovation reaches the leveling-off stage, investments in seeking additional adopters are usually wasted.



Original innovations and their diffusion are often responses to crises. In the absence of crises, the risks associated with innovations appear to outweigh the benefits. Innovative ideas that are relatively easy to describe, characterize, and build into narratives are more readily accepted.

Original innovators and early adopters reap the greatest rewards and have the highest status. They tend to be younger and come from newer institutions. Innovators inside institutions are perceived by others and by themselves as deviant. Most professionals may not be innovators but most are carriers, as anyone who has attended a professional conference can attest. Consultants are also innovation carriers who interact with top management and are part of the attempts to “cause” innovation from the top down, innovations that seldom last.

#### IV

One would assume that the good empirical research on innovation in private, public, and nonprofit institutions as well as the considerable scholarly research on the diffusion of innovation would have informed the best practice/benchmarking and other managed innovation literature. Wrong. Understandings of corporate innovation in the business ethos are based on the repetition of stories such as the Xerox Narrative, uncritical analysis, short-run conclusions resting on thin data, and myth. Written by management consultants, business gurus, and ghost writers for CEOs, the best practice/benchmarking literature has the mesmerizing influence of enthusiasm for perceived successful innovations and the beguiling quality that jargon and salesmanship can bring to a subject. The lack of detailed knowledge of presumed successful innovations and their verification has not held back best practice/benchmarking momentum.

All three of the cited studies of business innovation and the 1999 Paul Light study of government and nonprofit innovation, as well as many studies of the diffusion of organizational innovation, form a pattern. It reveals that innovation can occur in organized and managed settings but that there is a disconnect between the logic of managed innovation and,

as Light puts it, sustained innovation (1998). Both approaches share an emphasis on the organization's mission, on encouraging some risk taking and forgiving failed risks, on outcomes and performance, on a call for ideas or a marketplace of ideas, and on customers or clients.

The disconnect lies in the emphasis on traditional strategic planning, on the formalized importing of the innovations of others, on goal clarification, on order, on the visions of leaders, and on the advice of consultants found in the logic of managed innovation.

The evidence for sustained innovation, on the other hand, describes relatively untidy, even chaotic work environments. Such creative places avoid an overemphasis on planning, and emphasize doing and experimenting. While there is an emphasis on performance, it falls far short of a preoccupation with measuring it. Such creative places practice high decentralization, employee ownership of the processes of work, and leadership that is at once decisive and highly democratic and communicative. Put simply, one approach assumes *innovation* can be managed, and the other assumes the *organization* must be managed, and if it is rightly managed it will innovate.

## V

The staying power of the Xerox Narrative is associated with something new, itself an innovation. Xerox did not plan or anticipate it, but management consultants selling best practice/benchmarking certainly capitalized on it. This innovation was gaining momentum in the early 1990s, climbing rapidly up the S-curve at about the time the Xerox Narrative was taking form. Its name: ranking.

The ranking business has soared in the past ten years. A week does not pass without a newspaper headline about the most dangerous cities, the best HMOs, the worst airline on-time records, the best companies to work for, and which local schools are best and worst. Magazine covers hold out the promise that one will learn which are the best colleges, best graduate schools, best retirement places, safest cities, and other forms of "news you can use." Our work world is filled

with measures of performance, outcomes, results, quality, and descriptions of best, as distinct from worst, practices. Using the logic of organizational performance evaluations, school districts are being discredited, deans are being replaced because their schools slipped in ranking, and mayors, hospital administrators, university presidents, and school superintendents are crowing about being number one or in the top ten. Or they are on the defensive, describing what must be done in the future to improve rankings. Rankings are the fashion of the day. They are everywhere and they are powerful. It seems that everything can be ranked.

Consumer products have been ranked for a long time. Cars, refrigerators, toothpaste, and bread makers are tangible, easily tested, and easy to compare. Consumer product measurement is nuanced, providing not just an overall ranking but measures of reliability, appearance, convenience, safety, and cost. Who among us does not delight in a best buy? In response to corporate power and advertising, the Consumers Union was born in 1936 and is still the best known and most respected consumer product evaluation outfit. There is no question that firms producing tangible consumer products are influenced by the evaluation and ranking of their products, inasmuch as their success in the marketplace may turn on how the ranking of their products influence consumer choices.

Awards, prizes, and honors are a form of measurement, evaluation, and comparison, based on the assumption that such awards recognize the best of something. The Nobel Prizes are loose forms of measurement, setting the bar for standards and accomplishment. Consider the Pulitzer Prize in journalism, the Academy Awards in motion pictures, the Newbery Award for children's literature, the National League of Cities All America Cities Awards, the Harvard Government Innovation Awards, and the MacArthur "genius" grants. In virtually every field of human endeavor there is now at least one prize, award, or honor. These awards are very popular and often influential to those who compete (usually rather subtly) for them. James D. Watson's wonderful little book *The Double Helix* is at one level the first description of DNA, but at another level it is a description of a race with other scientists, particularly Linus Pauling, to get there first, knowing a Nobel Prize was waiting.

Systems of accreditation are used in virtually every professional field as measures of organizational adherence, or failure to measure up to agreed-upon standards. Certification and licensing serve as measures of individual competence in these same fields. If a hospital or a school is disaccredited or an individual is disbarred or has his or her license revoked, it is a measure of having fallen below standards of acceptable professional competence.

Tests are the coin of the realm for those who seek university and graduate school admissions and those who seek to enter the professions. Testing, including test preparation, is now a very big business. Tests are not only individual, they are aggregated to determine the assumed quality of an entire public school or public school district, or the quality of an entering law or business school class compared to the aggregate test scores of those admitted to other law or business schools. Universities and testing organizations are giant gatekeepers not just to the professions, but as arbiters of a student's future prospects (Lehman, 1999; Sacks, 2000).

We now expect the measurement of results in virtually all fields of human endeavor, including such difficult-to-predict fields as scientific research and development — as in measuring the progress toward the cure for AIDS and in philanthropy — as in setting goals for grants and measuring progress toward goals resulting from grants.

Rankings and performance evaluations are made by many organizations. Universities measure student performance. Professional associations such as the Bar Association, the American Medical Association, and many others act as self-governing peers, setting the standards and keeping the gates, as well as participating in, and controlling, accreditation. Many government agencies, particularly at the state and federal levels, control licensing and certification of both individuals and organizations in a form of regulatory control and oversight, often with publicly appointed boards and commissions. In 1993, the federal government passed, and President Clinton signed, the Government Results and Performance Act (GRPA), which requires all agencies to develop long-range strategic plans and specific goals, as well as descriptions of the measures to be used to gauge progress toward those goals.

