

## Chapter 12

# Performance Measurement

## The Critical Competence

*They say that figures rule the world. I do not know if this is true, but I do know that figures tell us if it is well or poorly ruled.*

GOETHE

---

***Performance Measurement creates information about the results of public activities. This enables officials to hold organizations accountable and to introduce consequences for performance. It helps citizens and customers judge the value that government creates for them. And it provides managers and employees with the data they need to improve their performance.***

---

..... **I**s government performing well or poorly? Performance measurement answers that question by creating information about the results produced by departments or agencies, programs, work teams, and even individual employees. This information is indispensable if you want to use the Consequences Strategy to transform government. If you do not measure performance, you cannot manage for it, reward it, contract for it, or even identify the bottom lines for which public organizations will be held accountable. In short, if you cannot measure performance, you cannot tie incentives to it.

Performance measurement is also a core competence needed to implement the other four strategies. To use the Core Strategy, you need information about how well both steering policies and rowing institutions are working. A performance measurement system allows legislators and elected executives to specify the results they want and to determine which organizations are delivering those results.



[Click Here to order](#)

To use the Customer Strategy, you need information about how well an organization's performance satisfies its customers. With that information, customers can make informed choices about which organization to use; thus, organizations become more accountable to them.

To use the Control Strategy, you need to replace traditional hierarchical control systems with a new form of control: accountability for results. Decentralization of authority without reliable performance information can quickly lead to problems.

Finally, performance information helps immensely in changing an organization's culture. In the words of Duke professor Robert Behn, it helps everyone use "the same definition of winning." It helps employees understand how their work contributes to the organization's success. And it helps reinforce the values of improvement and innovation.

Clearly, performance measurement plays a central role in the process of reinvention. By itself, however, it is insufficient. It creates awareness but not always action. In the 1970s, New York City began generating a thick volume of performance data about city agencies every year. But until the Giuliani administration began using this information to manage performance, it mostly gathered dust. Few managers and no elected officials used it to make decisions about budgets or personnel—so most public employees ignored it.

In short, performance measurement points the way, but it doesn't necessarily move anyone's feet—particularly when there are bureaucratic and political brick walls standing in the way of change. This is why we call performance measurement a *core competence* of government rather than a tool that activates people. To force significant changes in organizational behavior, you need incentives that give employees a reason to respond to performance data. "Without the consequences," says Don Forbes, former director of the Oregon Department of Transportation, "you shouldn't bother measuring performance."

---

## UNDERSTANDING THE ARCHITECTURE OF PERFORMANCE MEASUREMENT

Governments measure all kinds of things. They typically drown themselves in information. But almost all of it concerns *inputs*. In the past, real information about outputs and outcomes was rare.

Fortunately, this is changing rapidly. Beginning in the 1980s, entire nations, states, and local governments have committed themselves to measuring performance, including New Zealand; the U.K. (both the national government and many local authorities); Australia; U.S. states such as Oregon, Florida, Texas, North Carolina, Minnesota, Iowa, and Arizona; and dozens of cities and counties. The U.S. government took the pledge in 1993, when Congress passed the Government Performance and Results Act (GPRA), originally drafted by John Mercer, a former mayor of Sunnyvale, California, who had gone to work for the U.S. Senate Governmental Affairs Committee.



[Click Here to order](#)

The GPRA experience demonstrates just how difficult performance measurement is, particularly at the beginning. In early 1994, twenty-one federal agencies became GPRA pilots. They were on their own; they received no additional resources and—by design—no assistance from the Office of Management and Budget. They had four months to produce their plans for performance measurement.

When they submitted their plans, a panel of experts convened by the National Academy of Public Administration (NAPA) examined them. It found that many were still confused about what performance measurement was, even after months of trying to understand it. For most, the subject was completely new—as it is for most public organizations. To make matters worse, experts offer different and sometimes conflicting explanations and definitions. “Around the world,” writes NAPA fellow Donald Kettl, “measuring government performance is like the weather. Everyone talks about it. . . . But there is no consensus on how to do it.”

The first step to cut through this confusion is to define a “conceptual architecture” that explains what you want to measure and how you will measure it. Our examination of performance measurement systems suggests that this architecture has five linked components:

<b>Component</b>	<b>Examples</b>
Policy Outcomes	Unemployment rate, water quality, literacy rate
Program or Strategy Outcomes	Job training participants hired, reduction in industrial water pollution, increase in reading levels at fourth grade
Outputs	Unemployed people trained, water pollution permits granted, students taught in after-school reading program
Processes	Recruitment, registration, training classes; educating businesses, intake, permit processing; selection of students, instruction, self-paced learning with software
Inputs	Employees, budgets, equipment, contractors

The first component is called *policy outcomes*. These indicate the effectiveness of government policies in achieving the basic goals of a nation; a state, province, region, or county; or a community. For example, economic policy



[Click Here to order](#)

outcomes include unemployment rates, inflation rates, poverty levels, and trade balances. Environmental policy outcomes include public health, air and water pollution levels, soil erosion, and the like.

Typically, policy outcome goals come in two varieties: long-term (10–20 year) outcome goals, and intermediate (2–5 year) outcome goals, which are intended to contribute to achieving the long-term goals. For example, a long-term outcome goal might be to stabilize global warming; related intermediate goals might be to decrease oil consumption and to decrease the total number of miles driven by automobiles. Meeting these intermediate goals is not sufficient, because other things may be happening to intensify the problem. But unless we measure them—and others—we won't have a short-term handle on whether we are progressing toward the long-term goal.

Citizens judge politicians' performance, at least to some extent, by looking at how well they deliver these policy outcomes: Is the economy good or getting better? What is happening to the crime rate? How clean is the air we breathe?

Policy outcomes are shaped by many factors, some of which are outside the control of government. Public leaders try to affect them by creating *program and strategy outcomes*, the next component in the measurement architecture. Program and strategy outcomes indicate the effectiveness of government programs, strategies, regulations, or other activities at achieving the desired policy outcomes. (These don't necessarily correspond to organizations: they may involve several organizations; one organization may house several programs; and one program may involve multiple strategies.) One program outcome goal, for example, might be to place 80 percent of all participants in a government training program in jobs. If that goal is achieved, it will contribute to the policy outcome goal of lowering unemployment.

