

The upper left quadrant measures typically include the number of customers and activities. These can be broken out by customer characteristic or type of activity.

The upper right quadrant measures include a set of Common Measures that apply to many different programs, such as turnover rate and unit cost. There are also "Activity Specific" measures in this quadrant. For each activity in the upper left quadrant, there are one or more measures in the upper right quadrant that tell how well that particular activity is performed. For example, if the activity is the number of telephone calls answered, then the activity specific measure might be the percent answered on the 1st or 2nd ring. If the activity is the number of hospital discharge plans produced, then the activity specific measure might be the percent produced on time and with family participation.

The lower quadrant measures almost always come in pairs of number and percentage. For example, the number and percent who graduated high school or the number and percent who got off of alcohol and drugs. The lower quadrant measures almost always have to do with one of four dimensions of better-offness: skills / knowledge, attitude / opinion, behavior and circumstance. Think of these categories as rocks to look under when trying to find performance measures.

Skills and knowledge measures include achievement test scores for students, and percent of participants in parent training who show improved skills.

Attitude and opinion measures include percent of students with high personal ambitions and percent of customers who believe the service helped them with their problems.

Behavior measures include percent of students with good school attendance and percent of public housing residents who pay their rent on time.

Circumstance (or condition) measures include percent of customers who are employed in jobs above the minimum wage, percent of customers in stable housing and percent of road-miles in good condition

All programs have multiple measures in the lower right quadrant. Some of these measures will be more important than others. Try to identify as many measures as possible before winnowing them down to the most important measures. The method described in Appendix G can help you identify measures and winnow them down to the 3 to 5 most important headline measures and a Data Development Agenda.

Using performance measures to improve performance: the 7 Questions

Now that we have established the three types of performance measures, we are ready to address the talk to action thinking process for Performance Accountability. This is very similar to the population process presented in Chapter 3, with two important differences. Performance Accountability begins with the program's customers, while Population Accountability begins with a population in a geographic area. And the method

used to identify performance measures is different from the method used to establish results and indicators.

The full Performance Accountability talk to action thinking process is shown as a schematic in Figure 4.17 and in the form of 7 Questions in Figure 4.18. The 4-Quadrant method for identifying performance measures, presented in Appendix G, can help you get a beginning set of answers to questions 1, 2 and 3 and a head start on question 4.

1. Who are our customers? Customers are the people whose lives are affected, for better or worse, by the actions of the program. Most programs have more than one customer group, and sometimes these customer groups have competing or even conflicting interests. My favorite example of this comes from the Alaska Fish and Game Commission where their customers include the fishermen and the fish. You will never find two customer groups with less in common. They resolved this conflict by choosing "sustainable yield" as one of their most important lower right quadrant measures, a measure that showed how well they balanced the competing interests of their different customer groups.

I encountered another interesting debate about customers in a state department of education. One group of staff argued that the state department had no direct contact with students and therefore the customers must be the superintendents, principals and teachers with whom they worked. A second group of staff argued that the principle purpose of the department was student achievement and therefore the department's customers must be the students. The correct answer is both. The lives of both students and teachers are affected by the actions of the department of education. The issue was settled by calling superintendents, principals, and teachers direct customers and students indirect customers.

Your program or agency will also have multiple customer groups and you need to develop a complete list of who these groups are. If it is helpful you can distinguish between direct and indirect customers, primary and secondary customers, or internal and external customers.

There are several other important ways in which customers can be defined. There are times when organizations are customers. A United Way providing support to its grantees can consider the grantee organizations themselves as customers. This allows us to look at lower right quadrant measures about the behavior or achievement of these organizations. For example, the percentage of organizations that are in good financial condition or the percentage that have identified and are using performance measures in a continuous improvement process.

For services that build and maintain infrastructure, such as departments of transportation or water and sewer divisions, the roads, bridges and pipelines can be treated as customers for purposes of identifying performance measures. This allows us to assess the performance of these organizations based on the condition of the infrastructure system components. For example the percentage of roads and bridges rated in good condition, or the rate of water main breaks.

2. How can we measure if our customers are better off? These are the lower right quadrant *Is anyone better off?* measures. If your program does a really good job, how are your customers' lives better? Conversely, if you did a really bad job, how would their lives be worse? Think about the most meaningful measures, whether you have data or not. Remember not to exclude measures because you don't fully control performance. Measures in this category will always be those for which you have the least control.

3. How can we measure if we're delivering services well? These are the upper right quadrant measures that tell if the program's activities are performed well. These measures are usually about what staff do, and how well the functions of the program are performed. Think about the most meaningful measures whether you have data or not.

4. How are we doing on the most important of these measures? There are two parts to this question. From the measures identified in questions 2 and 3 about customers and service delivery, what are the 3 to 5 most important headline measures. Try to get a healthy mix of measures from each category.

For each of these measures, create a baseline that shows the history of performance and a forecast of where you are heading if you don't do anything more or different. If the data is not handy, then make your best guess about what this graph would look like. Tell the story behind this picture. Why are things getting better or worse? What are the causes and forces at work that explain the picture of your performance? Take credit for what's working and be honest about what's not working.