Many nonprofit organizations practice rankings and performance evaluations, including the professional associations described above. The Consumers Union and the Underwriters Laboratories (U.L.) are nonprofits. Philanthropies are also nonprofits and appear to be moving in the direction of funding organizations that do rankings and performance evaluations and ask them to measure the results. Virtually all government grants now require outcome evaluations. The rankings or report cards on the quality of state and large-city administration, conducted by the Maxwell School of Citizenship and Public Affairs of Syracuse University and published in the privately held *Governing Magazine*, are funded by the Pew Charitable Trusts. The Harvard Government Innovation Awards, funded by the Ford Foundation, are published annually and distributed broadly in an effort to encourage a diffusion of innovation. The largest testing organizations, the Educational Testing Service, and the field or profession-specific testing organizations, are almost all nonprofits. But the testing and educational preparation organizations such as Kaplan and Sylvan are profit-making (they hope) corporations.

Private firms have moved vigorously into the ranking and performance evaluation business, particularly the magazine companies. The regular *U.S. News & World Report* and *Newsweek* rankings in higher education have swamped the *Gorman Report* in the way that large firms swamp small companies. Many private rankers are freelance cottage industries such as Gorman. In the public or governmental world, the best known of these is the Morgan-Quintno Company in Lawrence, Kansas, which publishes annual rankings of virtually every bit of available data comparing the 50 states and the American cities. They also rank cities by safety, civility, and so forth. Libraries are Morgan-Quintno's largest customers because their publications are annual and have many of the characteristics of reference books and almanacs.

Interest groups (which are usually nonprofit organizations) also engage in rankings. The Sierra Club ranks legislators one way, while forest, mineral, and hunting groups rank them another. The pro- and anti-gun control groups engage in the war of the rankings. Less becoming is the evidence that the commercial package carriers (UPS, FedEx, etc.) do annual and

always negative evaluations of the U.S. Postal Service, and they use these evaluations on Capitol Hill to attempt to diminish the credibility of the U.S. Postal Service. The commercially owned electric utilities use rankings and performance evaluations to discredit the publicly owned electric utilities, always urging that they be privatized, which would, of course, make them easy prey for large commercial electric utilities. There is even some evidence that the relatively new private education industry has, at times, been associated with negative evaluations of public schools, and particularly inner-city public schools.

Rankings and performance evaluation are done in a variety of ways. For tangible consumer products, there are actual laboratory tests. For measures of individual scholarly development, there are general and field-specific tests at virtually all levels of schooling. The tests themselves are regularly evaluated and critiqued. In recent years, testing has been moving from pen-and-paper technology to Internet platforms. The rankings of entire schools or programs are simply aggregates of the individual scores of students or applicants, compared with previous years and with other schools or programs. States gather those data and set benchmarks; schools that fall below benchmarks may be discredited. Based on aggregate test scores in successive years, measured against a preestablished and fixed minimum standard (a benchmark) in the summer of 2000, the state of Missouri put the schools of both Kansas City and St. Louis on probation and threatened them with discreditation. That Kansas City and St. Louis are the two large inner-city, primarily minority school districts in Missouri — the only ones discredited — has resulted in second thoughts among the legislators who wrote the law and the public servants who must implement it. Second thoughts appear to take this form: At the time it seemed to be a good idea. Now what do we do? Doubtless, Missouri will need to fix the ranking scheme as New York State fixed its teacher accreditation test (more than half failed) and Texas fixed its student graduation tests program.

The staple of many ranking systems is the opinion questionnaire or survey of citizen, client, or customer satisfaction. *U.S. News & World Report* and *Newsweek* use surveys in their college and university rankings, as do organizations that rank air-

travel satisfaction or government-service satisfaction.

Finally, in the more subjective fields of rankings, forms of peer review are widely used. Awards, prizes, and honors almost always use peer evaluations. The choices of articles to be accepted or rejected for publication in leading scholarly journals, a form of ranking, are always made by qualified peers. Tenure may be in the balance for the young professor. The choice of government grant recipients is always peer reviewed, endowing peers in each discipline with the power to measure quality.

Assumptions made on the part of those who do the rankings and give the awards determine the results. In much the same way that the audience must accept the straight line in order for the joke to be funny, the salience of rankings rest on the assumptions made by the rankers. Ranking assumptions are, of course, the issue with respect to the alleged cultural bias of student testing as well as the issue in virtually all criticism of rankings. To deal with challenges to assumptions, ranking organizations use prestigious universities to set the assumptions and make the rankings or choose the winners.

For cities and states, the Maxwell School used the following good-management assumptions:

- Financial management systems that include multiyear perspective on budgeting;
- Mechanisms that preserve stability and fiscal health;
- Sufficient financial information available to policymakers, managers, and citizens, and appropriate control over financial operations;
- Human resource systems that emphasize strategic analysis, attracting needed employees, maintaining a skilled work force and a civil service structure;
- Information technology systems stressing a coherent architecture for information technology systems, meaningful, multiyear information technology planning, and information technology training;
- Capital management systems including a thorough analysis of future needs, monitoring, and evaluation of capital projects, and the maintenance of capital assets; and
- Managing for results including strategic planning, data that can measure progress, use of results data by managers for policy-making and management, and clear communication

with stakeholders.

Using these criteria, jurisdictions are graded in each category and overall. An editorial in the *Kansas City Star* (2000) considered the grade given Kansas City, Missouri:

Kansas Citians shouldn't be too pleased by the recent grade handed out to City Hall for the quality of its governmental operations. Kansas City received an overall grade of B-, placing it right in the middle of the pack of 35 cities that were exhaustively evaluated by the Maxwell School of Citizenship and Public Affairs at Syracuse University and *Governing Magazine*. The overall grade for the city was disappointing in this respect: The veteran management team led by City Manager Bob Collins should have done a better job handling longstanding problems in the areas outlined by the survey. The flip side of this argument is that Collins has the opportunity to put in place new, change-oriented managers as department directors leave. The most recent departure was of Fire Chief Rick Brisbin. Staying on a positive note, Mayor Kay Barnes, Collins, and other city officials recently have highlighted ways to improve the delivery of basic services and how public employees are trained. If these programs are carried out, Kansas Citians will be getting more efficient use of their tax dollars. And local government's next report card — scheduled to be delivered in three years — could contain a higher grade.

No doubt other cities (as well as states and federal agencies) that received mediocre grades responded in the same way. Clearly, the rankings put pressure on jurisdictions to improve their management, or at least get their management to conform to the assumptions made by the rankings.

In recent years, Kansas City has purposely chosen to invest more in neighborhood community services. Given limited financial resources, will the relatively poor rankings of the management of Kansas City put pressure on its leaders to invest more in information technology and other management needs and less in community services? Better management is the agenda of the funders of this ranking program, apparent-

ly without regard to the fact that better management has tradeoff costs, or that management must compete with other government needs.

