Chapter 3:

POPULATION ACCOUNTABILITY

for the well-being of whole populations in a community, city, county, state or nation

In this chapter, we address how partners can work together to improve results for people in communities, cities, counties, states and nations.

Getting from talk to action

Let's start by taking a quick spin through the steps of the population accountability process. Then we'll go back and look at some of the key concepts in detail. The full thinking process is shown as a schematic (Figure 3.2) and in the form of 7 common sense questions (Figure 3.3).

Population: The population accountability process starts, as you might imagine, with identifying a population. Populations can be any whole population or subpopulation in a geographic area. The geographic area can be anything from the entire world to a nation, state, region, county, city or neighborhood. Within these geographic areas, the population can vary from all residents to any identifiable subpopulation. So we might talk about all residents of California, all elders in Miami, all overweight people in South Beach. The only populations not addressed are service populations, actual recipients of a particular service, which is the subject of performance accountability in the next chapter.³⁶



Figure 3.1

Results: Next, we ask what results (conditions of well-being stated in plain language) we want for this population? In the leaking roof example, the implied result was "a non-leaking" roof. The California law that created the state's First 5 commissions³⁷ included the result that *Children Enter*

^{36.} Note that target populations, the intended group from which customers are drawn, is part of population and not performance accountability.

^{37.} In 1998, California voters approved Proposition 10 to tax tobacco products and use the proceeds for services for children ages 0 – 5. The California law created a state First 5 Commission, and First 5 commissions for each of California's 58 counties.

School Healthy and Ready to Learn. Many states, counties, cities and communities have established a more complete set of results, addressing such conditions as *Children Born Healthy, Children Succeeding in School, Safe Communities, A Clean Environment, All Residents Living in Safe, Stable and Affordable Housing.* The process can start with one result or many results. The only requirement is that results must be stated in plain language and not be about a program or about data.

Experience: How do we experience the results we want? How do we see, hear or feel them? If we are concerned about children being safe, for example, we might experience this by talking to our children at breakfast about whether they feel safe going to school or coming home from school. We might observe if kids are wearing bike helmets as they ride around the neighborhood. When Jerry Brown was running to be mayor of Oakland, California he said that he would know his neighborhood was safe when he could walk around without seeing so many grates on the windows. These are experiential versions of safety. There is no government report on window grates.

Experience is the bridge between results and indicators.³⁸ It grounds the work in our daily lives. Each experience is a pointer to a potential indicator. If we experience *People in Safe, Stable, Af-fordable Housing* by seeing neighbors who own their own homes, then we can count the percent of people who own their own homes. Experience can also be a temporary substitute for data when there is no data. We can develop action plans designed to improve how we experience these conditions, while we work to create data that will help us track our progress.

Indicators: If we can identify results in plain language and in experiential terms, then we can look for data that tell us if we're getting these results or not. We are looking for 3 to 5 indicators for each result. The section below on Choosing Indicators presents a step by step method for doing this.

Data Development Agenda: An important by-product of selecting indicators is the creation of a Data Development Agenda, a prioritized list of where we need new or improved data. We never have all the data that we need, and we must continually work to get better data. Getting good data becomes more difficult as we move to progressively smaller geographic areas. National and state data is easiest to get. Neighborhood data is hardest.

Baselines: For each indicator we must create a baseline. Baselines have two parts: an historical part that tells where we have been and a forecast part that says where we are headed if we don't do any-thing differently. Baselines allow us to define success as turning the curve away from the baseline or beating the baseline.

Story behind the baselines: What is the story behind the baselines? If juvenile crime is rising in our community, why? If water quality is declining, why? What are the causes and forces at work behind these conditions? In public health they call this step epidemiology. Epidemiologists seek to understand the causes of a disease as pointers toward a possible cure. Similarly, if we understand what is causing crime or poor water quality, this will help us decide what actions to take that will best address these causes. Many processes skip this step and go directly to talking about actions.

^{38.} Thanks to Jolie Bain Pillsbury for helping me understand the importance of using experience as a bridge from results to indicators.

This is like a doctor prescribing medicine before diagnosing the disease. In medicine, the proper sequence is diagnosis before treatment. In Results Accountability, it's story before action.

The word "story" is deliberately chosen here. Telling stories is the oldest form of communication, the oldest form of retained knowledge, and the oldest way in which we transform life experience into useful lessons. The idea of telling stories allows each partner to explain her or his perspective on how we got where we are today. This is the place to take stock of both positive and negative forces, what is working and what is not working. It is common to find many different opinions about causes. But it is <u>not</u> necessary for partners to reach agreement on a single story. Diverse points of view are assets to be respected, not obstacles to be overcome.

Many processes come to a dead halt because they require that all the partners agree with each other before going on to the next step. In Results Accountability, diversity of opinion is the sign of a healthy process. You want as much information as you can get about causes, so that you have lots of choices about actions. For example, when working to reduce teen pregnancy, people from the faith community might argue that teen pregnancy is caused by moral decay. While public health folks might argue that it is caused by lack of access to reproductive health services. This is OK, because when it comes to action, the faith community can preach abstinence and the public health people can hand out condoms. Constructive diversity of opinions about causes leads to a richer action plan with a greater chance of success.

Information and research agenda about causes: Trying to understand the story behind current conditions, will generate a need for more information. An Information and Research Agenda is a disciplined way of pursuing unanswered questions about causes. That agenda can guide information gathering between meetings, and, if resources are available, the actual commissioning of research. We often make incorrect assumptions about causes. For example, a study in Santa Cruz California challenged the view of homeless people as mostly single male alcoholics, and showed the high percent of evicted and displaced families with children.³⁹ Research can force us to question our biases. Research organizations, including local colleges and universities, have a very important role to play in doing this work.

This is also where **needs assessment** fits in. Many frameworks place needs assessment at the beginning of the process. But needs-based decision making processes tend to get mired in the simple fact that we can never meet all the needs. Incrementally meeting a greater percent of need misses the point. We meet needs for a reason and that reason is the improved well-being of children, adults, families and communities. Results and indicators should come first. Needs assessment comes later as a tool to help understand causes and craft solutions. Even in this role, needs assessment has limitations. Needs assessments predictably come to the conclusion that the problem is lack of services and the solution is more services. Results Accountability instead challenges people to think more deeply about causes and to consider both service and non-service solutions that will make a difference.