Identify where you need new and better data and create a Data Development Agenda. Identify questions you need answered in order to fully understand the story behind the baseline. Put these questions into an Information and Research Agenda about causes.

5. Who are the partners who have a role to play in doing better? Consider partners inside and outside the organization who might be able to help improve performance. Push beyond the usual suspects. Think about long-shots here. Make sure you include young people, parents and other customers. It is sometimes useful to distinguish between active and non-active partners, so that you can work to engage the non-active partners as part of your action plan.

6. What works to do better, including no-cost and low-cost ideas? There are two natural pointers to what works. Each cause points to actions you could take to address that cause, and each partner has something to contribute. Consider what the research says about what works. Be creative and think of non-conventional solutions. Insist that people think about no-cost low-cost actions. If you need more information about what works, make these questions Part 2 of your Information and Research Agenda.

7. What do we propose to do? This is the most important question. It is the action part of getting from talk to action. You can use the criteria *specificity, leverage, values and reach* (see page 44) to choose the most powerful actions from the possibilities identified in ques-

tion 6 about what could work to do better. Organize these actions into a plan that specifies the person responsible for each task, the start and end dates and necessary resources.

These questions are designed to be answered quickly and then repeated. It is possible to take a pass at all 7 questions in anywhere from 20 minutes to an hour.

Use the 7 Questions in every supervisory relationship

The 7 questions should be asked and answered on a regular basis in every supervisory relationship throughout the system. A supervisor should have regular conferences with subordinates. Imagine that you handed out the 7 Questions a few days before your next supervisory conference. You say, "We're going to take the first 15 minutes of our time together to talk about your best answers to these questions. I don't expect you to get it perfect. Just take a crack at it. Then we're going to have that discussion every time we get together for the rest of your life." If you were to do this, you would begin to change the culture of your organization.

Performance Accountability is about culture change. One of the things we've learned from Chaos and Complexity Theory is that culture change doesn't happen in complex organizations through big events, like one or two training sessions. Culture change happens with little events that happen thousands of times in the life of an organization. Asking and answering the 7 Questions at every intersection of supervision is a little event that could happen thousands of times in the life of your organization (from the executive to the front line) that would profoundly change how people think about customers (Question 1), how they think about data (Questions 2,3, and 4) and how they think about using data to improve performance (Questions 5, 6, and 7).

The idea of using the 7 Questions in every supervisory relationship is important for another reason. It means that each supervisory relationship is the unit of change for implementing Performance Accountability. If there are 100 supervisory relationships in the organization, then each relationship represents 1% of full implementation.

Using the 7 questions is not rocket science. Managers and supervisors can be trained to do this quickly and will get good at it through practice. This presents a large number of options regarding how to phase in Performance Accountability. It is not necessary to direct everyone to use the 7 questions all at once. It is not necessary to run huge training sessions and hire lots of consultants to hold supervisors' hands. Nor is it necessary to institute a new burdensome regimen of forms to be filled out. **You can simply start using this method, right now, one supervisor at a time.**

Get to the point planning

Notice that the 7 Questions skip past mission, vision, values, purpose, problems, issues, goals, objectives, needs assessments and environmental scans and go directly to performance measures. This goes against the orthodoxy of the planning profession

which dictates that some combination of these other processes must be completed before getting to measurement. Baloney.

Many people seem to think that the agency's mission statement must be perfected before anything else can be done. I have seen this process take six months or more. It may be a good idea for an agency to have a mission statement, but it is unnecessarily time consuming and burdensome to try to develop performance measures from these statements. Unless you are thinking of creating a brand new agency, most programs have established purposes and can begin using performance measures immediately.

Think about it this way. Programs are created through a population level decision process that determines whether the program should **exist or not**. Performance Accountability picks up at this point, takes the program's existence as a given, and moves to the next step of answering whether it's **working or not**. In this construct, it can be argued that **mission statements are by-products of selecting performance measures**, not the other way around. After the most important lower right quadrant performance measures are identified, they can be used to help craft a mission statement. For example, staff in a job training program will readily identify the following two performance measures: the percent of people placed in good paying jobs, and the percent who keep those jobs at least six months. This could then help them craft the following mission statement: "Our mission is to help people get and keep good paying jobs."

Traditional planning systems waste an inordinate amount of time before people are allowed to talk about performance measures. By going straight to the business of selecting performance measures, we ease the frustration and cynicism that goes with complex planning processes. We go directly to the heart of Performance Accountability, **the disciplined use of data in day-to-day management**.

Monthly or quarterly review of performance

Managers should review performance on a monthly or quarterly basis in staff meetings and individual supervisory conferences. There should be three or four charts on the wall for the most important measures. In supervisory conferences, the charts should be on a single piece of paper.

The agenda should track the seven Performance Accountability questions:

1. Customers: Start each meeting by remembering who the customers are. It's amazing how quickly this can get confused or forgotten. The list of customers could also be a chart on the wall.

2 and 3. Choice of measures: How can we measure if our customers are better off? How can we measure if we're delivering service well? Revisit your previous answers to these questions. Consider the status of your Data Development Agenda here. Choice of measures is very important in the early stages of doing this work and can be addressed less frequently as the work proceeds.