The assumptions of rankers are seldom challenged. Nevertheless, some ranking assumptions are certainly open to criticism — as, for example, the continuing claim that research is more important than teaching in giving faculty tenure, or that student test scores are more important than grade point averages for university and professional school admissions. Ranking assumptions can be the battleground for policy debates. But more often, rankings and the assumptions behind them are simply assumed to be valid.

It is also abundantly evident in the response of the *Kansas City Star* to the report card the city received, that such matters are highly political. Those who were not supportive of the incumbent city manager used these rankings as ammunition, and he has been replaced. If rankings do not improve, whoever challenges the mayor in the next election will doubtless use the report card grades as an argument for why the city needs a new mayor. Mayors, governors, deans, university presidents, and many others subject to the whims of rankers find themselves vulnerable and, in many cases, without the resources to make things better. Claims of objectivity notwithstanding, all rankings are based on assumptions, and no rankings are neutral. Ranking can be perilous to careers and to public perceptions of institutions.

James Fellows, the editor of *U.S. News & World Report*, from 1995 to 1998, is one of the strongest national advocates of the so-called civic journalism and “news you can use.” He argues that rankings are just such news. And as anyone who has walked by a magazine rack in the last five years would conclude, ranking is very good journalism business. Fellows said that their annual issue that ranks schools and colleges is, to *U.S. News & World Report*, what the annual swimsuit issue is to *Sports Illustrated*.

We do not know, and may never know, what effect ranking has on quality however defined. Such things are tough to measure.

William Gormley and David Weimer’s *Organizational Report Cards* (1999) is the most thorough consideration of those aspects of the measurement business that are distinctly organizational. They review organizations that do perfor-

mance measurement in child care, air travel, higher education, economic development, employment and training, hospitals, HMOs, insurance companies, nursing homes, and public schools. They are generally positive in their evaluation of the ranking and performance measurement movement and conclude that “organizational report cards are here to stay. Demands for systematic performance data from policymakers, citizens, managers, and consumers will guarantee that. The content, format, and impact of report cards however, will generate considerable debate.”

Rankings and report cards are, as Gormley and Weimer (1999) rightly claim, a kind of market surrogate in a nonmarket setting. The authors also claim that more marketlike competition resulting from ranking stimulates improvement. A highly ranked city, school district, or business, logically should try to improve its ranking by improving its quality. And a well-ranked city will feel pressure from below and respond with its own improvement — assuming it has the resources. Competition of this kind has always been with us, but rankings may sharpen and focus that competition. Rankings may also result in “teaching to the test,” in a desperate attempt to lift “quality.”

It is easy, however, to miss the point: many rankings do not really rank quality; they rank impressions of quality. Numbers are not reality; they represent reality. Rankings are, in fact, surrogates for quality, used in the absence of any reliable way to measure quality. Rankings processes pick up not only reflections of quality, but reflections of status, bias, political clout, and ignorance. We have learned that rankings attract a lot of attention. We may agree that some rankings are casual social science at best, but they are exciting and fun. Rankings can be a postmodern wedding of simplified social science and entertainment. Complicated matters are distilled to simple good, better, and best categories and put into sound bites that busy folks can easily digest without being conceptually overburdened. Rankings can also be genuine objective comparisons of the effectiveness of institutions assuming there are reliable data and assuming the assumptions built in to ranking criteria are reasonable. Ranking, like most things, can be well or poorly done.

It is now commonplace for schools, cities, nonprofits, busi-

nesses, and universities to set out specific ranking objectives in their strategic plans, to marry rankings and institutional objectives. A business or law school will, for example, list among its strategic objectives, the achievement of a ranking in the top 10 or 20. Businesses will set market share and profit goals — a kind of ranking objective. The power of rankings has stimulated the logic that institutions can improve their standing by better management and particularly by attempting to directly manage the processes of organizational innovation.

Every state, every city, every university has similar stories to tell. In the ranking game, assumptions are the predicate to rankings and they are never neutral. Rankings often constitute a formidable pressure to conform to ranking assumptions. Rankings are not friendly to innovation, creativity, risk taking, or doing things in new ways. When rankings get tightly linked to efforts to further organizational innovation, as is now everywhere evident, the results are usually less, rather than more, innovation.

## VI

Ranking, comparing, and measuring constitute a search for reason and understanding that grew out of Enlightenment notions of the scientific basis of useful knowledge. Measuring is embedded in modernist assumptions of the human capacity to demystify and, thereby, to know in objective ways.

Performance measurement, best practice, benchmarking, and the influence of rankings are all broadly a part of the family of strategic planning. The strategic-planning approach to guiding organizational change is, without question, a useful organizational tool. It may be that the processes of strategic planning, as compared to the strategic plan itself, have the most lasting effect on the organization, particularly if the strategic-planning process is highly participative and mostly bottom-up (Bryson, 1995). It is generally established that a well-designed strategic-planning process can encourage interaction between departments, enhance levels of commitment to shared objectives, help make organizational goals more congruent, sort out more important goals, and challenge goal

displacement. This is especially true for organizations with relatively unambiguous purposes, little goal conflict among principals, clear control over resources and staff, and predictable outcomes. Formal strategic planning loses salience in organizational settings that lack these characteristics. For example, government purposes are richly ambiguous, and government organizations are filled with political, regional, philosophical, and other conflicts between and among elected legislators and executives. Such organizations are often bereft of bureaucratic capacity to control work processes and outcomes. Strategic planning is, therefore, difficult to apply in the world of government management.

Nevertheless, virtually all approaches to innovation in the public sector begin with the logic of strategic planning. The 1993 Government Performance and Results Act (GPRA), for example, obliges all federal agencies to develop a strategic plan that brings precision to ambiguous law, reconciles competing objectives, and sets out formulae by which goal progress will be measured. The success of the federal strategic-planning processes mandated by GPRA is mixed, but public managers are making every creative effort to carry out the law (Radin, 2000). Finally, there are indications that managers are unable to resist inclinations to have the institutional vision formulated primarily at the top and then to sell it to the rank and file through strategic planning (Bryson, 1995).

Ranking, report carding, best practices, and benchmarking are all systems of comparison, measurement, and evaluation. In the nonmarket world of public affairs — governments, nonprofits and nongovernmental institutions, universities, and philanthropies — these comparative systems are modern market surrogates. We may not buy cities or universities, or judge the United Way or the MacArthur Foundation in the same way we decide to buy a Ford. We do, however, use their services and often depend on them. It is reasonable, then, to know how to differentiate among them and make informed choices in the absence of a market. Systems of ranking, report carding, best practices, and benchmarking presume to help us do that.