A public organization's success in creating program outcomes depends on the *strategies* it uses. For example, street sweeping is one strategy for producing clean streets; an antilittering campaign would be a different strategy. A sanitation department can measure the outcome of each strategy: how clean the streets are just after the department sweeps them and how clean they are at other times. In some organizations, it is useful to measure both program and strategy outcomes, because they are so distinct. In others, they are essentially the same, so it makes sense to measure only one. Even when they are distinct, measuring both may introduce too much complexity. This is why we treat program and strategy outcomes as one component of the architecture.

An organization's success in creating positive strategy and program outcomes depends on its *outputs*—the actual work products it produces. Sanitation departments produce outputs like miles of streets swept and numbers of household garbage pickups. The Social Security Administration produces ben-



[Click Here to order](#)

efits checks. Environmental agencies produce regulations, inspections, and fines. Police and fire departments produce emergency responses, arrests, traffic tickets, and the like. These outputs lead to program and strategy outcomes.

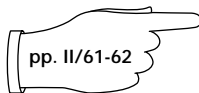
Agencies can usually measure the quantity and cost of their outputs and the efficiency with which they are produced. They can also measure their quality (for example, their accuracy and timeliness) and their effectiveness (the degree to which they produce the desired outcomes). Government agencies usually have substantial control over outputs. But often they provide grants to other levels of government or contracts to private organizations to produce outputs. In such cases, they depend on these other organizations to produce services of a certain quantity, cost, quality, and so on.

Outputs are created by *processes*, or activities. (Some jurisdictions distinguish between activities, which they define broadly, and processes, which are sub-elements of the activities.) These are the production methods of government, the work that is actually performed. Social Security Administration employees determine applicants' eligibility for benefits, establish benefit levels, process and deliver payments, and check for fraud. Sanitation workers set up collection routes and schedules, pick up trash, and haul it off to landfills. Law enforcement officers patrol communities, respond to complaints, subdue perpetrators, read them their rights, take them in to be booked, file reports, and testify in court. Performance measures for processes include efficiency (how much they cost to perform), quality (for example, how much time they take), and effectiveness (how often they produce the right output).

Processes depend on *inputs*—the resources that are required to create them. Employees, salaries, information, offices, computers, money, garbage trucks, uniforms, guns, and prisons are all examples of inputs. Here quantity, cost, efficiency, and quality can be measured.

The five components of performance measurement are connected one after another and work together, like a production line: inputs create processes, which create outputs, which determine strategy and program outcomes, which impact policy outcomes.

Inputs → Processes → Outputs → Strategy/Program Outcomes → Policy Outcomes



Performance measurement systems should start by defining policy outcome goals and then work their way down—as we explained in Chapter Five. Few governments have developed systems that link each of these components. Those that do refine this architecture in different ways and use different terms to describe it. In New Zealand, for instance, the government now uses a framework of long-term “goals” (policy outcomes), “strategic priorities” (the most important intermediate policy outcomes for the next three to five years), “key



[Click Here to order](#)

result areas” (ideally, each department’s most important program and strategy outcomes), and “outputs.” Sunnyvale uses “program outcomes,” but calls outputs “products” and calls processes “activities” and “subactivities.” You should use whatever language works best for you; if your employees intuitively grasp what a product is but not an output, call them products.

It is the focus on results rather than activities that is most important. We stay away from the traditional management-by-objective phrase “goals and objectives,” for example, because so many organizations use this language to measure activities rather than results. We purposefully use language—such as *output goals* and *outcome goals*—that highlights the difference between means and ends.

When people ask how well government is performing, typically they want to know what impact it is having: what are the policy and program outcomes? But when people want to improve program outcomes, they must turn to government’s strategies and outputs. Is it using the best strategies? Are they producing the right outputs? Would different outputs and strategies lead to better outcomes?

When people want to change or improve government’s outputs, they must change its processes and activities. Unless activities and processes change, outputs remain the same.

Finally, new processes often require new inputs, and improved processes often require improved inputs—or better use of existing inputs.

The other critical architecture question is what to measure about each of these elements. Your specific measures, often called indicators, can gauge the quantity, efficiency, effectiveness, quality, or cost-effectiveness of an outcome, output, or process. Unfortunately, few people have ever carefully defined what each of these terms means. Yet lack of precision can quickly lead to confusion. It can also create blind spots: areas no one measures because there is no label for them in the measurement system.

There is no final arbiter of the meaning of these terms, no *Webster’s Dictionary* of performance measurement. However, we offer below our best effort to define them in a way that focuses each term on a distinct issue—and thus gives you the broadest possible set of terms with which to measure your processes, outputs, and outcomes.

**Quantity:** How much of an output you produce: how many people are trained, how many permits are processed, and so on. This is the least important element on our list.

**Efficiency:** The cost per unit of process or output: that is, the ratio of inputs to outputs or processes. Most processes produce outputs, so measuring the cost of those outputs is, in effect, the same as measuring the cost of the process. But sometimes it takes multiple processes to produce a single out-



[Click Here to order](#)

put. In welfare offices, for example, eligibility workers may need to complete several different processes to produce an output, such as a determination of eligibility. Thus to improve their overall efficiency they might want to measure the efficiency of each process. (*Productivity*, the ratio of inputs to outputs—that is, how many outputs you can produce for how many dollars—is a subset of efficiency.)

**Effectiveness:** How successful inputs and processes are at producing desired outputs, outputs are at producing desired program or strategy outcomes, and programs or strategies are at producing desired policy outcomes. You can do a splendid job of hitting output goals—such as a certain number of people graduating from a training program, or a certain number of students passing a course—while failing completely to produce the outcomes you want, such as people finding jobs or acquiring skills. Effectiveness is the most important thing you can measure. If a program or process is not effective, why worry about how efficient it is?

**Quality:** How well an activity or process is performed or an output is produced. This is not quite synonymous with effectiveness, because one can measure the quality of a process or output without measuring its effectiveness. For example, one can measure the quality of telephone service delivered by an office: how quickly they answer the phone, how helpful they are, how satisfied the people on the other end of the conversation are. This won't necessarily tell you how *effective* their phone service is, however. In fact, you can provide high-quality services that are not effective at achieving your desired outcomes. For example, a welfare eligibility office could offer customer service of such high quality that welfare applicants and recipients find it appealing to remain on welfare—frustrating the desired outcome of moving people into jobs.