^{39.} Applied Survey Research, Santa Cruz County Homeless 2000: Census and Needs Assessment, 2000, www.appliedsurveyresearch.org.

Partners: Who are the partners who have a role to play in doing better? As we discussed earlier, no one program or agency can do it alone. The work requires contributions from a wide array of partners, public and private, across the community. In most processes, thinking about and assembling partners is one of the first steps taken, not something here in the middle. (See "Bringing people together to do this work," in Chapter 7). However, the work of adding partners is never finished. At each pass through the decision process, it is important to consider who is still needed at the table. The action plan should always have a component that addresses the recruitment and engagement of new partners. Remember that in practice, you never have everyone at the table. Processes that can't do anything until everyone is at the table, typically don't do anything. **Inclusion is a process not** an end point. One of the tricks used in the Turn the Curve exercises in Appendix E is to ask participants to wear "two hats." In considering both the story and what works, participants are asked to represent both their own point of view and also the viewpoint of a partner not otherwise represented at the table. This enriches the discussion, and also makes the work more fun. It is important to press for consideration of non-traditional partners. When working on children succeeding in school, consider partners outside the education system, like the business community, media or the children's extended family. When working on healthy children, consider partners outside the health system. Think as broadly as possible and consider potential partners based on what they have to contribute, not on the likelihood of getting their contribution. A good set of results and a solid action plan will be a magnet for people and resources. One small neighborhood family support center in Baltimore City listed over 300 people and organizations on its partner list.

Don't forget about fiscal people. Fiscal partners are often left out of the discussions until the very end of the process, when they are then asked to help finance the efforts chosen by the group. At this point it is often difficult to get their cooperation. It is better for you, and for them, to be involved early on.

What works: We now come to the central question: "What works to do better?" There are two natural pointers to answer this question. First, each part of the story behind the curve points to an action. If we know that one of the causes of teen pregnancy is that kids don't have enough to do after school, then this is a pointer to supervised recreation. If one of the reasons for increased fear of crime is poor lighting in the neighborhood, then this is a pointer to improved lighting. The second natural pointer to action comes from the partners list. Each partner and potential partner has something important to contribute to turning the curve. Systematically consider each possible contribution.

Look at the research for what has worked in other places including best and promising practice. There is a growing body of work about programs and services that have worked, and important resources on the web. The National Governor's Association and the Rand Corporation websites provide information about best practices. There is much less research on how complex strategies work to produce change in complex environments. But pathwaystooutcomes.org, a joint project of Harvard University and the Annie E. Casey Foundation, is an excellent resource for considering whole strategies. A more complete list of such resources is given in Appendix J.

Research is important, but it is also important that the thinking of the group not be <u>limited</u> by the research. My friends in the academic community sometimes blanch when I say this. But the research world can only tell us a fraction of what we need to know. We've got to make sure we use our own common sense, our own life experience, and our own knowledge of the communities in which we live. Something that has worked somewhere else might not work so well in your community. There must be room for learning and innovation.

Diversity of opinion about what works is a strength, not a deficit. While some agreement will be needed when it comes to allocating resources, there does not have to be a win / lose fight over every action. Some partners will contribute their own money for their part of the action plan and some solutions will involve no cost or low cost ideas.

No-cost and Low-cost ideas: Not everything is about money. Some of the most important things you can do sometimes require little or no money. This point was driven home to me during a work-shop in El Dorado County, California.. In the afternoon of the workshop, we did a turn the curve exercise,⁴⁰ where people worked in groups of six to turn a curve that they thought was important. That day people wanted to work on the teen suicide rate. In the preceding two years, there had been an alarming number of teen suicides in El Dorado County. The groups went through the whole Results Accountability thinking process in an hour. They came back and put their best ideas up on the walls. As we looked around the room at these ideas, half to two thirds of them were no-cost or low-cost ideas. They were things that could be done the next day without another dime of money, things like distributing a list of the warning signs of suicide to all the high school students and teachers and providing training in how to recognize those signs.

This experience led me to another rule. Any plan that does not have some significant component that is no-cost or low-cost is not complete. People should be sent back to do more thinking. If you give people permission to think about no-cost and low-cost ideas, you will be amazed at what they come up with. Oddly enough, people need permission to think this way. We have been so thoroughly trained to think about more money for more programs, that we are sometimes blind to obvious inexpensive actions that can make a difference.

Information and research agenda about solutions: The second part of the Information and Research Agenda is about what has worked in other places and the success stories where people have tackled similar challenges. Academic partners can be a great resource in gathering this information. Part of the action plan could include fund raising to commission the research that is needed.⁴¹

Strategy and Action Plan: Planning without action is meaningless. If you do a good job of thinking about what works, you will come up with more ideas than you can actually do in a year. You will need a set of criteria to set priorities, create an action plan and budget, and get started.

Priority setting is always based on criteria. Sometimes the criteria are explicit, as in review processes for solicited proposals. Sometimes the criteria are implicit, hidden in the preferences and prejudices of the decision makers, as in budget processes. The best group processes use explicit criteria for making choices. Here are four criteria that could be used to select the most powerful ac-

^{40.} There are two versions of the Turn the Curve exercise, one about population accountability and the other about performance accountability. Both versions are given in the Appendix E.

^{41.} My favorite quote about un-needed research was penned on a doctoral dissertation: "Fills a much needed gap in the literature."

tions.⁴² Note that the rating process provides an opportunity to fix proposals that get low ratings on any of the criteria. A low leverage proposal can be improved to make it more powerful. A proposal that is not specific enough can be made more specific.

Specificity: Is the idea specific enough to be implemented? Can it actually be done? "Everyone should have housing" is vague and rhetorical. "Building ten new units of low income housing" is a specific action that can be implemented. This is a threshold question, because it is hard to judge an action on the other criteria if it is not specific enough to be accomplished.

Leverage: How much difference will the proposed action make on results, indicators, and turning the curve? This is the most important of the criteria. It doesn't matter how well an idea scores on the other criteria if it won't make any difference. Investing in quality child care is a high leverage action to get children ready for school. Handing out parenting literature at the county fair is not necessarily a bad idea, but it's a low leverage action.

Values: Is it consistent with our personal and community values? There are many actions that are specific and high leverage but not consistent with our values. If you want to improve school attendance rates then kick out all the troublemakers. If you want to turn the curve on rates of entry into foster care, then slow down or stop doing investigations of child abuse.⁴³ These are specific high leverage actions that will make a difference but should not be done because they're wrong.