## VII

Max Weber, the first and greatest student of complex organizations described such organizations as iron cages, so efficient and powerful in their capacity to control men and women that, once established, the momentum of bureaucratization is irreversible (Weber, 1952). Generations of organization theorists have verified and elaborated the iron cage hypothesis, explaining why organizational innovation is less, rather than more, likely because of best practices, benchmarking, and their links to ranking.

One would logically assume that organizational change is driven by competition and by the need for efficiency. Instead, organizational change in the iron cage simply makes organizations more similar without necessarily making them more efficient (DiMaggio and Powell, 1983). This is why. *Isomorphism*, a bit of sociological jargon, is the concept that best captures the iron cage and the process of organizational homogenization. Isomorphism explains the pattern of increasing similarity, homogeneity, and congruence between and among organizations in similar fields — all research universities, all armies, all software companies, all HMOs, and so forth. Isomorphism in fields of similar organizations leads to homogeneity among those organizations in their structures, technologies, cultures, and outputs. The adoption of civil service reform in U.S. cities illustrates this process. Early civil service reforms were related to internal government needs and strongly influenced by such city characteristics as the size of the immigrant population, political reform movements, socioeconomic composition, and city size (Tolbert and Zucken, 1983). Later reforms had less to do with urban characteristics than with broadly held assumptions about the legitimate structural form for municipal administration. This is because cities have become increasingly alike. As they copy and mimic one another, they cease to innovate in significant ways and enter Weber's iron cage. The same isomorphic process is found in virtually all organizational fields, largely driven by three forces: coercion, mimicry, and contextual norms.

Coercive isomorphism is caused by pressure exerted on an organization by another organization on which the first is

dependent, and by the force of cultural expectations in the society within which the organization functions. In some circumstances, organizational change is a direct response to government mandate: manufacturers adopt new pollution-control technologies to conform to environmental regulations; nonprofits maintain accounts and hire accountants in order to meet tax law requirements; and, organizations employ affirmative-action officers to fend off allegations of discrimination. Schools mainstream special students and hire special education teachers, cultivate PTAs, and administrators who get along with them, and promulgate curricula that conform to state standards.

The existence of a common legal environment affects many aspects of an organization's behavior and structure. The United States has a complex, rationalized system of contract law that establishes organizational controls to honor legal commitments. Other legal and technical requirements of the state — the budget cycle, the fiscal year, annual reports, and financial reporting requirements that ensure eligibility for the receipt of federal contracts or funds — all cause organizations to behave in similar ways. A lot of organizational similarity can be traced to the forces of coercive isomorphism.

When organizational technologies are poorly understood, when goals are ambiguous or when the environment is uncertain, organizations tend to model themselves on other organizations (March and Olsen, 1976). The advantages of mimicking behavior are considerable; when an organization faces ambiguous problems or unclear solutions, it is cheaper and quicker to model other organizations with similar problems. Finding more fundamental organizational approaches is expensive and takes time (Cyert and March, 1963).

Homogeneity in organizational structures and behavior is also a function of the limited choices available. "Large organizations choose from a relatively small set of major consulting firms, which, like Johnny Appleseeds, spread a few organizational models throughout the land. Such models are powerful because structural changes are observable, whereas changes in policy and strategy are less-easily noticed. The history of management reform in American government agencies, which are noted for their goal ambiguity, is almost a textbook case of isomorphic modeling, from the Planning-Programming-

Gadgeting model of the McNamara era to the zero-based budgeting of the Carter administration,” to contracting out in the Reagan era, from government in the early Clinton years to the current search for innovation by measuring results and best practice/benchmarking (DiMaggio and Powell, 1983). Organizations tend to model themselves after similar organizations in their field that they perceive to be more legitimate or successful. Mimicking other such organizations is a practice that appears to have less to do with innovation and more to do with a search for improved reputation or standing.

A third feature of the iron cage of isomorphism is normative and stems primarily from professionalization. Two aspects of professionalization are important sources of isomorphism. One is the formal education of professionals and the other is the growth of professional networks that span organizations and act as agents for the diffusion of ideas. Universities and professional training institutions are important centers for the development of organizational norms among professional managers and their staffs. Professional and trade associations define and promulgate normative rules about organizational and professional behavior. They certify, accredit, and legitimate. Such processes fill a reservoir with rather similar individuals who occupy similar positions across a range of organizations. This similarity in professional orientation and disposition will often trump unique organizational traditions and controls (Perrow, 1974).

Many professional career tracks are so closely guarded, both at entry and throughout career progression, that individuals who make it to the top are virtually indistinguishable. There is, for example, an absence of variation among Fortune 500 board members (Hirsch and Whisler, 1982). Individuals in similar organizations are socialized to common expectations about their personal behavior, appropriate style of dress, organizational vocabularies (Williamson, 1975), and standard methods of speaking, joking, or addressing others (Ouchi, 1980). To the extent managers and key staff are drawn from the same universities and filtered through a common set of attributes, they will tend to view problems in a similar fashion, see the same policies, procedures, and structures as normatively sanctioned and legitimated, and approach decisions in much the same way. Together, they build the iron cage of isomorphism.

The exchange of information among professionals helps contribute to a commonly recognized hierarchy of status. This status-ordering occurs through both formal and informal means. The designation of a few large firms in an industry here as key bargaining agents in union-management negotiations make these central firms pivotal in other respects as well. Government recognition of key firms or organizations through the grant or contract process and by awards and prizes such as the Baldrige, gives these organizations legitimacy and visibility and leads competing firms to copy aspects of their structure or operating procedures in the hope of obtaining similar rewards. Professional and trade associations provide other arenas in which prestige organizations are recognized and their personnel given positions of substantive or ceremonial influence. The stature of managers in highly visible businesses is reinforced by representation on the boards of other organizations, particularly grant-giving philanthropies. On these boards, they often act as agents of diffusion, suggesting that nonprofit organizations use business strategies such as best practice/benchmarking. Executives and members of the boards of directors of leading nonprofits such as the Ford Foundation sit on government and foundation grant award panels, consult as government- or foundation-financed management advisers to smaller nonprofits, or sit on smaller organizations' boards, while their stature is reinforced and enlarged by the grants their nonprofits receive from government, corporate, and foundation-funding sources (DiMaggio, 1983).

Upwardly mobile managers and staff seek to secure positions in central organizations in order to further their own careers. Aspiring managers undergo socialization into the norms and mores of the organizations they hope to join. Career paths may also involve movement from entry positions in the center organizations to middle-management positions in peripheral organizations. The flow of personnel within an organizational field is further encouraged by structural homogenization — for example, the existence of common career titles and paths (such as assistant, associate, and full professor) with meanings that are commonly understood.