The British Treasury's publication *Executive Agencies: A Guide to Setting Targets and Measuring Performance* provides useful lists of aspects of quality you can measure. They include the following:

*Quality of Outputs*

- Accuracy (for example, the error rate).
- Did the output meet technical specifications and standards?
- Clarity of information or advice.
- Response to errors, complaints, or failures.
- Customer rating of general quality.



[Click Here to order](#)

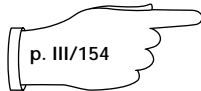


*Quality of Processes*

- Timeliness (for example, waiting times, turnaround times, and processing times).
- Backlogs.
- Responsiveness (for example, timely feedback to customers).
- Pleasantness, helpfulness of staff.
- Accessibility of services to customers.
- Availability of products, such as documents and forms.

**Cost-effectiveness:** The ratio of inputs to outcomes—that is, what level of outcomes can you achieve for the dollars you spend? This is what the British mean by “value for money.” It answers the question “How effective is your spending?” It measures not how many outputs you can produce for your dollar but how much value you produce. As with effectiveness, you can measure the cost-effectiveness of *strategies, programs, and policies*.

Table 12.1 provides examples of these categories, applied to each level of the performance measurement architecture. This is quite a complete framework, which should help you see all the particular elements you could include in your measurement system. *But don't assume that you should use all of it.* Pick and choose what is most important for your organization to measure; otherwise your staff will be overwhelmed. You may only want to measure process data when you're trying to redesign a process, for example. You may even need to minimize how much output data you collect, except where you see outcomes faltering. Information is only worth something if people use it, and if they're drowning in it, they won't use it.



Some people use the metaphor of a dashboard to explain this idea. When you drive a car, you need to know a few key things: How fast are you going? How much gas do you have left? Is the engine overheating? And do you have enough oil pressure? You don't need to know what the pressure is in each cylinder or what the air pressure is in each tire. You wouldn't want to drive without a dashboard—but if it were as complex as an airplane's dashboard, you'd probably never use it. At each level of your system, you want people to have a dashboard that works for them.



[Click Here to order](#)



## The Consequences Strategy

Measures of →	Quantity	Efficiency	Effectiveness	Quality	Cost-effectiveness
<b>Policy Outcome Goals:</b> Clean Air  Low crime rates  Skilled workforce	NA (Not Applicable)	NA	Air pollution level  Crime rate  Percentage of workforce with high school and college degrees	NA	Air pollution level ÷ cost  Crime rate ÷ cost  Degree percentage ÷ cost
<b>Programs/Strategy Outcome Goals:</b> Environmental Protection Agency: Reduction in industrial pollution  Police Department: Reduction in violent crime rate  State Colleges: Increase in percentage of college entrants who graduate	NA	NA	Volume of industrial emissions  Violent crime rate  Percentage of college entrants who graduate in X years	NA	Industrial emissions volume ÷ cost  Violent crime rate reduction ÷ cost  Percentage of college entrants who graduate in X years ÷ cost
<b>Outputs:</b> Air pollution permits processed  Arrests made  Students graduated	Number of permits processed  Number of arrests  Number of graduates	Cost per permit processed  Cost per arrest  Cost per graduate	Volume of industrial emissions  Percentage of convictions  Percentage of college entrants who graduate in X years	Percentage of permits processed by deadline  Percentage of arrests thrown out by courts  Satisfaction level of graduates, on survey	NA
<b>Processes:</b> Educating businesses; intake; permit review	Number of businesses advised; number of permits reviewed	Cost per business advised; cost per permit review	Percentage of businesses that complete permit application properly; percentage of permit decisions overturned on appeal	Satisfaction of businesses with advising process; number of complaints; average time required for permit review	NA

Table 12.1. Performance Measurement Matrix



[Click Here to order](#)

## The Consequences Strategy

Measures of →	Quantity	Efficiency	Effectiveness	Quality	Cost-effectiveness
<b>Processes:</b> Investigation; arrest; booking  Registration; courses; advising	Number of investigations, arrests, bookings  Number of courses; number of advising hours	Cost per investigation, arrest, bookings  Cost per course; cost per student advised	Percentage of investigations that lead to arrest  Percentage of students registered in classes desired; student ratings of courses	Error rate in booking  Student ratings of registration staff, faculty, advisors; percentage of students registered on time	
<b>Inputs:</b> Employees, salaries, equipment, overhead	Number of employees; cost of salaries, equipment and overhead	Cost per employee; percentage of indirect (overhead) costs to direct costs	Percentage of employees who have mastered required skills; percentage of equipment downtime	Employee satisfaction level; average rating of employees in 360 degree evaluation; employee ratings of equipment quality	NA

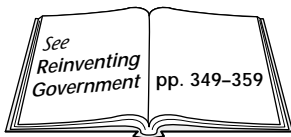
Table 12.1. Performance Measurement Matrix, cont'd.

## IMPLEMENTING PERFORMANCE MEASUREMENT: LESSONS LEARNED

Performance measurement is not easy. It typically takes about three years to develop an adequate set of performance measures. The first time around, most agencies find themselves measuring inputs and processes, not outputs and outcomes. They often go through several iterations—nudged along by a neutral body that has authority to approve their measures—before they get the focus squarely on results.

Performance measurement requires a combination of technical, managerial, and political savvy. Because it employs highly rational, scientific observation and analysis, it demands technical expertise. Because it must be implemented in a messy, changing environment, it requires managerial skill. And because it influences public perceptions of elected officials—as well as elected officials' decisions about policy and budgets—it requires political sensitivity.

*Reinventing Government* addressed some of the most important challenges in an appendix, “The Art of Performance Measurement.” It contained a glossary of terms and a list of books from which you could learn more. It also provided advice to practitioners, including the first four lessons on our list. See pp. IV/71-76 in this book for a few additional pointers.