Performance Accountability

Getting from Talk to Action

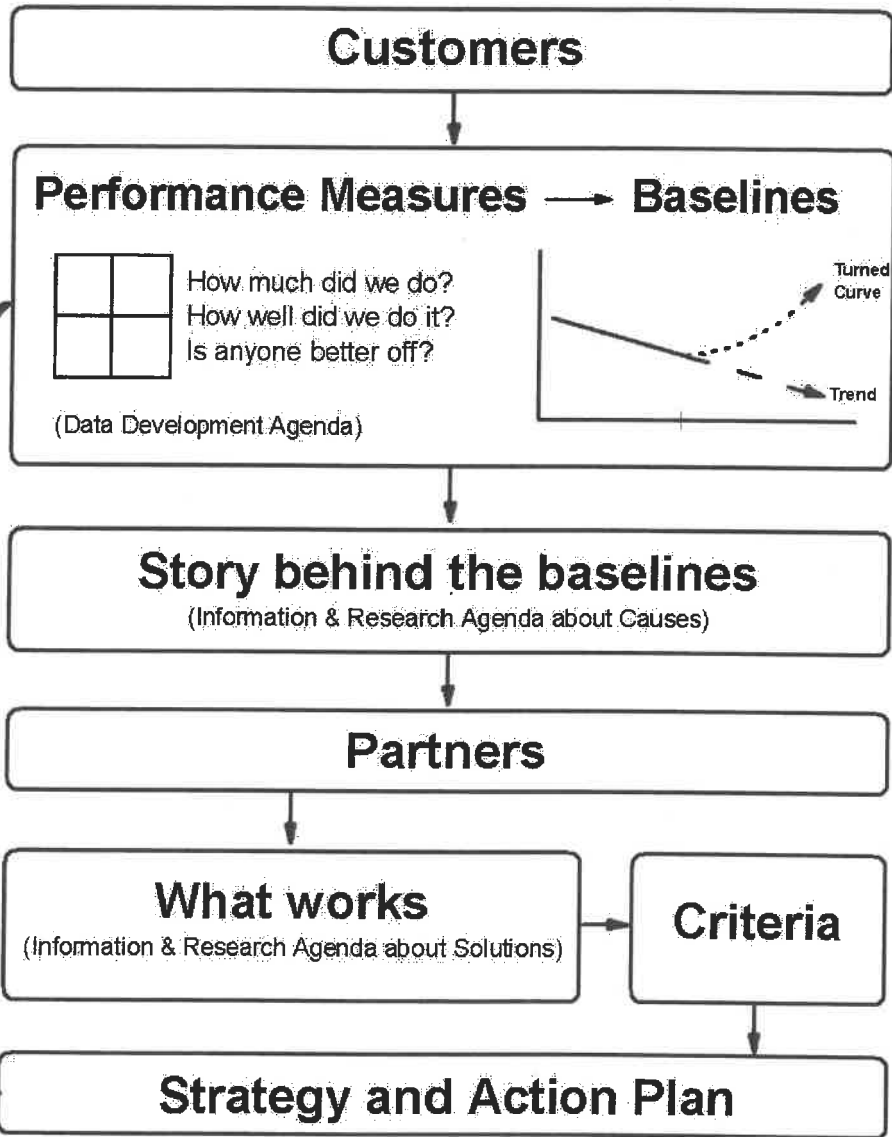


Figure 4.17

The 7 Performance Accountability Questions

1. Who are our customers?
2. How can we measure if our customers are better off?
3. How can we measure if we are delivering services well ?
4. How are we doing on the most important of these measures?
5. Who are the partners that have a role to play in doing better?
6. What works to do better, including no-cost and low-cost ideas?
7. What do we propose to do?

Figure 4.18

4. Current performance: Evaluate performance on the most important measures. Discuss the story behind current performance and progress since the last meeting. Do not go around the room and have everyone report on what they did since the last meeting. Instead use this time for a diagnostic discussion. You have to understand what is causing current performance before you can do better.

5. Partners: Evaluate how successful you have been in engaging new partners and what they contribute. Identify partners that need to be recruited or engaged.

6. What works to do better: Consider what could work to do better. Be creative about possibilities. Make sure you consider no-cost and low-cost actions. Encourage people to suggest "off the wall" ideas.

7. Action plan: Decide what you will do before the next meeting to improve performance. Regarding written monthly or quarterly performance reports - keep them very short, one page of bullet points, or don't do them at all. In all of this work, form should follow function. **Don't create paper if you don't need it.** It's OK to simply keep notes on what people agreed to do. Try these methods without any paper for a while and then decide what form of written material would be useful. Remember one of the cardinal rules of good management: IF IT'S NOT USEFUL, DON'T DO IT.

Comparing performance

There are three ways to compare performance: **to yourself, to others and to standards.** The first and most important comparison is to yourself, using baselines that show the history and forecast of your performance.

Comparing to yourself is a common sense approach that is used all the time in the business world. If we were attending the annual sales meeting of a corporation there would be a baseline chart at the front of the room. It would show the history of sales and where sales were heading if we just continued our current growth rate. The company president would push all of us to do better than the baseline in the coming months.