Organizational fields that include a large professionally trained labor force will be driven primarily by status

competition. Organizational prestige and resources are key elements in attracting professionals. This process encourages homogenization as organizations seek to ensure that they can provide the same benefits and services as their competitors.

In the iron cage, these isomorphic processes proceed in the absence of evidence that they increase organizational efficiency. (DiMaggio, 1983; DiMaggio and Powell, 1983). To the extent that organizational effectiveness is enhanced, the reason is often that organizations are rewarded for their similarity to other organizations in their fields. This similarity can make it easier for organizations to transact with other organizations, to attract career-minded staff, to be acknowledged as legitimate and reputable, and to fit into categories that define eligibility for public and private grants and contracts. None of this, however, ensures that conformist organizations are more efficient than their more deviant peers. And worse, there is no evidence that isomorphic processes enhance prospects for organizational innovation and creativity; indeed, logic suggests the opposite because isomorphic processes punish deviation.

Here are some key hypotheses drawn from the research on institutional isomorphism:

- The more uncertain the relationship between means and ends, the greater the extent to which organizations will model themselves after organizations perceived to be successful.
- The more ambiguous the goals of an organization, the greater the extent to which it will model itself after organizations that it perceives are successful. There are two reasons for this. First, organizations with ambiguous and disputed goals are likely to be dependent on appearances for legitimacy. Such organizations may find it to their advantage to meet the expectations of important constituencies about how they should be designed and run. In most situations, reliance on established, legitimated procedures enhances organizational legitimacy and survival characteristics. This explains why one prestigious research university looks almost exactly like all the others. A second reason is found in situations where conflict over organizational goals is repressed in the interest of harmony; thus participants find it easier to mimic other

organizations than to make decisions on the basis of systematic analysis of goals since such analysis is painful, disruptive, time consuming, and expensive.

- The greater the reliance on academic credentials in choosing managerial and staff personnel, the greater the extent to which an organization will become like other organizations in its field.
- The greater the participation of organizational managers in trade and professional associations, the more likely the organization will be, or will become, like other organizations in its field. This hypothesis is parallel to the institutional view that the more elaborate the networks among organizations and their members, the more likely organizations in that environment will come to resemble each other.
- The greater the extent to which an organizational field is dependent on a single (or several similar) source(s) of support for vital resources, the higher the level of isomorphism.
- The greater the extent to which the organizations in a field transact with agencies of the state, the greater the extent of isomorphism in the field as a whole. The current popularity of governmental contracting and outsourcing is, therefore, likely contributing to isomorphism.

Prizes, rankings, and test scores are representations of status and legitimacy — bars in the iron cage of isomorphism. There is every possibility that the impetus to best practices and benchmarking is driven as much by the desire to increase an institution's reputational capital as it is by a genuine desire to innovate. The iron cage of isomorphism explains the cycles of management fads and the attraction of executives in prestige organizations to consultants who claim a leading-edge innovation, a new concept, or an answer to a vexing question. The paradox of the iron cage is this: in their pursuit of breakthroughs, creativity, and innovations, organizations usually converge into an essentially indistinguishable lump. In this lump, they compete for the slightest differentiation — which is ranked first, second, third — when in fact the distinctions between them are very slight, if not without difference at all. Isomorphism is actually furthered by the combination of managed innovation and ranking, a perverse result exactly

the opposite of the reported purposes of benchmarking and best practice.

This critique of best practice/benchmarking and managed innovation, may help readers gain perspective and reconsider the logic, strength, and usefulness of these concepts. But criticism is easy. If not managed innovation by best practice and benchmarks, then what? If not the iron cage, then what? Are there better ideas for achieving institutional innovation and, if so, is there any evidence that such ideas work better than best practice/benchmarking? The answer is yes.

## VIII

For purposes of comparison, consider the description of best practice/benchmarking to be the *managed innovation* model.

Next, consider research findings on innovation and the diffusion of innovation as the *sustaining innovation* model (Light, 1998).

How do these models differ with regard to organization and reorganization, leadership, responsibility, goals, criteria for performance, norms, and values, the propensity to take risks and, most importantly, the propensity to innovate? The table that appears on the next page presents a simplified comparison of the two models. The comparative emphasis has partly to do with differences between the use of the logic of managed innovation in business and corporate settings and the logic of sustaining innovation found in business as well as governmental, nonprofit, and philanthropic settings.

**Organization and Reorganization.** Virtually all descriptions of organizing or reorganizing for best practice/benchmarking are expressed in terms of systems of management, control, and delegated authority. There is little evidence of the application of managed-innovation logic resulting in decentralization and greater work-group autonomy. Innovation research, however, indicates that innovations are more likely to occur in flexible, loosely coupled institutional (or noninstitutional) settings.

Research on the diffusion of innovation indicates that there is a wide range of organizational strategies that work —

## Comparing Managed Innovation and Sustaining Innovation Models

	Managed Innovation Model	Sustained Innovation Model
<b>Organization:</b>	Designed, orderly, hierarchical	Loosely coupled, untidy
<b>Reorganization:</b>	Designed and planned	The formalization of developments that are already taking place
<b>A Leader is:</b>	The architect The change agent The visionary	The “gardener” — supports rather than directs
<b>Responsibility:</b>	For results To Stockholders To customer, citizens, clients	To the group and in the group To the institution and in the institution To external stakeholders, present and future
<b>Planning:</b>	Traditional strategic management and goal-setting and measuring results against those goals.	Innovating in a generally agreed-upon direction.
<b>Goals:</b>	Seek clarity, prioritize, focus on core mission. Set goals, then implement them.  Linear	Some purposeful ambiguity  Solutions search for problems just as often as problems search for solutions.  Questions search for answers, answers search for questions.  Goals lead to action, action leads to goals.  Curvilinear
<b>Criteria for Performance:</b>		
<b>Time</b>	Shorter	Longer
<b>Precision</b>	Greater	Lesser
<b>Results</b>	Measuring efficiency, the bottom line	Instrumental effectiveness
<b>Outcomes</b>	Measured precisely	Estimated, debated, A garbage can
<b>Process</b>	Means-ends chain	
<b>Values:</b>	Managed, guided, vision	Historical, emergent, natural, shared
<b>Propensity to risk:</b>	Lesser	Greater
<b>Propensity to innovate:</b>	Short, rapid, shallow cycles of innovation	Long, slow, deep cycles of innovations

managed innovation, model laws and charters, policy entrepreneurs, advocacy groups and, in modern times, rankings. But all tend toward isomorphism.

There are differing perspectives on reorganization. Managed-innovation enthusiasts are likely to describe, advocate, and carry out reorganization to attempt to achieve either innovation or the diffusion of innovation. Research on institutional behavior describes successful and enduring reorganization as an iterative process of organizational adaptation only thought to be a reorganization after it is well along. Sustaining innovation assumes that form follows function.