[Click Here to order](#)

### IMPLEMENTING PERFORMANCE MEASUREMENT: LESSONS LEARNED

1. Measure qualitatively, not quantitatively.
2. Anticipate powerful resistance to performance measurement.
3. Keep the measurement function politically independent and impartial.
4. Watch out for perverse incentives.
5. Begin with policy outcome goals and work downward.
6. Don't just measure efficiency.
7. Measure customer satisfaction.
8. Involve your customers in choosing performance indicators.
9. Watch out for overkill: don't try to measure everything.
10. To get a full picture, create a balanced scorecard.
11. Phase in the new system gradually.
12. Continuously improve your system.
13. Don't attach consequences to performance too soon.
14. Standardize, but don't centralize.
15. Build a quality gatekeeper into your system without taking too much control away from the agencies.
16. Use experts in the design stage.
17. Top managers have to get their hands dirty.
18. Communicate with the unions.
19. Train, train, train.
20. Automate the system.
21. Make performance data usable and visible.
22. Beware of tempting measurement shortcuts.
23. Expect setbacks.
24. If you're not going to use performance data to make decisions, don't bother collecting it.



[Click Here to order](#)

**1. Measure qualitatively, not just quantitatively.**

Some results are impossible or impractical to quantify and should be measured qualitatively. Don't just go by the numbers; use on-site observations, peer review, professional evaluations, employee interviews, customer focus groups, and other methods of gathering information.

**2. Anticipate powerful resistance to performance measurement.**

Performance information can be extremely threatening to service providers—contractors, grantees, or public employees—who worry about how well they will perform. They will mobilize political opposition to the process. Bringing them in to help design the measurement system can help address their concerns and defuse their hostility.

**3. Keep the measurement function politically independent and impartial.**

Performance data must be trustworthy; it must be free from political or bureaucratic taint. It is important to maintain the integrity of the data gathering and analysis processes. Organizations should collect most data themselves, so they get quick feedback and can learn from it. But make sure an independent unit, such as a city auditor, audits the data for accuracy.

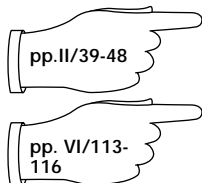
**4. Watch out for perverse incentives.**

Sometimes using the wrong performance measures leads organizations to do the wrong things. For instance, if a police department is measured on the number of arrests it makes every month, its officers may make hasty arrests and work hard on cases that are easy to solve. To avoid perverse incentives, test them first for problems.

**5. Begin with policy outcome goals and work downward.**

“You can't measure for results if you don't know where you are going,” notes Camille Barnett, former city manager of Austin, Texas. Hence you should begin the measurement process at the top, with your overall mission and policy outcome goals. Once you have done this, define missions and outcome goals for your steering organizations. Then move to the operational agency level. Define each agency's mission: just what is it supposed to accomplish? Then begin defining key outcome goals for each of the agency's programs and/or strategies. (Don't try to do this for the agency as a whole, unless it is very small. When large agencies try to define agency-wide outcome goals, they often discover that it doesn't work, because the agency is really an agglomeration of disparate programs and strategies.)

Although starting at the top may seem obvious, many organizations learn it the hard way. The Oregon State Police provide a perfect example. Its top



[Click Here to order](#)

managers started by drafting about 270 measures. When they asked field staff for feedback, it was negative: employees couldn't connect the measures to the agency's priorities. Management stopped the process, realizing that it had to begin with a clear understanding of the department's goals. "We have to clarify some organizational goals and values as early as possible, and it has to come from the top," Major Lee Erickson told us.

*Otherwise there is no sense of direction. It's kind of like a bunch of horses milling around the infield. If you get them hitched up and pulling in the same direction, there's real power. But we don't have clarity of purpose.*

Most of the U.S. GPRA pilot agencies experienced a similar problem: NAPA found that their measures were unrelated to their missions.

Once you have program and/or strategy outcome goals, you can build down through the organization, measuring outputs, processes, units, teams, and—if appropriate—individuals. If you try to build up from units, teams, and individuals, you will probably discover at some point that your measures have little to do with the outcomes you're after.

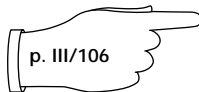
#### 6. *Don't just measure efficiency.*

Often finance agencies—treasury departments and offices of management and budget—lead the charge on performance measurement. This was the case, for example, in Australia, the U.K., and New Zealand. Quite naturally, their focus was on the cost and efficiency of government outputs.

Unfortunately, this approach ignores quality, effectiveness, and cost-effectiveness. As we asked earlier, what good does it do to make services more efficient if they are ineffective—if they are not achieving the desired outcomes?

This problem is quite common. After almost 15 years of performance measurement, Sunnyvale decided it was too focused on productivity—often to the exclusion of effectiveness. So it revamped its system, which it now calls "outcome management." In Australia, a review of performance reporting found that, after nearly a decade of finance-driven efforts, agencies collected little data on outcomes. The same thing happened in the U.K.

Focusing only on financial measures can also generate resistance among employees and managers. Employees are typically suspicious of efforts to improve "productivity," because they assume that means making them work harder—or cutbacks and layoffs. In contrast, they usually like efforts to improve "quality" or "effectiveness," because many of them want to improve service to their customers.



[Click Here to order](#)

### 7. *Measure customer satisfaction.*

Perhaps the most difficult step for many organizations is finding quantifiable measures of their effectiveness. Education officials endlessly debate the value of standardized tests and search for alternatives. Police departments debate whether the crime rate is the key measure of their effectiveness—or whether it is response time, or something else. Often, we have found, the best way to cut through such debates is to focus on customer satisfaction. How safe do the citizens feel? How satisfied are parents, students, and employers with the schools? For much of what government does, former Sunnyvale city manager Tom Lewcock points out, “A satisfied customer is the ultimate performance measure.” Sunnyvale now requires every city service to measure its customer satisfaction, including satisfaction with the timeliness, convenience, and accuracy of the service.

A classic example occurred during a seminar David Osborne conducted in Brazil. The director of a state industrial development corporation said he was having trouble figuring out how to measure the performance of his recruiting teams, whose job was to convince foreign companies to open plants in his state. Usually the companies chose their locations based on factors outside the control of these teams, like the cost of land, labor, and taxes, so success or failure was not a very good indicator of the teams’ performance. He was perplexed about what to measure. Osborne suggested that he ask the foreign firms to rate his teams’ performance. As it often does, the simplicity of this solution opened up a whole new vista for him. By the next day he had decided to use customer feedback to benchmark against his competition—and to build a financial reward system based on the results.