Reach⁴⁴: Is it feasible and affordable? Can it actually be done and when? No-cost and low-cost actions will rate higher here. Actions that require significant new resources will rate lower. Actions where there is a clear lead person or organization will rate higher. Actions where everyone says it's a good idea but no one wants to do it will rate lower.

Each proposed action item can be rated "high," "medium" or "low" on each criteria. The best pattern is obviously HHHH. This pattern is rare. More often you find actions that are strong on some criteria and weak on others. Consider the action items that rate highest on specificity, leverage and values, and space them out over a multi-year period. Actions that rate highest on the first three criteria can be done this year and next year. Lower rated actions can be done in the next 3 to 5 years.

When this rating process is used with a large group, it is best to do the rating work in small groups⁴⁵ first and then compare the choices of the groups. Often the small groups will separately identify the same priorities. This convergence of opinion can help develop a larger group consensus. When giving instructions to the small groups ask for one person in each group to be the "ombudsman"

^{42.} You can use the pneumonic SILVER to remember these criteria.

^{43.} There is one state that is reputed to have slowed down child abuse investigations in the early 1990's to control foster care caseloads.

^{44.} The use of the word "reach" here is a deliberate reference to Lisbeth Schorr's *Within Our Reach* where she argues that we really know a lot about what works, and our failures are most often failures of will and not knowledge.

^{45.} The ideal size for small group work is 6. It is large enough to provide for diversity of opinion and small enough that no one can hang back and not participate.

for the different criteria. This helps assure that each criteria is fully considered and helps create a lively debate.

Budget: While you can and should get started with no-cost and low-cost actions, many of the most important actions will require money. How do you finance an action plan? Contrary to what you may think, the single most important financing strategy is having a compelling vision. **Powerful visions are magnets for resources.** People with money want to buy success. The most common financing mistake is to go after the money first and then figure out what to do with it. In Results Accountability, the vision and action plan come first, and financing follows. The second most common mistake is to seek a single source of support for an action plan. People sometimes have had success with only one or two financing strategies, such as getting a foundation grant, or donations from the business community, and they limit their thinking to these approaches. A financing plan must be developed as a <u>package</u> of resources. It is a fundamental rule in financing Self Assessment" in Appendix I provides a systematic method for thinking broadly about financing options for an action plan.⁴⁶

^{46.} For more information on this approach see "The Cosmology of Financing," www.resultsaccountability. com.

Population Accountability

Getting from Talk to Action



Figure 3.2

The 7 Population Accountability Questions

- 1. What are the quality of life conditions we want for the children, adults and families who live in our community?
- 2. What would these conditions look like if we could see them?
- 3. How can we measure these conditions?
- 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?



Considering the talk to action process as a whole: How is the Results Accountability thinking process different from other processes? It's quite simple. Results Accountability takes the definition of success and moves it to the top of the page, in plain language and in measurable terms, and uses that definition of success to drive action plans and budgets. That's not a new method at all. Businesses have been working that way for thousands of years.

This connection between Results Accountability and business was driven home by an experience I had one day in a doctor's office. You know how piles of magazines collect in doctors' waiting rooms. That day on the top of the pile next to me was a very fancy publication from Merrill Lynch Associates, stamped CONFIDENTIAL in big red letters. I couldn't resist looking at it. What I found inside was one page per company. Each multi-billion multi-national company was reduced to a single page. The top of the page showed the numbers about the company's performance. The middle of the page showed the story behind those numbers: what was happening in the market place, what was happening in the company's management. The conclusion was an action plan, one of three words, "buy, sell or hold."⁴⁷ I was worried that this doctor had this confidential document sitting out in his office. So I took it. I dare not reproduce a page here. But there are countless examples of businesses using this same thinking process.

Action items after your first meeting: You can get through the entire Results Accountability thinking process in less than an hour. The Turn the Curve Exercise is one way to do this. Results Accountability does not require 3 months getting step one perfect, and 3 months getting step two perfect, and by the time you get to step three you forgot why you were doing the work in the first place. Results Accountability thinking is designed to be done quickly and then repeated. Each time you repeat the thinking process, your action plan gets better. This means that any group can have an action plan after the first meeting. The first action plan will not be complete, of course. But it's a start. The actions in the first-meeting action plan are usually no-cost and low-cost ideas. As you refine the plan in later meetings and begin to gather resources, the action plan will get better.

What we usually do is skip everything at the beginning of the process and go straight to what works. We have a fight about everyone's favorite program. We divide up whatever money is available. Then three years later we call in an evaluator and ask if the evaluator could please fill in the top of the chart. It's too late. This business of skipping ahead to solutions is very tempting and very common. Much of our political discourse is about means and not ends, about actions and not results. It takes discipline and even a little courage to help a group of partners think about results, indicators and causes, before they craft solutions.

The Most Powerful Question: "What would it take to succeed?" This is a better question than "what works?" I believe this is the most powerful question facing our society today. When the Allies invaded Normandy, they did not ask, "What works?" They asked, "What would it take to succeed?" When NASA planned its lunar missions, they thought about what it would take to succeed and they succeeded. We need to bring this same approach to the biggest challenges in our society today. What would it take to end or significantly reduce poverty? What would it take to end or significantly reduce poverty? What would it take to end or significantly reduce a sustainable environment?

^{47.} Thanks to Merrill Lynch for not suing me for describing their report.

The "what would it take?" question is about ambition. Ambitions don't guarantee success, but they make it possible. A sense of ambition shows up in the thinking of Geoffrey Canada, creator of the Harlem Project. "He wants to prove that poor children, and especially poor black children, can succeed – and not just the smartest or the most motivated or the ones with the most attentive parents, but all of them, in big numbers."⁴⁸ This should be the way we all think.

Some people say that it is too hard to reduce teen pregnancy, juvenile crime, and poverty. "Can't we work on something easier?" You can keep picking easier and easier things to work on, and pretty soon you have something that's not very important. There will always be a need for short term successes but ultimately we will be judged as a society by progress on the big stuff. We will never make progress on the big stuff unless we face it head on and ask "What would it take?"