**Leadership.** The Xerox Narrative, as well as its spawn, includes heroic models of the great leader, the architect of survival, the agent of change, the visionary. All the research on the diffusion of innovation describes the linkage between leaders or elites, management consultants, policy advocates, and interest groups. Managed innovation simply assumes strong, centralized, heroic leadership. It is often the case that particularly creative persons with breakthrough ideas start organizations — Steve Jobs at Apple, Bill Gates at Microsoft — and become the heroic leaders in organization narratives. The problem, of course, is sustaining momentum, not only in the production of the original product, but the development of new innovative products. Organizations in the public and nonprofit sectors almost always receive a new leader into an ongoing institutional culture. While new businesses result from the creative energy of a single leader, the Department of Agriculture, the state of Kansas, or the Ford Foundation are not phased out to be replaced by new organizations led by creative leaders. So, as we say in politics and public management, “ya dance with who brung ya.”

Research on sustaining innovation, on the other hand, describes leadership thus:

The institution is a political and moral order, a collection of long-lasting standard operating procedures — reflecting values, principles, and beliefs that are shared. The primary task of the leader is to guarantee enough order and autonomy to enable the pursuit of collective purposes. Leaders are “gardeners” — they support rather than direct. They are obligated

to defend uniform and collective standards of appropriateness, with reference to what is best for the institution.

— Johan P. Olsen

The metaphor of the leader as gardener simply would not do in the corporate world of managed innovation. As the logic of managed innovation migrates to governments, nonprofits, and philanthropies, assumptions about leadership take on greater importance. Nonbusiness institutions are inclined to multiple and contrasting leaders, and loose power structures that dissipate centralized power, rendering heroic leadership assumptions of little use. It should not be imagined, however, that there are no great leaders in government and nonprofit institutions, because there are (Cooper and Wright, 1992). However, these great leaders, as the best research on public leadership shows, are more like gardeners than high-profile visionary heroes.

Smaller and newer nonprofits such as community development corporations, particularly those that survive and flourish, are a bit like smaller aggressive corporations and are a major exception to this generalization. They are often created and led by high-energy, strong, heroic-leader types. To endure, over time, they must routinize charisma, as Max Weber noted more than 50 years ago.

**Responsibility and Goals.** No phrase better describes the modern conception of responsibility than Al Davis' now famous "just win baby." It is the bottom line that matters most. In our time, the bottom line is summed every quarter, every month, every day. Searching elsewhere for best practices and applying them through benchmarking is a bit like catching the next good wave in a patterned response to short-run performance pressures.

Diffusion research shows that institutional participants resist purported innovations they have not participated in creating. And institutions will ultimately reject purported innovations that run counter to institutional culture and norms. There is, in the sustained innovations model, a deep commitment to institutional performance, but that performance grows out of a shared sense of responsibility in the work-group, from the

institution to the individual and from the individual to the institution. As W. Edwards Deming (1988) carefully explained in his total quality management (TQM) theory, high performance flows from shared work-group responsibility for performance. Innovation happens in a steady iterative trial-and-error fashion, but it does happen. Once such innovation occurs, it becomes a collectively owned and enduring part of the institution.

Traditional strategic-planning logic — clarifying goals, choosing the most important among them, settling on one or two goals that best represent the institution's core purpose, setting up measures of how the institution will know how well it is achieving these goals, looking at other institutions to find their best practices and importing the best, and using benchmarks — is at the heart of the managed-innovation model. It is an organizational adaptation of stimulus-response, means and ends, and logical positivist rationality. Based on empirical observation, scholars and researchers long ago modified and softened this model with the logic of “muddling through” and buffered rationality (Lindbloom, 1959). The common, empirically tested model is described in formal terms as “successive limited comparisons” and is compared with the “rational-comprehensive model,” which now has virtually no empirical warrant. The best of modern strategic-planning literature reads very much like successive limited comparisons and the sustaining-innovations model (Bryson, 1995).

Managed-innovation enthusiasts and salespersons have mostly ignored such dull academic findings and press on with the rhetoric of more or less pure rational-comprehensive strategic planning.

Two ghosts rest in the logic of strategic planning and managed innovation. First is the assumption that goals are knowable in the existential sense. On the one hand, leadership theorists see goals emerging in the institutional visions of leaders, but the ghost of megalomania haunts institutional purposes (not to mention the public interest) with personal ambition and hubris. On the other hand, TQM theorists find goals in customer opinions and preferences. The original-customer focus-group ideas in TQM were designed to deal with that ghost, but TQM assumes that customers *already* know and understand their needs and interests and have thought through

all the possibilities. Development of the Internet and e-mail, for example, are not so much responses to customer suggestions for product improvement as they are innovations that, when presented to possible customers, were in time found to be useful. It was not customer response or strategic planning or management that resulted in these innovations; it was creative people in flexible institutional settings seeing the possibilities. Innovation often takes the form of discovering answers for as yet unasked questions, services for which there is not yet a demand, solutions searching for problems, arrows seeking bull's-eyes, and bull's-eyes waiting for arrows. Collective institutional deliberation and agreement on sensible organizational *actions* are more likely to put goals into perspective and make them useful than are formal strategic-planning, goal-clarification exercises. Goals that are too precise may be the enemy of innovation, thus turning the logic of managed innovation on its head.

Research also indicates that a certain level of goal ambiguity leaves wiggle room, invites possibilities, and generally opens things up. Techniques for the effective management and navigation of ambiguity are part of virtually all management textbooks and courses. It is useful to reflect on “Frederickson’s rule”: Goals are deceptive — the unaimed arrow never misses.

**Criteria for Performance.** How can institutions know how well they are doing? At one level this is a temporal question — how well are we doing today, today compared to yesterday, today compared to a year ago, or today projected a month, a year, a decade ahead? The managed-innovation model is impatient and the whole logic of best practice and benchmarking is in a hurry. For several years, the Xerox Narrative was confirmed by events. If written today, a Xerox Narrative would come to the opposite conclusion. Sustaining-innovation assumptions are, it must be admitted, luxurious with time.

Deliberation takes time. Finding consensus takes time. Training takes time. Gardening takes time. Nevertheless, institutions that practice widespread open deliberation, moving only after genuine consensus, investing in their people through training and, particularly, through job security, are more likely to nourish innovation as are institutions that presume to manage innovations. And innovations, thus nourished, are much more likely to last.