Customer satisfaction is not the only useful measure of effectiveness: you will need to measure other things. But it is often the most important one. It is also an attention grabber. Elected officials, public employees, and the public will pay attention to customer ratings. Customer feedback will not only help people understand the value of a performance measurement system, it will build support for developing one.

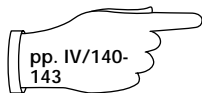
### 8. *Involve your customers in choosing performance indicators.*

Customers often understand better than anyone what else you should measure—because they know what is important to them and what isn’t. Unless customers are involved, says Camille Barnett, performance measurement becomes little more than “professionals and bureaucrats deciding what is best for the public.”

In Phoenix, City Auditor Jim Flanagan persuaded five departments to meet with their customers to develop measures. “We sat down in a room with people who run these agencies and people who are customers and simply discussed



[Click Here to order](#)



what quality is and what the indicators of success are,” he explains. “These were some of the best conversations ever held in Phoenix.” Other organizations create ongoing customer councils or boards to help with tasks like this.

**9. Watch out for overkill: don't try to measure everything.**

Don't measure what you won't use, and stop measuring what you don't need. Otherwise you will end up creating an enormous, unwelcome burden for your organization. “Perhaps our most glaring mistake was getting lost in the data,” confesses Craig Gerhart, who spearheaded the performance measurement effort in Prince William County, Virginia. “We overcounted; we overmeasured. We measured a lot of things because they were countable, and we mistook that for progress.”

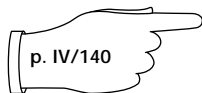
The point is not to measure everything. As officials from Atlantic Richfield told the U.S. federal government, measurement should be “a vehicle for focusing people's attention on the factors *critical* to the success of the organization.”

The British Next Steps Team makes the same argument. Based on the experience of executive agencies, it recommends four to seven key measures for any particular unit. Certainly you should avoid having any unit target more than 10 measures, or it will lose sight of what is most important.

One way to restrain “measurement creep”—the endless demands for more and more measures—is to identify the cost of measuring things. That way, when a legislator or manager willy-nilly orders up a new measure, you can ask how to pay for the additional effort, or which existing measures should be dropped.

It's also important to keep the measurement system from getting too complex. If people can't understand it, they won't use it. The Oregon Department of Transportation developed a system to weight the importance of different indicators and calculate a weighted average for each activity—a system we loved when we first saw it, because it was so sophisticated. But after several years, department officials abandoned it. It turned out that employees didn't use it, because they were befuddled by its complexity.

You have to find the right balance between too few measures and too many measures and between reporting systems that are too simple and those that are too complex. This is an art, not a science. You have to find the dashboards that work for your agency, your customers, and your funders.



**10. To get a full picture, create a balanced scorecard.**

As the British Treasury explains, “Any one measure or indicator, taken in isolation, may give a misleading picture. To form a complete understanding of what is happening, you need to look at the measures and indicators as a whole.” Some targets will drive you to achieve one objective at the expense of another:

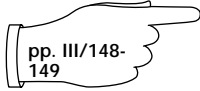


[Click Here to order](#)



increased efficiency, for example, at the expense of a higher error rate. “Timeliness, quality and cost are always in contention with each other, and the impact of improving any one or two must be weighed in relation to the expense to the third,” adds the U.S. Treasury.

The solution is what many call a “balanced scorecard,” which touches all the important bases without creating so many measures that its complexity overwhelms people. For a look at all the elements you can choose from, see Table 12.1.



### **11. Phase in the new system gradually.**

Don’t try to put an entire performance measurement system in place across your whole organization all at once. Pick a few programs to start with. “We didn’t do this, but I wish we had,” says Craig Gerhart, Prince William County’s budget director. When you try to tackle everything at once, the process may collapse under its own weight. Looking back at the collapse of the U.S. government’s complex Program Planning and Budgeting System in the early 1970s, experts in the U.S. Office of Management and Budget developed a carefully phased implementation for the 1993 Government Performance Results Act.

Implementing in stages allows you to show people how the new system will work and what it will be used for. You can use volunteer agencies to get the bugs out early on and build some successes. Phasing in also allows you to give employees a greater sense of control over the system; they can critique the early efforts and improve them.

### **12. Continuously improve your system.**

Because performance measurement systems are technical and complex, it is tempting to try to get the design perfect at the beginning. In reality, that is impossible.

“You have to go through several cycles to get good measures,” cautions Craig Holt, formerly a top manager in the Oregon Transportation Department. “We have our people check them and redo them after about six months.” Holt learned this the hard way; after the department “locked in” in some measures, he found out they weren’t the right ones. Typically, organizations go through three or four cycles before they are comfortable with their measures—and then they keep refining them periodically. As Sylvie Trosa concluded after her analysis of the Next Steps program, “Experience shows that implementing a good target setting process shared by all staff (a management culture), needs at least three years.”

Even when you develop appropriate performance measures, the data you gather may throw you a curve ball. For example, a sanitation department may



[Click Here to order](#)

find that a ticker-tape parade or spells of bad weather will throw their data off. The water department in Portland, Oregon, was rated highly until suddenly, in 1992, only 30 percent of citizens said they were satisfied with it. The reason: the previous summer's drought had led to severe restrictions on water use. "People could not water their lawns or wash their cars, so when they got the survey, they said, 'That damned water department!' and gave it bad ratings," explains Richard Tracy, director of city government audits.

All of this means that getting performance measurement "right" is an evolving process filled with constant adjustments—a process that never ends. When Sunnyvale began in 1980, City Manager Tom Lewcock advised his staff not to expect the system to be perfect—just to get it started and then improve it. They began by focusing on efficiency, then in the 1990s discovered the need to measure customer satisfaction. In 1995, they added a focus on outcome measures.

Over those 15 years, Sunnyvale learned a related lesson: Don't let your system get too rigid. When John Christian took over the Parks and Recreation Department in the early 1990s, he brought a focus on customer satisfaction with him. Yet the performance system forced him to concentrate on producing specific outputs—specific recreation programs and courses—regardless of what customers wanted. He found it difficult to change those targets—and impossible to do it quickly. As another manager put it, "We need to use the performance system as a tool, but it's become a straitjacket."