Research proven practice: I am very troubled by the movement to fund only research-proven practice. This movement's intent to take advantage of proven methods and make the best possible use of scarce dollars is laudable. But there are two problems. First, the research world does not have all the answers. There are many important and powerful ideas that have never been fully tested by research. If we only fund programs that have been tested, we cut ourselves off from these other ideas. Second, thinking beyond research is often the most creative edge of the work. We must allow for the development of new knowledge. I have no problem with giving preference to research, but we must also allow research to be adapted to the unique circumstances of each community, and we must give people room to experiment and learn.

Core strategies vs. result-specific strategies: Many if not most actions contribute to more than one result.⁴⁹ Creating a community center could help make the community safer, help children succeed in school, and help parents get jobs. Groups that are working on more than one result at a time should identify the core strategies that cut across results and those that are specific to one result. Written plans can be coded to show core actions that may deserve a higher priority in developing the action plan.

Where do you start? It doesn't matter where you start. Start anywhere. Start where the passion is. Whatever you do will have beneficial effects across a range of results. Con Hogan, former Secretary of the Vermont Agency of Human Services, has a great way to summarize this: "Anywhere leads to everywhere." If you improve the teen pregnancy rate you will also improve the high school graduation rate. If you reduce teen smoking you will impact rates of low birthweight babies. All of these conditions of well-being are tightly interconnected. Touch the web anywhere and the effects are felt throughout.

How to choose population results

Choosing results is as much a political process as it is a technical process. What set of statements captures the most important hopes for our children, families, community, city, county, state or nation?

^{48.} New York Times Magazine, June 20, 2004.

^{49.} This is one of the consequences of the non-linear relationship between program and performance accountability.

One way to elicit this from a group is to put the following unfinished sentences on the board:

- We want children who are...
- We want families that are...
- We want to live in a nation, state, county, city or community that is...

The group is asked to complete these sentences with adjectives descriptive of well-being. For example, "We want children who are safe, happy, confident and ready for school." "We want families that are safe, stable, connected, self sufficient." "We want to live in a community that is clean and safe, where neighbors know each other." Each completed sentence is a result. After the group lists its ideas, three small groups can be formed to edit these down to a shorter, more easily communicated list. The children's group is asked to produce 5 statements. The family group is asked to produce 3 statements. The community group is asked to produce 2 statements. At the end of about two hours, there is a list of 10 results that represents the group's consensus.⁵⁰

There are three elements that are typically found in broadly based processes that have successfully developed results and indicators.

- There is a sponsoring group that consists of high level, well respected leaders. In Georgia, the Family and Children's Policy Council that developed Georgia's results for children and families had leaders from the legislature, the governor's cabinet, the business community and the faith community.
- If you pick the right leaders, they won't have time to do all the work. The group must have staff to support the work of the leadership group. This is anywhere from a half-time person to 5 or more staff. The leadership considers the staff's recommendations and creates a working list of results and indicators
- The leadership group sponsors an inclusive comment process with public hearings and focus groups. The leadership group considers this input, makes adjustments to their initial proposals, and issues a "final" list of results.

Figure 3.4 shows examples of results that come from these kind of processes. Notice that these are very different in content and style, but in each case they are plain language statements of well-being.

^{50.} This process is more fully described in the exercise "A Results List from Scratch," www.www.raguide. org.

Examples of Results					
Georgia	Placer Co. California ⁵¹	Vermont			
 Healthy children Children ready for school Children succeeding in school Strong families Self-sufficient families 	1. Safe 2. Healthy 3. At home 4. In school 5. Out of trouble	 Families, youth, and individuals are engaged in their community's decisions and activities. Pregnant women and young children thrive. Children are ready for school. Children succeed in school. Children live in stable and supported families. Youth choose healthy behaviors. Youth successfully transition to adulthood. Adults lead healthy and productive lives. Elders and people with disabilities live with dignity and independence in settings they prefer. Communities provide safety and support to families and individuals. 			



Avoid referencing services in results statements: One of the most important characteristics of a well constructed result statement is that it is not about data and not about service. In the Results Accountability framework, the need for data is addressed by indicators, and services are best considered as means and not ends.

People often propose results that make reference to a specific service. One of the most common is "All pre-school children receive high quality child care."⁵² Whenever you use this kind of statement as a starting point for the Results Accountability process, it quickly leads to not-very-helpful circular thinking.

Result: All children receive high quality child care **Indicator**: Percent of children who receive high quality child care **Story**: Not enough high quality child care **What works**: More high quality child care.

^{51.} See David Gray's excellent book: *Safe, Healthy, At Home, In School and Out of Trouble: Making child and family services work for children and families,* 2004.

^{52.} Not all child care is created equal, and when the term child care is used here, it refers to high quality child care.

It is better to think of any service, no matter how important, as a means, not an end in itself. So, in this case, child care can and should be part of the discussion of what works to advance the result "all children ready for school."

The same rule against referencing services in results statements also applies to indicators. If the indicator is "percent of people receiving service" then the same circular thinking leads to a strategy to provide more of that service. Service thinking alone is too narrow to produce a good action plan. Getting children ready for school, for example will require the contributions of many partners, such as parents, grandparents, the business community, the media, that are not part of the formal service system.

One immediate benefit of this shift in thinking is that it puts advocacy for that particular service on a different and more solid footing. Rather that arguing that child care is a good in and of itself, the argument shifts to how child care contributes to the higher purpose of assuring that all children are ready for and succeeding in school. Child care becomes more than just another service competing for money, but part of a larger strategy to achieve something everyone can readily support. It provides space for us to think deeply about other actions required to achieve these ends. We remove a burden that is unfairly placed on child care as the only path to school readiness and school success for children.

This same shift in emphasis can strengthen advocacy for other services. Instead of arguing for more child protection staff, frame the discussion around child safety. If the public accepts the importance of child safety, they will be more likely to support funding for the programs that promote safety.⁵³

Another way in which services are referenced in results statements involves a description of the intended quality of government services. For example, one county sets out *Modern and Responsive County Government* as one of several population results. This is arguably a result, since it is a condition of well-being for people in a geographic area. The problem is that virtually every piece of data that describes this condition will be a performance measure for the county service system. This blurs the line between population and performance accountability and runs the risk that people will see good county services as an end and not a means. The better way to handle this is to produce documents that have clearly separate sections on population accountability that addresses the quality of life for all people in the area and performance accountability that addresses the performance of the people's government and its programs. In Chapter 6 we will look at budgeting and planning formats that clearly separate population and performance accountability.