Child support collections are known to increase in proportion to the growth in the population and the growth in wages. It is possible to forecast the extent to which these forces will drive up collections. Let's say that this forecasted increase is 2% for the next year. If the manager of the child support enforcement program achieves a 2% increase in collections next year, that manager has not actually improved the program's performance. The manager must do better than 2% to show real performance improvement.

All managers should have these kind of baseline charts on their walls or in their desk drawers. When I worked for the state of Maryland, the State Budget Secretary, Lou Stettler, had 15 of these charts arrayed around his office. He could look up at any time and see the baselines for the most important measures of the state's fiscal health.

Comparing to others: When you compare the performance of a group of programs providing the same service, you usually get the majority of programs clustered in the mid range of performance, with a few performing much better and a few performing much worse. What happens when you reward the high performers and punish the low performers? Before you answer this question, you have to know **why** one group is doing well and the other group is doing poorly.

The high performing group may be doing well because they have all the easy cases and the low performing group may have all the tough cases. If you reward one and punish the other, you send a simple message throughout the service system: "Skim the easy cases for yourself. Dump the tough cases on someone else." We see examples of this skimming all the time. There was a managed health care provider in Florida that was enrolling elders in their health care plan. They put their enrollment office on the third floor and made sure the elevator was always broken. If you could walk up three flights of stairs they wanted you in their plan. It was a brilliant scheme to skim the elder population for people healthy enough to walk up three flights of stairs. Illegal but brilliant.

Comparing to standards: Finally we can compare to standards. There are many well-established standards for how well service is delivered. We have standards for processing applications in 30 days. We have standards for staffing ratios in child care centers and nursing homes. We have standards for handicapped accessibility. There are many useful standards about how service is delivered because there are clear cases where we know what good service looks like.

Standards for measures of whether customers are better off are much more difficult to establish. It's the same problem we encounter when comparing programs to each other. The mixture of easy and hard cases makes it difficult to determine a fair standard for all providers. If you have 15 different welfare to work providers all with different proportions of easy and hard cases, it is not fair to set a single standard (e.g. 70% placement in jobs) for all of them. It will be easy for some to meet this standard and impossible for others. The concept of turning the curve is a much better way to gauge progress on lower right quadrant measures than whether that program has met a fixed standard.

When standards are used to stretch organizations toward higher performance, this should never be done without also using baselines to give credit for incremental improvement. Failing to do this can lead to dysfunctional behavior. A hospital was being judged against standards for in-patient death rates by type of operation. They figured out how to beat the system. If a patient was about to die, they were discharged so it wouldn't show up on their statistics. This hospital looked great on inpatient death rates. It's not a hospital you would want to go to.⁸⁷

87. This story is cited in *Governing Magazine*/May 1998 management column entitled "Poisoned Measures." The magazine admits that it may be apocryphal.

This illustrates another important point. If you manage your program using only one measure, you can look good on that measure at the expense of something else. Maybe the hospital should have tracked a second measure about the well-being of patients 6 weeks after discharge. If the discharged patients were all dead, this would hint at a problem. The hospital's goal should have been to optimize performance simultaneously on in-patient death rates **and** post-discharge recovery.

The danger of perverse incentives can be found in just about any system. If you're a high school principal and want to improve the school's attendance rate, then kick out all the troublemakers. The attendance rate will go up. If you are a social services director and want to bring down foster care caseloads, slow down child abuse investigations. The foster care caseload will begin to decline.

We are not trying to find one magic measure for a program, but rather a **balanced set of three to five measures** taken from *How well did we do it?* and *Is anyone better off?* These can be used to give a "checks and balances" view of program performance. Managers can work toward improving performance on all of them simultaneously.

Using baselines to set targets without creating fear of punishment

When people hear the words "performance measurement" it triggers an association with all the bad things that have ever happened to them since childhood. In almost everyone's experience, the matter of measurement has been connected at one time or another to punishment. This goes back to our earliest memories of school grades and continues right into the workplace. It is no wonder that there is an atmosphere of fear surrounding this work.

Setting targets has been especially problematic. There is a long history of setting unrealistic performance targets and then beating people up when they fail to meet them. This practice is based on the misguided belief that such targets coupled with fear of punishment will lead to better performance. In practice, unrealistic targets detract from credibility. And fear turns out to be one of the worst ways for managers to motivate people. People working within this punishment culture will try to pick measures they look good on or set targets they can easily meet. The measures are rarely the most important ones and the targets are meaningless. The organization appears to be practicing Performance Accountability, but it's a waste of time. There is a better way.

We must first acknowledge that getting people to use data to improve performance is a culture change. Disciplined use of data is a new experience for many managers. In developing a Performance Accountability system, data should first be used in private consultations between employees and their supervisors. This can provide an incubation period, where managers get familiar with how to interpret data and use data to improve performance.⁸⁸

^{88.} See the discussion of "Inside story - outside story" below.

Performance Accountability should be part of a larger organizational development effort to create a healthy workplace. You can have the best designed accountability system in the world and it will not work in a sick environment. The purpose of Performance Accountability must be improved performance. Recognition should be provided for improved performance as well as outstanding performance.