"We change our quality performance measures constantly," says General Michael Loh of the Air Combat Command (ACC). "Every time we review them—every quarter—we go back and change them and make sure we're measuring the right outputs and outcomes." Between 1990 and 1995, he says, the ACC changed about 50 percent of its measures to reflect changes in its basic mission.

But be careful not to change too much. You will want to be able to compare performance over time, and you can do this only if you have a core set of measures that remain unchanged.

### **13. Don't attach consequences to performance too soon.**

Although it is crucial to tie performance to consequences, there are several reasons to wait until the bugs have been worked out of the measurement system. We suggest that you build consequences in *after* the typical two- or three-year start-up period.

If you build consequences in too soon, people will argue endlessly about the validity of measures and data, because the stakes are so high. They will also fear that the system will be used to cut agency budgets and staffing—and thus resist it.



[Click Here to order](#)

There is also a good chance that an agency will discover that its performance is not very good. This is exactly what everyone needs to know, so that improvement can begin. But finding out poses risks. Everyone else—elected officials, the media, customers, the public—may find out, too. Then the blaming will start, particularly if the groundwork for understanding and using performance measurement has not been well laid. Attaching consequences early on only exacerbates this tendency.

#### **14. Standardize, but don't centralize.**

“Do not create a bureaucracy to administer or monitor a performance measurement program.” This warning—from officials of Boeing who were responding to a federal government survey on implementing performance measurement—is important. Every experience we have studied teaches the same lesson: if you centralize control of performance measurement—in a budget unit or departmental headquarters, for instance—those who must collect and use the data may never come to “own” it. Instead of welcoming performance measurement as a way of improving their work, they will resent it. They will view it as another administrative control imposed on them from above, and ignore or resist it. When his Oregon State Police made this mistake, says Major Lee Erickson, they ran into a wall of passive resistance. The attitude was, “It’s just one more thing headquarters needs; they’re cranking out numbers.” To get employee buy-in, get employees involved in creating the system, and let them run the system.

At the same time, you can't let each agency or unit go its own way on performance measurement. There have to be some basic standards, to ensure both quality and consistency. Otherwise, the information will become a Tower of Babel: many languages, no communication.

This is precisely what happened, initially, in Oregon. Budget Director Michael Marsh invited agencies to volunteer to create performance measurement systems. When they responded, he allowed them to invent their own systems. Different agencies built very different systems for tracking performance and reporting it. That made it difficult to aggregate data and report it as a whole, and it confused legislators.

The solution, Marsh decided, was to get agencies to agree on a basic framework—a set of standards—for what would be measured. A neutral agency such as a budget office or auditor's office can take responsibility for developing the framework and for helping other organizations use it. But it should be flexible: agencies should be free to vary what they measure, when appropriate.

Marsh found that to make performance measurement work, he needed both a commitment from key central agencies and leadership within the line



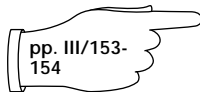
[Click Here to order](#)

agencies. To get the latter, he convinced agencies to create steering committees made up of the key champions of measurement: high-level managers; budget, personnel, and information systems staff; and a performance coordinator who became the in-house expert. The coordinator, who quarterbacked the implementation and training, had to have access to the agency director.

**15. Build a quality gatekeeper into your system without taking too much control away from the agencies.**

If you want useful measures that focus on the right things, you will need to put ownership of the measures in agency hands without abandoning quality control. In Prince William County, for example, Budget Director Craig Gerhart asked middle managers to get together with employees and customers and design appropriate performance measures—but no one had to approve their measures. “With this approach,” Gerhart says, “You are going to get a lot of bathwater with the babies. We acquired some really bad measures, and [several years later] we still have some.”

The best solution is to ask agency leadership to drive the process, but have a central team with expertise act as a coach and gatekeeper—helping agencies develop measures, pushing them to improve, and signing off when they have finally developed a useful set of measures. The gatekeeper should require that both agency employees and customers participate in developing the measures. It should require agencies to evaluate their measures periodically, with the help of a coach and a customer group. And it should ensure that someone audits the performance data that agencies produce—just as one would with financial data.



**16. Use experts in the design stage.**

Experts—consultants or practitioners with experience in performance measurement—can help you avoid many pitfalls. They can help you understand what developing a performance measurement system is all about, and they can help you solve technical measurement problems. But don't let them hand you an off-the-shelf system; make sure your system is customized to meet your needs.

**17. Top managers have to get their hands dirty.**

If top managers are not firmly committed to performance measurement, it will not work. But commitment is not enough. Executives have to get their hands dirty building the system. In many of the federal GPRA pilot agencies, NAPA found that top managers were largely absent from the process. This deprived the design teams of critical knowledge about the agencies' missions and objectives—the foundation of performance measurement.



[Click Here to order](#)

When top managers don't participate in the process, they also miss a critical learning experience. They don't learn how difficult measurement is, nor why patience is required. They don't recognize that they are not the only users of the system, that other potential users—legislators, the budget office, top executives, employees, customers, partners, and interest groups—are interested in different kinds of information. If top managers fail to understand these realities, they are likely to underinvest in the development process.

### **18. Communicate with the unions.**

Public employee unions often resist performance measurement for fear that it will lead to performance pay. This would break a fundamental union precept—"the same pay for the same work."

Most organizations don't use performance measurement to appraise individual workers—only individual *managers*. They do use it, however, to measure the performance of units and work teams. This does not usually cause problems with unions, unless it leads to performance pay.

The solution is to communicate clearly with the unions about your intentions, from the beginning. If you intend to create a performance bonus system, ask the union to participate in the design process. If it refuses, involve employees.

### **19. Train, train, train.**

As we keep saying, performance measurement is difficult. If you don't educate your people in how to do it, it will be even more difficult. "In the majority of pilot plans," NAPA reports, "the program people did not appear to have a full understanding of outcome measurement, the need to identify the various categories of their customers, [or] how to identify service quality and outcome indicators."

In Oregon, training began with a small group of volunteers from state agencies. It took about five months, after which the graduates became mentors for a second group. Doing it that way helped make the training feel like working with "a sister agency," says Mike Marsh.

### **20. Automate the system.**

Recognize that performance measurement is an *information* system. Unless you use information technology to run it, its care and feeding will be extremely labor-intensive. Practitioners advise that as you design your performance measurement system, you should automate the data collection and reporting processes as much as possible.