As you might expect, there are exceptions to these rules. In much of the world, education services are severely restricted for young girls and women.⁵⁴ The numbers and percent of young girls and women eligible for and participating in the education system becomes every bit as important as the educational results themselves. There are times when the availability of an important service is a kind of result. However, this construction should be avoided whenever possible.

^{53.} This is not much different from the ancient rules of advertising, where you must first convince people they need new clothes, a car or golf clubs, before offering up the particular one you want them to buy.

^{54.} This is a major priority of the UN World Population Fund and statistics are available from their website.

Another common mistake is to structure the quality of life report card using service headings like: Health, Education, Environment, and Safety. People sometimes think these are results, but they are not. Health, education, environment and safety are categories of service or labels for the organizational components of government. Using government service and organizational categories reinforces the old way of thinking that pins responsibility for a population condition on one agency. The solution to this problem is quite simple:

Instead of "Health" use All Children and Adults are Physically and Mentally Healthy. Instead of "Education" use All Children are Ready for and Succeed in School. Instead of "Environment" use A Clean and Sustainable Environment. Instead of "Safety" use Safe Communities.

The difference in wording may not seem like much, but the underlying message is very important. When true results statements are used, it is much easier to see that responsibility for progress does not, and cannot, rest with any one department of government or even with government alone. Results have tremendous power to motivate people and bring people together. Services categories do not.

Another common problem with result statements has to do with the words "improve, increase and decrease." Consider the result *Improved High School Graduation*. There are two problems here. First, the statement implicitly includes the high school graduation <u>rate</u>. Data statements are always best addressed as indicators, not results. Secondly, the word "improve" may signal a lower level of ambition than the group intends. What if the graduation rate goes from 50% to 51%? This passes the test for improvement, but is certainly not what anyone hopes to achieve. A better and simpler construction would be to make the result: *All Children Succeed in School* and one of the indicators the high school graduation rate. The use of the word "all" is optional, but it signals our ambition that all children succeed, that it is not OK for us to allow some children to fail. This does not mean that we expect to fully achieve this ambition. We will get incrementally closer by turning the curve on the high school graduation rate. But the most important statement of our purpose is not limited by settling for the achievement of a particular percent or amount of increase.

Avoid the "King of the Hill" syndrome where one result is given pre-eminent importance over all others. In one county, the county commissioners authorized a significant effort to advance *Family Self Sufficiency*. A range of other results (*Children are Healthy and Safe*, etc.) were then identified as subordinate to self-sufficiency. Clearly health and safety both contribute to family self-sufficiency. But self-sufficiency also contributes to health and safety. When any one result is designated king of the hill, all other results can be viewed as subordinate contributors. This relationship is artificial and unhelpful. When more than one result is articulated, it is better to view them as co-equal, not hierarchical.

Avoid creating multiple levels of results: Aside from the artifice of one pre-eminent result, there are situations where lists of results have two (or even three) levels. So for example:

Level 1 Result: Clean Environment Level 2 Result: Clean Air Level 2 Result: Clean Water

Each level is a plain language statement of well-being. But multiple layers make the work more complicated. People sometimes further confound the situation by calling the different levels by different names. Level 1 results are called "Results." But level 2 results are called something else such as "Objectives," or "Goals." Since both Level 1 and Level 2 are actually conditions of well-being, per our definition of results above, the introduction of another label for this same idea completely undermines any chance for discipline in language and thinking.

From a practical standpoint, it is best to avoid all this complexity by having one level of results and then using indicators to address what would otherwise be a second or third level. The example above would become:

Result 1: Clean Environment Indicator 1: Percent of days with clean air Indicator 2: Percent of stream miles meeting water quality standards

Clarity and simplicity are virtues in choosing results and indicators.

How to choose indicators to measure results

For each result, there is a set of indicators that reflect the extent to which the result is being achieved. Figure 3.5 shows some examples of indicators:

Examples of Indicators				
Montgomery Co. Ohio ⁵⁵	Vermont ⁵⁶			
<u>Result</u> : Stable Families	<u>Result</u> : Children live in Stable Supported Families			
Indicators:	Indicators:			
1. Percent of first births where both par-	1. Percent of children in poverty.			
ents completed high school, parents are married (at any time from conception	2. Percent of children in families receiving food stamps.			
to birth), and the mother is at least 20	3. Percent of child support paid.			
years old.	4. Percent of parentage established for out-of-wed-			
2. Rate of substantiated child abuse	lock child support cases.			
and neglect.	5. Rate of substantiated child abuse and ne-			
3 & 4. Deaths to children ages $0 - 17$ that	glect.			
were ruled preventable, or somewhat preventable, by the Child Fatality Re-	6. Percent of children ages 5 – 17 in families receiving welfare.			
view Board.	7. Rate of out-of-home placements.			
5. Domestic violence deaths.	8. Average number of moves within the child sub- stitute care system.			

Figure 3.5

^{55.} Montgomery County Family and Children First Council, 2004 Progress Report, Outcomes, Indicators and Strategic Community Initiatives, December 2004.

^{56.} Vermont Agency of Human Services, 2004 Community Profiles, January 2005, www.ahs.state.vt.us.

Notice the differences between how these very similar results are measured. Montgomery County uses five measures. Vermont uses eight. Only one measure (rate of substantiated child abuse and neglect) is used in both places. Vermont lists poverty statistics here while Montgomery County lists poverty rate under a separate result: *Economic Self-Sufficiency*. Both lists are powerful and appropriate. There is not a right or wrong set of measures for any given result. Choosing indicators will always be a matter of approximation and compromise, and different groups will come up with different answers about what indicators to use.

How do you choose the best indicators to represent a result? Any choice can be reduced to a set of criteria. The following three criteria have been used to choose indicators in many processes.

Communication Power: Does the indicator communicate to a broad and diverse audience? This criteria is sometimes called the *public square test*. If you had to stand in a public square⁵⁷ and explain the result to your neighbors, what two or three pieces of data would you use? Obviously you could bring a thick report and begin a long recitation. The crowd would thin out fast. It is hard for people to listen to more than a few pieces of data at one time. The data must be common sense, and compelling. The crime rate has communication power. The rate of successful adjudication does not.