Another way to avoid the punishment culture is to set targets in a fair and useful way. A target is a specific desired future level of achievement for an indicator or performance measure. Targets are sometimes time specific and sometimes not, "The target is a 95% graduation rate," or "The target is a 95% graduation rate by 2020." A standard is a type of target that has an established relationship to quality service or quality of life conditions. Standards can be minimum acceptable standards such as "no more than 24 hours to respond to a report of child abuse," or high standards to strive for such as "all invoices paid on time."

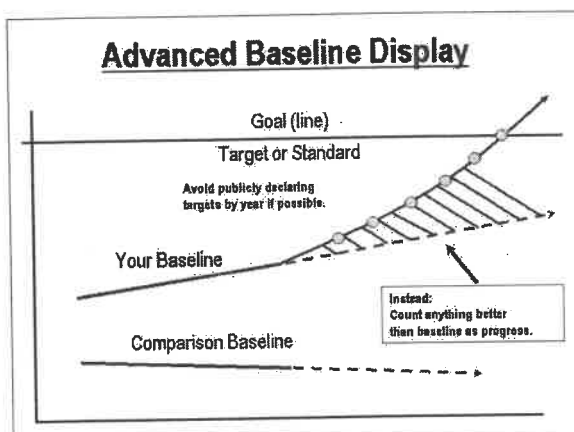


Figure 4.19

Figure 4.19 illustrates the concept of turning the curve. The performance baseline is shown with its history and forecast. A comparison baseline is shown for a similar program. A goal, target or standard is shown as a horizontal line. The idea is to turn the curve and literally cross the goal line as soon as possible. Anything better than baseline counts as progress.

The key to effective use of targets is a sense of fair play. Targets should always be set in relation to a baseline. When a date is set for achieving the target, it should look reasonable when viewed against the baseline history and forecast. There is nothing wrong with ambitious targets provided that they are achievable, and there is a method for recognizing incremental progress short of the target. The concept of turning the curve is such a method.

The schematic in Figure 4.19 shows a way to present the three types of comparisons (to yourself, to others and to standards) on the same graph. The performance baseline is presented with history and forecast. A comparison baseline is shown for a similar program. And a goal, target or standard is shown as a horizontal line. The idea is to turn the curve and literally cross the goal line as soon as possible. Anything better than baseline counts as progress.

This stands in contrast to the way we often set targets. We set a 5 year target and then declare year by year exactly where we are going to be along the pathway of improvement. The problem is that you can do a great job of turning the curve, but fall just a little short of where you said you were going to be, and you're a failure. In a small town in Australia, the newly assigned police officer promised a 50% reduction in crime in the first year. He only achieved a 43% reduction. The headlines were "Failed to meet target."

I recommend that you avoid **publicly** declaring **year by year** targets if possible. The only exceptions are when minimum standards of health and safety are involved. Year by year targets or month by month targets can be useful **internally** to motivate and stretch employee performance. But once you declare them publicly, you risk setting yourself up for failure.

This baseline format should become the standard way in which you present all of your most important data, whether your funders ask for it or not. Fill out all the forms that the funder requires, but provide this presentation as well. It allows you to claim credit for progress even if you do not reach some predetermined target. And it allows you to challenge targets that are unreasonable or unfair.

One of the best examples of this format is Healthy People 2010, published by the US Department of Health and Human Services. For each of the health indicators presented in this document, an historical baseline is presented, and a horizontal line is drawn representing the national goal for that indicator for 2010. It is easy to see if the nation is making progress toward this goal, and judge the chances of achieving it at the current pace of change.

Inside story - Outside story

Closely related to the matter of setting targets is the matter of how and when organizational performance is thrust into the public view. This is not always a matter of choice. School performance data is now routinely published in the newspapers whether the schools are ready for it or not. And performance that concerns unethical or illegal behavior should never be withheld or hidden. But, no one wants to look bad in public, and Performance Accountability systems sometimes fail because they raise the stakes too high and too fast.

Programs and agencies need to understand their performance and have a plan to fix it before they explain it publicly. While embarrassment is sometimes unavoidable, you cannot ask managers to support a system that **deliberately** embarrasses them.

Managers will sometimes need to have an "inside story" and an "outside story." The inside story must be the unadulterated truth, a straight look in the mirror about how well the program or agency is performing. When managers understand the truth about their performance, they can figure out what to do about it and how to tell their performance story to the outside world.

The problem arises when the only story a program or agency has is its "outside" story. This means that the manager and senior staff don't even know the truth about their own performance. **They are not just lying to the outside world, they are lying to themselves.** If Performance Accountability is going to work, then complete truthfulness is the only way.

We should require managers to have a completely truthful inside story, and then help them figure out how to tell that story to the outside world. This suggests a way of telling the outside story that is powerful and compelling and yet is rarely used. In a public setting with elected officials, stakeholders, or constituents, the presentation goes like this:

*"Here's where we are doing well.
Here's where we are not doing as well as we would like, and
Here's what we are doing to get better."*

A good offense is the best defense. Give them the good news first. Take credit for what you have accomplished. Then give them the bad news before they force it out of you. Then tell them what you plan to do to get better. You might even ask for their help in funding the necessary improvements.