### **21. Make performance data usable and visible.**

Communicating about performance measurement is like telling a story. To keep the audience interested, you need to make the story both entertaining



[Click Here to order](#)

and relevant to their needs. This is true whether the audience is employees, managers, central agencies, elected officials, the press, or the public. There are a number of useful rules of thumb for doing this:

- **Make sure the performance information is accurate.** It has to be “viewed as valid, reliable, and objective, not public relations,” cautions Richard Tracy, audit director for Portland, Oregon.
- **Explain the context of the performance information.** Recognize that most citizens and many elected officials don’t know the history of government programs and agencies. Therefore, you must often provide them with brief background summaries of programs, the rationale for using particular measures, historical data trends, and other explanatory information.
- **Tailor publication of performance reports to specific audiences.** Because different audiences care about different measures, one size does not fit all. For example, since the public cares about results, it will pay attention to outcome measures but rarely to process measures. Yet employees need process measures to figure out how to improve their organization’s performance. Gauge your audiences’ needs and tailor your reports for them. Don’t make them wade through a sea of information to find what they care about.

Every year, Portland’s city auditor, an elected official, produces a report on the city’s performance. It contains information about the city’s nine largest services, covering at least 80 percent of its staff and budget. Each service gets its own chapter, which describes its mission and background; reports on its process, output, and outcome measures; and gives the results of citizen satisfaction surveys. The city auditor presents the report to the city council, holds a press conference, and sends it to neighborhood associations, community groups, and libraries.

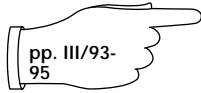
In 1995 and 1997, City Auditor Barbara Clark mailed residents a four-page, tabloid-size “Report to Citizens.” Much shorter than the detailed annual report, it contained easy-to-understand charts, maps, and information about the city’s spending, financial condition, and service results. It was not just a “good news” document: the 1995 version noted that the condition of the city’s streets had declined in 1994 and that city spending for fire, police, and sewer services was above the average of six comparable cities.

The idea was to increase citizen interest in the city’s performance, explains Richard Tracy, the audit director. “It’s a document the city can use to communicate with all sorts of people: citizens, outsiders, downtown businesses. We can give this to anyone and say, here’s what’s happening.”



[Click Here to order](#)





Local authorities in the U.K. go a step further. They publish similar documents but include comparative data supplied by the Audit Commission, which shows how local services compare in cost and quality to those of all other localities in their region.

### 22. *Beware of tempting measurement shortcuts.*

Because developing a performance measurement system is difficult and time-consuming, you will be tempted to make things easier, quicker, or less costly. Here are some temptations to avoid:

- **“Let’s just use the data we already have.”** Wrong! Much of it probably is not relevant; most of it is only about inputs and processes. The right thing to do is to stop gathering data that isn’t useful, so you will have time to gather data that is. This may require that the legislature and central administrative agencies eliminate some of their reporting requirements.
- **“Let’s just use year-end measures.”** Wrong! Managers need feedback more frequently than that. Some data is affected by seasonal variations or other factors that occur more often than once a year. And getting data more often will provide early warning signals and other important insights. Many agencies review performance data quarterly.
- **“Let’s not disaggregate the data too much.”** Wrong! Data should be disaggregated to the level of the unit responsible for performing the work, so it can be used to create consequences and improve performance. It should also be cut up by customer categories, by geographic area, and by any other categories that will illuminate the reasons behind different results.

### 23. *Expect setbacks.*

You will select some bad performance measures. Employees will resist collecting the data, and some managers will resist using it. People will complain about how much time it takes to develop and use the system. These problems are inherent in trying to develop a new organizational competence. Over time, you can work through them. Be patient. Acceptance will grow and flaws will fade. The one thing everyone who measures performance has discovered is that it is, to a degree, a self-correcting system: when you use the wrong measures, you immediately generate pressure to improve them.

### 24. *If you’re not going to use performance data to make decisions, don’t bother collecting it.*

If you’re not going to use performance data to reward employees and organizations, improve work processes, and allocate resources, don’t bother gathering it. It will become an expensive, time-consuming paper chase—one that builds cynicism rather than performance.



[Click Here to order](#)



### MEASURING PERFORMANCE: CHECKING YOUR ORGANIZATION'S PROGRESS

- Do people understand what performance measurement is and how it is used?
- Do people like the process that is being used to develop performance measurement?
- Do people believe that the measures used are valid?
- Does everyone in the organization receive performance measurement reports?
- Is the measurement data reliable and accurate?
- Do people think the report format is user-friendly?
- Is staff using the performance data to make management decisions?
- Is the data being used in budgeting?
- Is the data being used for determining the objectives of contracts and grants?
- Is the data being used to reward high performance by work groups?

—Adapted from the Oregon Department of Transportation

### RESOURCES ON PERFORMANCE MEASUREMENT



David N. Ammons. *Municipal Benchmarks: Assessing Local Performance and Establishing Community Standards*. Thousand Oaks, Calif.: Sage Publications, 1996. A comprehensive guide to measurement in local government, with thousands of specific measures used by cities.



*Balancing Measures: Best Practices in Performance Management*. Washington, D.C.: National Partnership for Reinventing Government (NPR), 1999. Available at [www.npr.gov/library/papers/bkgrd/balmeasure.html](http://www.npr.gov/library/papers/bkgrd/balmeasure.html), this is a comprehensive guide to creating a well-balanced set of performance measures, complete with dozens of case studies, resources, and links to useful Web pages.

Robert D. Behn. *Bottom-Line Government*. Durham, N.C.: The Governors Center at Duke University, 1994. Telephone: 919-613-7374. A thoughtful monograph on why governments need to measure their “bottom lines” and on the many challenges they encounter in doing so.

Jack A. Brizius and Michael D. Campbell. *Getting Results: A Guide for Government Accountability*. Washington, D.C.: Council of Governors' Policy Advisors, 1991. A basic guide to performance measurement at the state level, available from the National Governors Association in Washington, D.C. Phone: (202) 624-5300.