Proxy Power (or Representative Power): Does the indicator say something of central importance about the result? Can this measure stand as a proxy or representative for the plain language statement of well-being? We know, for example, that the percent of children reading at grade level in the 3rd grade is a powerful measure of school success. Children who can't read in 3rd grade have a much higher chance of failing in later grades and dropping out of school. So 3rd grade reading scores are a powerful proxy for the result *All Children Succeed in School.*"

The other part of proxy power has to do with the fact that **data tend to run in herds**. If one indicator is going in the right direction, usually others are as well. You do not need 20 indicators telling you the same thing. Pick the indicators that have the greatest proxy power, specifically those that are most likely to match the direction of the other indicators in the herd.

Data Power: Do we have quality data on a timely basis? Is the data reliable and consistent? To what extent do we have the data at the state, county, city and community levels?⁵⁸

Each indicator is rated High, Medium or Low on each criteria. We are looking for indicators that rank high on all three criteria. These are indicators that communicate well, that tell us something of central importance about the result, for which good data is available. If we can find indicators that have these three characteristics, there's a good chance they will work with our neighbors in the public square.

^{57.} Or perhaps the central atrium of your local shopping mall.

^{58.} Some places prefer to use a larger list of criteria, or phrase the criteria differently. In almost every case, these separate criteria are components of Communication, Proxy and Data power. For example, one county split Data power into four separate criteria: Valid, Available, Accurate, and Reliable. The same county thought "communication power" sounded too much like jargon. So they renamed it "Easily understood." This is all good. Whatever works.

There are two messages in this rating system. The first message is "Start with the best of what you have." The second message is "Get better."

Data is often used as an excuse, "We could be accountable for the children in our community when we have better data." If we are honest, we will admit that we said that 10 years ago. We'll say it again 10 years from now. Do not make the collection of new data a precondition for getting started.⁵⁹ You have to start with the data you have. In fact, only by starting with the data you have do you have any chance of generating the support you need to get better data.

Choosing Indicators Worksheet

Outcome or Result	Safe Community			
Candidate Indicators	Communication Power	Proxy Power	Data Power	
Measure 1	HML	HML	HML	
Measure 2 Measure 3	CH	Н	H	
Measure 4				
Measure 5	H	H	LR	
Measure 6 Measure 7			Data	
Measure 8			Development Agenda	



These three criteria lead to a three-part indicator list <u>for each result</u>:

<u>Primary or Headline Indicators</u>: The 3 or 4 most important measures, the ones that rise to the top in the rating process.

<u>Secondary Indicators</u>: Any other measures for which there is good data. We do not throw away good data. We will use these measures in assessing the story behind the baselines, and other parts of the process.

Data Development Agenda: The priorities for new and improved data.

This is not a process that is done once and is then finished. This is a living, changing list. As you develop new and better data, you may promote one measure to become a primary indicator and demote another measure to become a secondary indicator.

Baselines and the definition of success

Like other words in this book, the word "baseline" has many possible meanings. Results Accountability uses a definition that comes from the budget world. Budget baselines have two parts: an historical part that tells us where we've been, and a forecast part that shows where we're headed if we don't do something differently.

 ^{59.} The Results Accountability process can be started without any data at all. Groups can use the results and experience to drive the thinking process. Groups can also create working versions of indicator baselines based on group consensus about history and what the future will look like if we don't change.

Creating the forecast part of a baseline is an art, not a science.⁶⁰ There is not a single, right forecast. It is sometimes useful to show a range of forecast scenarios: high, medium, and low; or optimistic, pessimistic, and likely.

As chief financial officer of the Maryland Department of Human Resources, I was responsible for overseeing the department's forecasts of welfare caseloads, child support collections, foster care caseloads and child abuse and child neglect reports. These forecasts were needed to estimate the department's staffing requirements and the end-of-year surplus or deficit. Each quarter, we updated our forecasts through the end of the current and succeeding fiscal years. We tempered our review of the numbers with what we knew about policy changes, developments in the local departments, and what was happening on the street. The best forecasting work is based on this kind of open learning process.

Not all forecasts are straight lines. In economic forecasting, periods of growth are followed by periods of decline. Economic forecasts almost always involve estimating when the current period of growth or decline will slow down and turn in the other direction. Straight line forecasts also don't apply when the numbers in the historical part of the baseline are very small. For example, the number of suicides in a small geographic area often go up and down without any clear pattern. The forecast in this case might be a continuation of the up and down pattern in the same range. In Boston, the number of juvenile homicides ranged from 6 to 16 over a six year period in the late 1980's and early 1990's. While it was not possible to predict exactly where this line would go, there was little argument that it would continue in this range if something was not done to change it.

We are finding more and more reports that present the historical part of baselines but not the forecast part. The forecast part is essential for two reasons. First, it allows us to ask and answer the questions, "Is this future OK?" "Juvenile crime has been rising in our community. If we don't do something about it, this is where we're headed, folks. Is that OK?" "Water quality has been declining. If we don't do something about it, this is where we're headed. Is that OK?"

If people say "Yeah, that's OK with me," then you have no basis to move from talk to action. The first step in getting from talk to action is for someone to say



Baselines allow us to define success as turning the curve not point-to-point improvement.

Figure 3.7

^{60.} One of my favorite Yogi Berra quotes comes from one year in the 1950's, before the World Series, when Yogi was asked to forecast his performance in the upcoming series. He hesitated for a second and said, "You know forecasting is really difficult... especially about the future."

"That's <u>not OK</u>, we could do better." Stephen Covey talks about "pain and vision" as the ingredients of change.⁶¹ The forecast of an unacceptable future is the pain part of the equation.

The other reason why baselines with forecasts are so important is that they allow us to define success as **turning the curve** away from the baseline or beating the baseline. This is a very common sense idea that we see all the time in the rest of the world. The U.S. Congressional Budget Office routinely forecasts policy-neutral estimates of federal deficits against which tax and budget proposals are debated. Similar forecasting occurs routinely in business planning.

In most work on population and performance accountability, success is not defined this way. Success usually means <u>point to point</u> improvement, moving from where we are now to a new better place. Crime should go down. High school graduation should go up. Every measure has a success direction and success is moving to a new place in that direction. "We will be successful if crime is reduced by 20% in the next two years." "Our school will be successful if reading scores go up by 5% per year."

The problem is that the world doesn't work that way. Most conditions we are trying to change have a history. They have momentum. Often the trend has been headed in the wrong direction for a long time. Sometimes the best you can do is slow down the rate at which things are getting worse, before you can flatten the curve and turn it in the right direction.⁶² Every time we define success as point to point improvement, it is a setup for failure. Two years later, if we haven't hit the new lower (or higher) level, our work is declared a failure and we go on to the next fad. If we are serious about this work, we have to be more business-like about it. And business-like means using baselines and the concept of turning the curve to define success.