This is scary and somewhat risky. But this approach gets you two things. It gets you credit for what you're doing well. More importantly it gets you **credibility** for telling the truth even when it's difficult. And credibility is the single most precious commodity in management and politics.

Is my program working?

Did you skip ahead to this section? Do you have an urgent presentation to make? If the first purpose of Performance Accountability is to improve performance, then the next purpose is to demonstrate that the program is working.

The question of whether a program is working has two parts. The first part is about process and the second part is about customer results.

Process: Let's take the easier process questions first. "Is my program doing what it's supposed to do?" and "Is the program doing it well?" Most programs are designed to provide certain services in a certain way. Process evaluations can test if the program is operating true to its design. In a treatment program, for example, it is possible to determine if the treatment regimen in practice conforms to the program's design. In a school, it is possible to determine if teachers are in fact teaching the prescribed curriculum.

The next question is whether the program's prescribed functions are being performed well. Here we can use the *How well did we do it?* measures to determine if the functions are timely, accurate and properly staffed. What percent of applications are processed on time? What percent of staff have completed relevant training? Are staff workload ratios within an acceptable range? Problems related to program design most often require changes in policy such the hours the treatment clinic is open. Problems with program implementation most often require changes in practice such as hiring more staff.

Results: The second part of "Is my program working?" is about customer results. This is a much more important and much more difficult question to answer. "Are our cus-

tomers' lives getting better?" and "To what extent did the program cause this change?" What is the high school graduation rate? What percent of 3rd graders can read at grade level? Are job training graduates getting good paying jobs? We can test how we're doing on these measures and determine whether performance is getting better or worse.

There is an underlying question that is always present when considering customer results. Did my program **cause** the customer's improvement? This is the **most** difficult question to answer. There are a few reasonably straightforward methods that can be used to demonstrate that a program has had a positive effect on its customers. But any attempt to discern the extent to which the program caused this improvement runs up against profound questions about the very nature of causality in complex environments. We'll give four straightforward methods in this section. The nature of causality in complex environments will be a completely and definitively addressed in chapter 7.⁸⁹

1. Control group research: is the "gold standard"⁹⁰ for demonstrating cause and effect. Create two groups with comparable characteristics. Provide treatment to one. Deny treatment to the other. See if there's any difference in the customer, student or patient results. These methods are used all the time in medical research. The principal problems that go with control group research are expense and ethics. Control group research is very expensive. Most programs can't afford to do it. There are also legal and ethical restrictions in human research about withholding treatment that is believed to work.

There are some rare circumstances where control groups occur naturally and it is possible to create comparisons with relatively little effort and expense. When Pennsylvania implemented its first family preservation programs, there were not enough slots for all the families who qualified for the service, and a waiting list was created. Since families were not triaged from this list, the families waiting for service were very much like the families receiving service. An analysis showed that there were higher rates of foster care entry from the waiting list families than the families in the program. This suggested that the family preservation program was making a difference on the rate of entry into foster care. Using this data, it was possible to estimate the program's effect on foster care caseload and cost. Such comparisons can also occur between jurisdictions where some jurisdictions adopt a policy and others don't. States with motorcycle helmet laws, for example, show lower accident related medical and rehabilitation costs than those without.

2. Direct measurement: Another way to determine program contribution to customer change is by selecting performance measures that directly compare the customer's condition at two points in time. For example, a tutoring program could compare a student's grades at the beginning of the program with grades later in the school year. Any significant pattern of improvement would represent evidence of program contribution to the change.

^{89.} See "Demonstrating program contribution to community change." in Chapter 7.

^{90.} So called because it costs a lot of gold.

3. Comparisons to similar programs or populations: Other kinds of performance measurement comparisons can also help quantify the extent of program contribution to customer change. You can compare your performance to that of a similar program whose effect on customers has been formally tested. If your program's success rates fall near those of the comparable program, you can piggyback on their documentation to demonstrate your program's effect.

Still another comparison looks at your customers' results compared to relevant population characteristics in the areas from which your customers are drawn. If, for example, the rate of repeat teen pregnancy for the young women in your program is lower than the rate for other teens in their neighborhoods, you can argue that the program contributed to this difference.

These kinds of comparisons are among the most commonly used, but they are replete with difficulties. No two programs are ever exactly the same. Programs with the same names and designs still operate differently and serve different customers. The same problem applies to comparisons with neighborhood population characteristics. The case for comparability between programs or neighborhood populations must be compelling or at least credible to make these methods believable and useful.

4. Customer opinion: Finally, when all else fails, ask the customers if the service made a difference. "Did this training help you become a better teacher?" "Did these counseling sessions help improve your child's behavior at home and at school?" Like all customer satisfaction surveys, the data can be difficult to interpret and may be viewed with suspicion. People in positions of authority tend not to believe self reporting of benefits. But customers are the ultimate judge of a program's success and their opinions, compiled as data or given as stories, can be useful evidence that the program is working.