Performance assumptions, in managed innovation, turn on the bottom line as Baldrige Award criteria attest. In the business world, the bottom line is not only profit and/or stock value; it is also status and legitimacy. In the government and nonprofit world, the criteria for performance are efficiency (accomplishing the greatest possible results for the dollars available), economy (accomplishing some fixed task, such as janitorial services, for the least money, which explains the love affair with contracting out) and legitimacy. Profit, stock value, efficiency, economy, and external legitimacy all ignore the instrumental benefits and favorable effects on performance of widespread participation in institutional decision making, two-way loyalty between institutions and their employees, and a deeply shared set of agreed-upon values and commitments. Although such instrumental values have always been hard to measure, it is agreed among institutional scholars, as well as innovation scholars, that instrumental values define organizational culture and broadly influence all forms of institutional performance, including the propensity to innovate. To those invested in the institution, deliberative instrumentality is, all by itself, a bottom line. It is internal legitimacy.

All versions of managed innovation call for greater precision in the measurement of results. In technology, this makes great sense because we want software, medicine, and machines to do what they are supposed to do. The problem is exporting this logic to the murky world of human behavior and collective action. Virtually all research on the subject indicates that the challenges to precise measurement of performance in the public and nonprofit sectors are legion. On this point, I am partial to the wisdom of Sir Josiah Stamp (1927):

Public agencies are very keen on amassing statistics — they collect them, add them, raise them to the nth power, take the cube roots, and prepare wonderful diagrams. But what you must never forget is that every one of those figures comes in the first instance from the village watchman, who just puts down what he damn pleases.

All effective organizations must have precise revenues, expenditures, budgets, and instruments of measuring how well they are accomplishing their purposes. But, following

managed-innovation assumptions, it is understood that performance measures are exactly that — measures — and that they only *represent* reality. They are subject to scrutiny, to debate, and even to dismissal if they are nonsense. Performance measures seldom answer questions; they *are* questions. Under conditions of institutional innovation it is far better to approximate an answer to an important, if not fully defined, question than to search for an exact answer to an unimportant question simply because precise measurement is possible. Such answers can always be made more precise; they are still answers to the unimportant questions.

**Values.** No difference between the logic of managed innovation and sustained innovation is greater than their contrasting assumptions regarding values. Sustained-innovation values are understood to be historical, natural, shared, and enduring. Individuals, it is assumed, are attracted to the institution, at least in part, because of the values they see in its purposes. There is, for example, ample evidence that individuals, if given an open choice, will choose to associate with institutions that bolster their own beliefs and characteristics. Persons who go into social work, teaching, accounting, law, business management, the ministry, and public service differ. It has long been understood that institutions are not only places where people work together and share values, but also systems of shared meanings, a common language, and strongly reciprocal individual-institutional linkages by which institutional values influence the individual and vice versa. Institutional systems of shared meanings are layered, historical, evolving, autonomous work cultures that take on a natural organic quality. Such systems carry the institution's values and generally reflect the values of those who make up the institution. Such organizations are sometimes described as strong-culture institutions. Most research indicates that the effective management of strong-culture institutions is usually homegrown and guided by what Olsen, et al. call "the gardener" — which is to say, someone who is very familiar with all parts of the gardening process, works with what is at hand in the appropriate seasons, anticipates drought and insects and, at the harvest, is prepared for the next season. The gardener tends to the values of the garden by managing the garden. Leaders, following managed-innovation assumptions, assume

not only management of the garden but also set its values.

Examples of continuing innovation in strong-culture organizations in the private sector would include Nordstrom's culture of customer service, Ben and Jerry's social agenda, Procter and Gamble's attention to employee well-being, and Southwest Airline's emphasis on both fun and bargain fares. Public-sector strong-culture organizations include NASA, particularly in the 1960s and 1970s, the U.S. Secret Service, the U.S. Forest Service, the Federal Bureau of Prisons, in recent years the Social Security Administration, the General Accounting Office, and many police and fire departments. Strong culture in the nonprofit sector includes most large research universities and many leading liberal arts colleges, most large-scale philanthropies and many interest-based organizations such as the NRA and NOW. All reflect a continuing emphasis on a set of agreed-upon values and shared meanings. The problem is, they are also iron cages.

Strong-culture institutions do change and can encourage innovation, but they will resist attempts to manage the processes of change and innovation. Great strong-culture institutional leaders are both *in* the institution and *of* the institution, unquestioned adherents to its shared meanings. Such leaders can carefully and patiently nurture the processes of change and keep open the prospects for innovation.

**Propensity to Risk and to Innovate.** Results are never neutral and all solutions have some negative side-effects. Organizations in the modern marketplace, particularly in high technology, lean strongly in the direction of rapid change, strong leadership, a transient workforce, and the logic of managed innovation. They trade worker loyalty, brand identification, and long-term institution-building for rapid response and high flexibility. Management consultants and strong leaders in such settings may claim that they can do it without these tradeoffs, but the evidence is to the contrary. Businesses in more stable markets tend to exhibit fewer tradeoffs and institutionalize more — Wal-Mart, Procter and Gamble, and General Electric, for example.

Government institutions are, by definition, more stable and permanent, obviously giving away some rapid response, some malleability and, because of the separation of powers, some strong singular leadership, in return for order and predictability. Many of those who would import business reforms to govern-

ment, including TQM and managed innovation, will claim they can somehow obviate these tradeoffs. The evidence is not with them.

For our purposes, we must ask if one institutional type is more inclined to creativity, originality, and innovation than the other. The answer is that the two types are differently original and innovative. The threat of competition does appear to stimulate innovation in short, rapid cycles, particularly in the high-technology marketplace. But it also results in the iron cage of isomorphism. Innovations in the public sector tend to be slower, take longer, and last longer. Governments build their own iron cages, but it should not be assumed that governments are not creative or innovative. Just like businesses, some are and some are not.

Nonprofits are highly varied, ranging from rather stable institutions such as foundations and universities, to highly volatile human service contract organizations. The forms of innovation and tradeoff appear to be determined by context — nonmarket nonprofits acting more like public institutions in their innovative behavior and market-style nonprofits acting more like businesses. How they respond to the logic of managed innovation depends on their context, their culture, their leadership, and whether they seek short-run or long-range innovation. But nonprofits and even foundations appear to be homogenizing, converging, and entering the iron cage because rewards in the form of reputational capital go to the highly ranked, not the deviant.

## IX

In the relentless tide of reform and the search for innovation, nothing better illustrates the weakness of benchmarking and managed innovation than the latest idea. That idea, the balanced scorecard, is the brainchild of Robert Kaplan and David Norton (1996). Balanced scoring, not unlike grading or ranking, is used to measure not only the obvious financial bottom line in the short run, but also the effectiveness of management and employee growth and learning. Scoring is balanced, implying that benchmarking and managed innovation is not balanced because of its lack of attention to employee

growth, participative forms of organization, and organizational stability. Balanced scorecard advocates argue that too many organizations are underinvested in their employees and in issues of employee security and satisfaction. Better scores on investments in training, professional involvement, community participation, and workforce stability will, it is claimed, sharply improve the prospects for continuing innovation. Such an argument is also a critique of benchmarking. In the balanced scorecard, the emphasis on best practice is only in terms of comparative scores and not in terms of borrowing the innovations of others or directly presuming to manage innovation. Both are connected to the logic of ranking and lead to homogenization rather than differentiation.