Some people argue that the notion of turning the curve is too complicated and takes too long to work in a political environment. Impatience, driven by term limits and the frequent turnover of elected officials, has made the demand for short term improvement even stronger. However, most elected leaders want a way to assess the big picture, whether or not we, as city, county, state or nation are really making progress. Is all this money we're spending on programs getting us somewhere? Baselines can help answer this question. The political need for short term accomplishments will never change. But we can add a systematic way of assessing if all this short term activity adds up to anything.

Baseline curves don't always take a long time to turn. Changes in program performance measures and population indicators can happen much more quickly than people think. If you look at the stories in this book about Boston's work on juvenile homicide, or Tillamook County's work on teen pregnancy, you can see significant progress in 15 to 18 months.

One common mistake in creating baselines is to forecast where you want to go instead of where you think you're headed if you don't do something different. The purpose of the baseline forecast is to

^{61.} I'm pretty sure he said this.

^{62.} The idea of turning the curve does not depend on the baselines headed in the wrong direction. If the baseline is headed in the right direction, then turning the curve means accelerating the rate of improvement.

establish what a policy-neutral or budget-neutral future might look like, so that the discussion can focus on whether proposed actions will make things better or worse than this future. The policy-neutral path should always be clearly differentiated from the policy-change path.

Of course, any framework can be misused. In one county, a department of county government was trying to sell the county council on more money for a teen pregnancy prevention program. So they drew a chart with the history of teen pregnancy rates and two alternative forecasts. The first forecast, labeled "without investment" made a sharp, almost 90 degree turn upwards. The second forecast labeled "with investment" made an equally sharp turn in the other direction. It was a joke. Nobody believed either forecast. I don't remember if they got the money or not. But they should have been embarrassed. And I was em-



barrassed because they said they were using Results Accountability.

Baseline forecasts must be credible, otherwise the process collapses. It is important to involve key stakeholders in the creation of credible forecasts. Most states, counties and cities have formal forecasting processes for important revenue and expense components of the budget. In Florida, there is a joint executive and legislative branch committee that establishes projections for revenue and for key non-discretionary expenditures. This expertise exists in most state and local jurisdictions to one degree or another. Building a consensus on the forecast of "what will happen if we don't do something different," will pay dividends later in the process. It lays the groundwork for the discussion about what needs to be done.

Quality of life report cards for children, families and communities

Once a set of results and indicators has been developed, it is possible to create an annual report card for a city, county, state or nation. Looking at this kind of report card is like looking in a mirror. People see whether the community's quality of life is getting better or worse. Sometimes people look at the report card and they don't like what they see. This can be the first step in bringing people together to take action.⁶³

Such report cards have been developed for dozens, perhaps hundreds of cities, counties and communities across the United States. There are many different ways to construct a report card, but the best ones use a structure based on results and indicators. Vermont has produced some of the best work. The excerpt from the Vermont 2001 report (Figure 3.9) shows two facing pages for one of

^{63.} See "Bringing people together to do this work" in Chapter 7.

Vermont's ten outcomes.⁶⁴ The left side shows "Heartening Indicators," "Troublesome Indicators," and the "Story behind the curve." The second page shows the action plan in two parts: accomplishments from last year and recommended actions for next year. This is a brilliant construction that puts all the essential information for each outcome in a format that every community member can understand.

In 1991, the Annie E. Casey Foundation issued the first Kids Count report comparing all 50 states and the District of Columbia on 8 indicators of child and family well-being. Kids Count has grown to include a network of state partners that now issue reports for counties (and sometimes cities) within their states. These reports have helped to create a national movement toward regular periodic assessment of indicators of child and family well-being. Links to the Kids Count reports can be found on www.aecf.org. A partial list of exemplary report cards on child and family well-being includes those found in Contra Costa County, Los Angeles County, Marin County, San Mateo County, Santa Clara County, and Santa Cruz County, California; Montgomery County, Ohio; Philadelphia, Pennsylvania; and at both the state and local levels in Georgia, Oregon and Vermont. Links to other exemplary children's report cards can been found on www.raguide.org.

 ^{64.} Outcome-Based Planning: State Partners and Local Communities Working Together to Improve the Well-Being of All Vermonters: A Report from the State Team for Children, Families and Individuals,
 Vermont Agency of Human Services, February, 2001.



· There has been a significant decline in child abuse and neglect. A major effort by the Department of Social and Rehabilitation Services to encourage reporting, coupled with a broadening of policy in "risk-of-harm" cases, has resulted in an increase in both reports and victims in the past 12 months.

nancy programs.

· Child support collections have had significant increases related to an excellent automation system, mandatory wage withholding law, and the ability to attach assets.



- Children in custody move more frequently due to an increase and intensity of challenging behaviors.
- The lack of availability and affordability of housing continues to affect the number of people housed in temporary shelters in Vermont.

Recommendations	Accomplishments in 2000	Recommendations for 2001	Actions/Strategies 2001
Famílies achieve economic independence	 Conference on "living assets project" held in 2/2000 in- cluded in Welfare Reform legislation. 	Families live in safe and affordable housing	 Provide supports to help keep people in their homes. Provide case manage
	 Community Action Program agencies distributed livable wage information. 		ment services to support high-risk tenants.
	 CAP agencies, the Department of Prevention, Assistance, Transition and Health Access and Headstart distributed information about earned income tax credit. 		 Promote affordable housing in all local planning initiatives.
partners in developing	 Vermont Federation of Fami- lies are now operating in seven regions of the state. 	Families are supported by their communities	 Provide bette parenting support fo incarcerated parent and support their tran sition back to their families.
	 Parent to Parent has regional capacity. 		
	 The Parent-Child Centers have been doing Parents As Teachers and outreach to families about how services work. 		 Support and enhance the role of fathers in their children' lives.
	 Interagency Training Commit- tee began developing informa- tion to ensure accommoda- tions and support for parents with disabilities. This will be continued and enhanced by the Green Mountain Family Network Grant. 		 Family organization expand connection with Regional Partne ships.
			 Vermont Federation of Families will continue to expand family resource
	 Victim Advocacy programs have been strengthened. 		consultants to all regions.
	 The Legislature requested a substance abuse prevention intervention and treatment framework from the Depart- ment of Health for January, 2001. 		 Strengthen the cc laboration of familie court personnel ar service providers wi an annual "Meet t Judge" opportunity.
	 All regions of the State have Community Resource Guides in hard copy and on the internet. 		 Encourage hiring of experienced family members in peer support and mentor roles.
	 Eight Parent Leadership programs are in operation. 		 Increase the number of services and activities associated with afterschool programs.
	 There is an SRS Consumer Advisory Board in operation in each region. 		
	 Six Regional Partnerships were funded for family involvement through the Family Consor- tium and the tymaining Re- gional Partnerships are now applying for Green Mountain Family Network grants in the 		 Determine effecti ways to support t parents of preteens promoting positi youth development Green Mountain N work/Children
	second round.		UPstream Servic (CUPS) Family Cc sortium grants will pand family networ to all regions.