The performance of service systems

A service system is a set of programs or agencies with common customers and related purposes. The education service system includes pre-school child care, primary and secondary education, adult education programs, and colleges and universities. The health care system includes public and private sector doctors, hospitals, clinics, rehabilitation services, long term care, and support services like laboratories and ambulances. The out of home care system for children includes public child welfare services, juvenile justice, and mental health services, and residential and therapeutic providers supporting children in care. Subsets of these service systems can also constitute a service system. For example, the higher education system is often considered separately from the education system as a whole.

Performance measures can be developed for service systems. The most important measures can usually be derived by adding together comparable measures from the systems' component parts. In health care, it is possible to calculate a rate of emergency room usage for the system as a whole by adding together the usage data for individ-

ual hospitals. In higher education, the percent of freshman who reach graduation can be calculated by adding together the data for individual institutions.

There are some measures, however, that are only meaningful at the system level where more than one agency is involved. These include such measures as the average number of case workers assigned to the same family, or the average number of different offices that clients must visit each month to get the services they need. These kind of systemic measures cannot be produced by adding agency data but must be separately gathered across the system, usually on a sample basis.

These kind of cross-system measures give an operational definition to otherwise vague notions like coordinating services or service integration. If services are truly coordinated, then customers should have fewer case managers to deal with and fewer offices to visit. Managers of a service system can join together to use these measures to consider the system performance as a whole, and work to improve that performance.

But we already have so many reporting requirements imposed on us

Yes, everyone does. You spend so much time filling out forms for other people that you have no performance system for yourself at all. The sum total of your funders' reporting requirements should not be your performance system. Instead, you should create a performance system that makes sense to you and your staff, and it should actually be **useful** in running the program.

If you do this well, it will provide most of the information you need to meet your funders' reporting requirements. Meeting these requirements will become a **byproduct** of your system. Furthermore, you will be more successful with funders because your program performs better. Think about using this approach when designing the management information systems for your organization. Put your management needs first and allow the funders' forms to follow.⁹¹

How to implement Performance Accountability

The history of performance measurement goes like this: One morning the chief executive comes charging out of the shower all fired up about performance measurement, gets to the office, declares that starting tomorrow every unit of the agency will have performance measures. Three months later, the organization has generated a thousand pages of paper, most of it completely useless. Eventually the system collapses under its own weight until, several years later, the next executive comes charging out of the shower. This is the life cycle of performance measurement systems. Trying to implement Performance Accountability on a grand scale in a single stroke is almost always a mistake. Implementing Performance Accountability should proceed in stages and is best done both top down and bottom up.

^{91.} Another case of forms following function.

Let's start with the **bottom up**. Find the three or four best managers in your organization at the operational level. Provide support for these people in the form of a coach or mentor. Help them identify performance measures and use the 7 Questions on a regular basis. Don't prescribe forms and reports at first. Let them play with and adapt the process and see what forms emerge. Consider this an experiment where the purpose is to make the products useful to the manager. Post the most important performance baselines on the wall. When these people have a good story to tell, allow them to showcase their work using presentation methods that track the 7 Questions. (See the Turn the Curve report format Appendix E.) In other words, help them be successful.

Once you have examples of success, then grow the work up through the organization. The managers who do the best work can become coaches for those who come later. For those of your staff who are resistant or just don't get it, assign a coach, send them to training or send them to a desert island. Let people see that performance measures actually get used. Give the success stories to your public relations staff to use in newsletters and press releases.

Equally important is the work from the **top down**. If people in the organization are going to take the work seriously, then senior staff must model this behavior. The agency director should ask each of the people reporting directly to her or him to bring their best answers to the 7 Questions to their next one-on-one conference. In the first few meetings, focus on getting the right performance measures and establishing baselines. After this, focus on the questions about the story behind the baselines and what will be done to get better.

Make this discussion the **first** 15 minutes of every meeting, not something you get to if you have time. After a few sessions, the performance measures could become the basis for a short performance report that is used for setting organizational or personal goals. As senior managers prepare for their meeting with the boss, they will find they need to ask the same questions of the people who report to them. This begins the process of building down through the organization. After a few months, the second tier of managers can be required to use this same process with their staff.

This process will lead to consensus about what measures are most important for the organization and allow baselines for these measures to be updated on a regular basis. A few of these baselines should be posted on the wall in the conference room where staff meetings are held. Some might be permanently on display. Others might rotate through the different component parts of the organization. At the monthly staff meetings, staff can take turns making short 15-minute presentations about what they do by answering the 7 Questions and making reference to their baseline performance charts.

This approach sends two very important messages. First, performance measurement is taken seriously by management, and top managers are willing to practice what they preach. Secondly, and perhaps more importantly, performance measurement is part of day-to-day management, not a back-burner exercise.

Begin using the performance measures and analysis in the organization's internal budget deliberations. Using performance measures in the budget process makes another statement about its importance. Use the 7 Questions to structure internal budget deliberations. Keep it simple.⁹² Don't let the budget process add unnecessary complexity to the forms or process and don't let it become the primary reason for doing this work.

If you are a mid-level manager and work in an organization where you are the only one who seems to understand this approach, try it quietly for your part of the organization. Become an island of excellence. Take credit for the improvements in performance you are able to achieve. Begin presenting your work to others using RBA Performance Accountability methods. When top management notices, **and they will**, help them understand how this could be used more broadly in the organization.