The balanced scorecard is increasingly popular with management consultants and the businesses that hire them. The balanced scorecard has migrated to government and nonprofit organizations much as TQM and benchmarking did. The isomorphic effects of balanced scorecards are under way.

Innovation is both feasible and necessary. But serious organizational innovation is neither easy nor cheap. The primary work of the public and nonprofit sectors is to achieve public interest and to serve the people. Innovation in the service of the public interest should not be cheated. The potent mixture of rankings, performance measurement, and the logic of managed innovation cheats real, lasting innovation. In the same way that the overemphasis on testing and the linking of testing to ranking and status has caused schools and colleges to teach to the test, the assumptions of institutional rankings and prizes combined with the logic of managed innovation reduces the probability of innovation and especially lasting innovation. (Lehman, 1999; Sacks, 2000). To weaken the iron cage, it is essential to break the weld that links the instinct for organizational innovation to rankings, prizes, status, and legitimacy. Once that weld is broken, the organization can worry less about legitimacy and prestige and more about genuine creativity. Only then, can it break out of the iron cage. Virtually all the empirical evidence is to the effect that public-sector and nonprofit creativity is more likely, more enduring, and more effective when it follows the logic of sustained innovation. It is the way to ameliorate the effects of the iron cage, to become genuinely creative, original, and innovative.

## REFERENCES

- Blazey, Mark L. 1998. *Insights to Performance Excellence 1998: An Inside Look at the 1998 Baldrige Award Criteria*. Milwaukee, WI: ASQ Quality Press.
- Bryson, John M. 1995. *Strategic Planning for Public and Nonprofit Organizations*. San Francisco, CA: Jossey-Bass.
- Camp, Robert C. 1995. *Business Process Benchmarking*. Milwaukee, WI: ASQC.
- Christensen, Clayton M. 2000. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Cambridge, MA: Harvard Business School Press.
- Collins, James C. and Jerry L. Porras, 1994. *Built to Last: Successful Habits in Visionary Companies*. New York: Harper Business Publishing.
- Cooper, Terry L. and N. Dale Wright. 1992. *Exemplary Public Administrators: Character and Leadership in Government*. San Francisco, CA: Jossey-Bass.
- Cyert, Richard M. and James G. March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.
- Deming, W. Edwards. 1988. *Out of Crisis*. Cambridge, MA: MIT Press.
- DiMaggio, Paul J. 1983. "State Expansion and Organizational Fields," in *Organization Theory and Public Policy*, R. H. Hall and R. E. Quinn, eds. Beverly Hills, CA: SAGE 147-161.
- DiMaggio, Paul J. and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields," *American Sociological Review* 18:147-160.
- Frederickson, H. George and Jocelyn M. Johnston, eds. 1999. *Public Management Reform and Innovation: Research, Theory and Application*. Tuscaloosa, AL: The University of Alabama Press.
- Gorman, Jack. 1989. *The Gorman Report: A Rating of Graduate and Professional Programs in American and International Universities*. Los Angeles, CA: National Education Standards.
- Gormley, William T., Jr. and David L. Weimer. 1999. *Organizational Report Cards*. Cambridge, MA: Harvard University Press.
- Hirsch, Paul M. and Thomas Whisler, 1982. "The View from the Boardroom," a paper presented at the Academy of Management Meetings. New York, NY. *Kansas City Star*. February 5, 2000.
- Kaplan, Robert S. and David P. Norton. 1996. *The Balanced Scorecard: Translating Strategy into Action*. Cambridge, MA: Harvard Business School Press.
- Lehman, Nicholas. 1999. *The Big Test: The Secret History of the American Meritocracy*. New York: Farran, Shaws & Disou.

- Light, Paul C. 1998. *Sustaining Innovation: Creating Nonprofit and Government Organizations that Innovate Naturally*. San Francisco, CA: Jossey-Bass.
- Light, Paul C. 1999. *The True Size of Government*. Washington, DC: The Brookings Institution.
- Lindbloom, Charles E. 1959. "The Science of Muddling Through," *Public Administration Review* 19:79-88.
- Lynn, Lawrence, Jr. 1993. "A Challenge to Reinventing Government," Unpublished Paper presented to the National Academy of Public Administration, Washington, DC.
- March, James G. and Johan P. Olsen. 1976. *Ambiguity and Choice in Organizations*. Bergen, Norway: Universitetsforlaget.
- Nayak, P. Rangarath and John M. Ketteringham. 1994. *Breakthroughs!* San Diego, CA: Pfeiffer & Co.
- Olsen, Johan P. 1988. "Administrative Reform and Theories of Organization" in *Organizing Governance: Governing Organizations*, edited by Colin Campbell and B. Guy Peters. Pittsburgh, PA: University of Pittsburgh Press.
- Ouchi, William G. 1980. "Markets, Bureaucracies and Clans." *Administrative Science Quarterly* 25:129-141.
- Perrow, Charles W. 1974. "Is Business Really Changing?" *Organizational Dynamics*. Summer, 31-44.
- Radin, Beryl. 2000. "The Government Performance and Results Act and the Tradition of Federal Management Reform: Square Pegs in Round Holes?" *Journal of Public Administration Research and Theory* 10:1, January.
- Rogers, Everett M. 1962. *Diffusion of Innovations*. New York: The Free Press of Glencoe.
- Sacks, Peter. 2000. *The Standardized Mind: The High Price of America's Testing Culture and What We Can Do to Change It*. New York: Perseus Books.
- Stamp, Josiah, 1927. "The Statistical Verification of Social and Economic Theory." The Sidney Ball Lecture, Nov. 5, 1926. Barnett House Papers, No. 10, London, England: Oxford University Press.
- Tolbert, Pamela S. and Lynne G. Zucken, 1983. "Institutional Sources of Change in the Formal Structure of Organizations: The Diffusion of Civil Services Practices," *Administrative Science Quarterly* 28:22-39.
- Watson, Richard D. 1991. *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*. New York: Lawrence Bragg Mass Market Paperback.
- Weber, Max. 1952. *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribner.
- Williamson, Oliver E. 1975. *Markets and Hierarchies*. New York: Free Press.





*Kettering  
Foundation*

200 Commons Road, Dayton, Ohio 45459 • (937) 434-7300  
444 North Capitol Street N.W., Washington, D.C. 2001 • (202) 393-4478  
6 East 39th Street, New York, New York 10016 • (212) 686-7016