[Click Here to order](#)

*City of Portland Service Efforts and Accomplishments*. Portland, Ore.: Office of the City Auditor, published annually. The city began issuing these performance reports, described on p. III/160, in 1992. They are available at [www.ci.portland.or.us/auditor/pdxaudit.htm](http://www.ci.portland.or.us/auditor/pdxaudit.htm). Telephone: (503) 823-4005.

The Governmental Accounting Standards Board Web site: [www.gasb.org](http://www.gasb.org). An excellent and comprehensive collection of information on performance measurement in government.

Harry P. Hatry et al. *How Effective Are Your Community Services? Procedures for Measuring Their Quality*. 2nd ed. Washington, D.C.: Urban Institute and International City and County Management Association, 1992. An excellent how-to guidebook on measuring local government performance, chock full of specifics.

Harry P. Hatry et al. *Service Efforts and Accomplishments Reporting: Its Time Has Come, An Overview*. Norwalk, Conn.: Government Accounting Standards Board, 1990. An overview of performance measurement in many different services, at different levels of government.

The ICMA Center for Performance Measurement ([www.icma.org/performance](http://www.icma.org/performance)) offers a comprehensive program in comparative performance measurement for cities and counties. Components include on-site training for new participants, statistical data cleaning, and a private Web site that facilitates discussion among participating jurisdictions about best practices and innovative management techniques. The program brings together 120 jurisdictions in the U.S. and Canada (plus two pilots in Australia and New Zealand) to analyze performance in 15 different service areas. ICMA publications related to performance measurement, including an annual volume of comparative data on these services (*Comparative Performance Measurement: FY 1998 Data Report*) and *Accountability for Performance: Measurement and Monitoring in Local Government*, are available at [www.bookstore.icma.org](http://www.bookstore.icma.org) (type *performance measurement* in the search engine) or by calling (800) 745-8780.

Robert S. Kaplan and David P. Norton. *The Balanced Scorecard: Translating Strategy into Action*. Boston: Harvard Business School Press, 1996. A comprehensive introduction to the concept of balanced scorecards, based on business examples.

United Kingdom. Her Majesty's Treasury. *Executive Agencies: A Guide to Setting Targets and Measuring Performance*. London: Her Majesty's Stationery Office, 1992. An excellent nuts-and-bolts guide, which can be ordered from the Stationery Office at [www.itsofficial.net](http://www.itsofficial.net) (use the search function to search



[Click Here to order](#)

for the title).

Jonathan Walters. *Measuring Up: Governing's Guide to Performance Measurement for Geniuses (and Other Public Managers)*. Washington, D.C.: Governing Books, 1998. An entertaining and enlightening guide to the challenges and pitfalls of performance measurement, published by *Governing* magazine. To order call (800) 638-1710.



[Click Here to order](#)

From *The Reinventor's Fieldbook*, by David Osborne and Peter Plastrik. ©2000 by Osborne and Plastrik

# Notes

All quotations that are not attributed in the text or in these endnotes are from interviews with the authors or their associates. Only in cases where there might be some confusion about the source of a quotation have we indicated in a note that it came from an interview.

## Chapter Twelve

- P. III/140: Goethe quotation: Quoted in Roger J. Vaughan, “Is It Working? Measuring the Performance of Public Programs,” *The Entrepreneurial Economy Review* (July-August 1989). Published by the Corporation for Enterprise Development, Washington, D.C.
- P. III/141: Behn quotation: Robert Behn, *Bottom-Line Government* (Durham, N.C.: The Governors Center at Duke University, 1994), p. 12.
- P. III/142: “In early 1994, twenty-one federal agencies . . . Management and Budget”, *Toward Useful Performance Measurement: Lessons Learned from Initial Pilot Performance Plans Prepared Under the Government Performance and Results Act* (Washington, D.C.: National Academy of Public Administration, November 1994).
- P. III/142: “When they submitted their plans . . . trying to understand it”: Ibid.
- P. III/142: Kettl quotation: Donald F. Kettl, “Measuring Performance When There Is No Bottom Line” (paper prepared for a conference of the New Zealand Politics Research Group, Victoria University of Wellington, July 8, 1994).
- P. III/146: British Treasury’s quality measures (in box): United Kingdom, Her Majesty’s Treasury: *Executive Agencies: A Guide to Setting Targets and Measuring Performance* (London: Her Majesty’s Stationery Office, 1992), pp. 13–16.
- P. III/152: “Most of the U.S. GPRA . . . their missions”: *Toward Useful Performance Measurement*.
- P. III/152: “In Australia, a review of performance reporting . . . little data on outcomes”: Task Force on Management Improvement, *The Australian Public Service Reformed: An Evaluation of a Decade of Management Reform* (Canberra: Australian Government Printing Service, 1992), pp. 351–359.
- P. III/152: “The same thing happened in the U.K.”: See, for instance, Sylvie Trosa, *Next Steps: Moving On* (London: Office of Public Service and Science, February 1994).
- P. III/154: Atlantic Richfield officials quotation: Project USA, *Performance Measurement: Report on a Survey of Private Sector Performance Measures* (Washington, D.C.: U.S. Department of the Treasury, January 1993), p. D-1.
- P. III/154: “Based on the experience of executive agencies . . . any particular unit”: United Kingdom, Next Steps Team, *The Strategic Management of Agencies: Models for Management* (London: Her Majesty’s Stationery Office, 1995), p. 40.
- P. III/154: British Treasury quotation: United Kingdom, Her Majesty’s Treasury, *Executive Agencies: A Guide to Setting Targets and Measuring Performance*, p. 34.



[Click Here to order](#)

- P. III/155: U.S. Treasury quotation: Financial Management Service, *Performance Measurement Guide* (Washington, D.C.: U.S. Department of the Treasury, November 1993), p. 32.
- P. III/155: Sylvie Trosa quotation: Trosa, *Next Steps: Moving On*, p. 13.
- P. III/157: Boeing quotation: Project USA, *Performance Measurement: Report on a Survey of Private Sector Performance Measures*, p. D-1.
- P. III/159: NAPA quotation: *Toward Useful Performance Measurement*.
- P. III/160: Portland annual report: Interviews with Audit Director Richard Tracy.
- P. III/162: “Checking Your Organization’s Progress” box: Adapted from Craig Holt, “Review of the Performance Measurement Program Strategy,” internal memo, Oregon Department of Transportation, March 31, 1995.



[Click Here to order](#)