Figure 3.9

6

WHERE RESULTS ACCOUNTABILITY THINKING HAS WORKED

Vermont⁶⁵

One of the places that has made the most progress using Results Accountability is Vermont. In the early 1990's, Vermont identified ten "outcomes" for children, families and communities, and began issuing report cards, known as Community Profiles. The profiles are produced for each of the 59 school districts in the state and show baselines for over 70 indicators, organized by the state's ten outcomes.⁶⁶ The reports are produced by the Vermont Agency of Human Services in partnership with the Vermont Department of Education.

Vermont established the State Team for Children, Families and Individuals, to work on improving outcomes across the state.⁶⁷ Membership on the State Team includes a balance of state agency staff and representatives from the state's 12 regional partnerships. The State Team has 10 standing subcommittees, one for each of the ten outcomes. The subcommittees take turns reporting to the larger group over ten meetings per year. Each of the regional partnerships has a similar process to track indicator data, organize partners and take action. There have been significant improvements in many of the state's indicators. The state and regional partnerships have many Turn the Curve stories to tell. Here is one of them.



Figure 3.10

Vermont is a small state of only 630,000 people. The state has a high percent of older housing stock, particularly in rural areas. Many of the houses constructed prior to 1978 contained lead paint, which is especially toxic to toddlers and young children who ingest paint peelings. Lead poisoning

^{65.} Thanks to Con Hogan, David Murphey, and all the Vermont partners.

^{66.} The full set of these report cards can be seen at the website of the Vermont Agency of Human Services, www.ahs.state.vt.us

^{67.} We sometimes have trouble coming up with inspiring names for the groups that do this work. In one state the local family and children collaboratives are known as "Local Management Boards." Contrast this with "Family and Children First Councils" in Ohio, or "First 5" in California, or "Family Connections Collaboratives" in Georgia.

is known to have life-long damaging effects on a child's mental and physical development. Over time, the indicator "Lead Levels in Children" remained stubbornly unacceptable.

By the early 1990's there were several public health education efforts geared to bringing more awareness to this issue. In 1991, the state and local partnerships decided to include a focus on lead paint in their new strategic effort to visit the homes of families with newborn babies. These visits were designed to offer information, provide opportunities to sign up for needed services, and help make social connections for young families who often lead isolated lives.

Every family with a newborn (there are about 6,000 births per year in Vermont) was offered a visit by someone trained to make these visits. The family "visitors" came from different institutions depending on the district. In one district, the visitor was a school nurse. In other districts, a public health nurse or community volunteer played this role.

One of the conversations that occurred during these visits concerned lead paint.

Was this house built before the 1980's? If so, may we look at the paint around some of the windows? Are there any paint chips around these window that have come loose over the years? It is very likely that some of these chips may have been ground into dust over time, and it is also likely that your young toddler may be ingesting some of this lead filled dust. May we arrange for a lead screening for your child? There is a state lead abatement program that exists to help take care of problems like this. Would you like us to connect you with that program?

In addition to lead levels, a number of other indicators were expected to improve as a result of home visiting and related community efforts, including 2 year immunization rates, abuse rates for young children, and children's health coverage. Over time, the larger strategy of offering visits to families with newborns contributed strongly to turned indicator curves in these areas. Eighty eight percent of families accepted the home visits. In the case of the percent of children ages 1–5 with elevated lead levels, the curve turned dramatically, dropping from 11.3% in 1994 to 5.9% in 1999. The state percentages dropped significantly faster than national totals during this period, cutting the gap from approximately 6.9 percentage points in 1994 to approximately 3.3 percentage points in 1999-2000.⁶⁸

Vermont No-cost and Low-cost Stories

Vermont also offers some good examples of where no-cost and low-cost actions made a difference.

In the mid 1990's, Medicaid costs were rising rapidly in Vermont as they were in the rest of country. One of the largest components of Medicaid is nursing home care for elderly and disabled individuals. In one region of the state, these costs were growing much faster than in the rest of the state. The state asked the local regional partnership to bring people together to see why this was happening

^{68.} The percent of children 1 to 5 in the United States with elevated lead levels in their blood (above 10 ug/dL) was 4.4% for 1991 – 1994 and 2.2% for 1999-2000 according to the CDC National Health and Nutrition Examination Surveys (NHANES).

and what could be done to slow the rate of nursing home entries. As an added incentive, the state offered to share some of the Medicaid savings with the local group if the rate of increase could be slowed. The group soon discovered that one of the leading causes of entry into nursing care was hip fractures. This is a very serious injury for an older person. Often it is not possible for the person to return home after their hospital stay and nursing care becomes the only option. The group dug deeper into the story behind the curve. The group soon discovered that one of the leading cause of these falls was... winter. The elderly were slipping on the ice and snow. The group decided to organize a community-wide effort to shovel the walks of the older people in the community.

Another community with a similar challenge organized something they called "drive by shoveling." After a snow storm, three or four students and one teacher from the high school would drive around the town. If they found an un-shoveled walk, they would jump out, shovel the walk, get back in the car and drive off.

Still another no-cost effort involved the young people who served as baggers at the local grocery store. If an older person came through the line, they were asked to look down and see if the older person's shoelaces were untied. If the laces were untied, the young people were asked to get down on their hands and knees and tie them. No money was attached to any of these efforts. To my knowledge there is not yet a federal shoelace program. They represent the kind of common sense no-cost and low-cost contributions that everyday citizens can make to advance the well-being of their community.