⁹² Most people don't know the second half of "Keep it simple, stupid." (KISS). It's "Or face failure." (OFF).

WHERE RESULTS-BASED ACCOUNTABILITY THINKING HAS WORKED

North Lincolnshire Council, United Kingdom⁹³

The North Lincolnshire Council has responsibility for a wide range of public services for a population of approximately 150,000. One of those services is public housing. The Council directly owns and manages over 10,000 housing units, approximately 15% of the area's housing stock.

In 2001, the national Housing Inspectorate reviewed the North Lincolnshire public sector housing services and rated it as a "poor service." In 2003, the inspectors returned and found that the service had actually gotten worse. The housing service was near the bottom of virtually all of the national performance tables. In 2000, over 3.5% of the stock was empty, compared with the top performing authorities that had a vacancy rate of only 0.75%. The inspectors were so concerned about the overall state of the service that they threatened to take the service out of local control unless the Council could produce a convincing recovery plan.

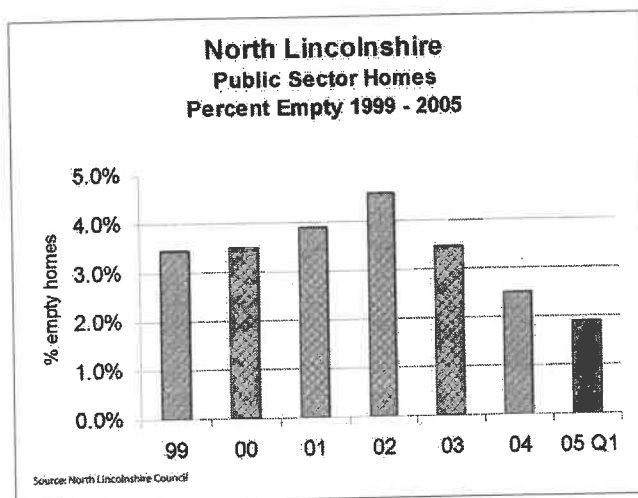


Figure 4.20

These findings were disappointing and embarrassing for the Council. Since the creation of the North Lincolnshire Council in 1996, elected members, managers and staff had worked hard to develop high-quality services. Some of the Council's services, notably social care services and educational services, were judged to be among the best in the country. In the weeks that followed the re-inspection, the social and housing department put together a Recovery and Continuous Improvement Plan. The purpose of this plan was to raise service performance to high levels and develop a capacity for self improvement that would carry the service beyond the life of the plan.

⁹³. Thanks to Nigel Richardson, Executive Director, Social and Housing Services and Mike Pinnock Head of Learning, Development and Support, and their colleagues at the North Lincolnshire Council.

Think outcomes: The newly appointed head of social and housing services, and one of his senior staff, drew on previous experience using Outcomes Based Accountability (OBA) in social care services. The new plan involved 16 tracks of improvement work. One of these was an effort to improve the turnaround of empty properties and make them ready for re-occupancy. Managers within the service had traditionally called these empty properties “voids.” The director insisted that staff start talking about “empty homes” to link the effort to the wider community outcome of helping people create decent and affordable homes.

Process makes perfect: A team of staff and tenants completely revamped the process of preparing vacant properties for re-letting to assure quality and also reduce turnaround times. They began classifying all properties according to the amount of work needed, creating a fast track for properties that required minimal work. Each property was assessed against a set of quality standards by tenant inspectors. Two tenant “champions” were nominated to lead the implementation of the work.

If you can't see it you can't manage it: In consultation with the operational managers, the information system team developed a web-based reporting system that produced a set of “desert island reports,” reports that the operational managers could not imagine living without. Each manager got a set of real-time reports that helped them manage their work and assess their progress. At the start of the work, managers didn't know how many empty homes they had. The new system told them the number of empty homes, the type of property, and when each property was going to be ready for inspection and occupancy. The reports helped operational staff anticipate problems and manage proactively.

The power of positive feedback: As the work progressed, senior managers and elected officials publicly praised and encouraged the staff and tenants involved. Press releases were issued and reports presented to the Council's cabinet meeting. The progress was made visible to residents and other stakeholders through a widely distributed monthly progress report. For many housing services staff, it was the first time they had heard anything good said about their work.

At the start of the improvement program, there were usually about 500 empty properties on any one day. In the first quarter of 2005, there were less than 200. This dramatic reduction has had a number of positive effects:

- 300 extra families have decent, affordable homes;
- 300 sets of neighbors are not living in fear of the crimes that empty properties attract;
- Neighborhoods are less vulnerable to the danger of market collapse that can occur when empty properties begin to accumulate in areas of low demand; and
- There are 300 extra streams of rent income which can be re-invested in housing stock improvements such as energy efficiency.

The success of this work gave staff the confidence to tackle other areas of weakness within the service. It began to restore the confidence of the tenants in the Council's competence as their landlord.

Two years after being judged a failing service, the Housing Inspectorate returned to review progress. Their judgment was that the service was a "fair service with excellent prospects." The report noted the role of strong leadership, a supportive relationship to staff, and a changed organizational culture as contributing to the change. Performance Accountability methods combined with strong leadership turned around a critical service.