

Measuring Outcomes

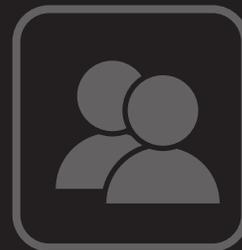
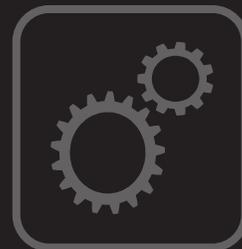
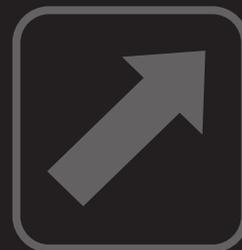
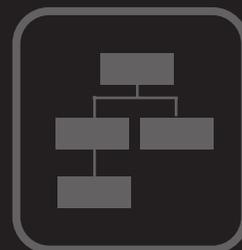
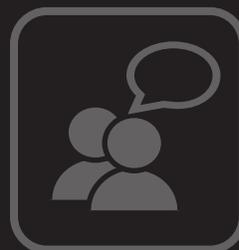


TABLE OF CONTENTS

- INTRODUCTION 4**

- OVERVIEW 5**
 - Why Measure Outcomes? 5
 - What is Outcome Measurement? 5
 - Some Words of Caution: Limitations of Measuring Outcomes 7
 - Effective Outcome Measurement Practices 8
 - An Outcome Framework..... 9

- GETTING STARTED..... 9**
 - 1. Where Should We Focus? 9
 - 2. What Do We Want to Accomplish with Our Outcome Measurement Efforts? 10
 - 4. How Can We Make This a Participatory Process? 11
 - 5. Is the Timing Right? 12
 - 6. What Resources Will We Need?..... 12

- PHASE 1: IDENTIFY OUTCOMES AND DEVELOP PERFORMANCE INDICATORS..... 13**
 - About Outcomes and Outcome Statements 13
 - Using Outcome Chains 14
 - Outcome Chain Example 15
 - Logic Models 16
 - Developing Performance Indicators..... 18

- PHASE 2: CREATE AND IMPLEMENT A DATA COLLECTION PLAN 21**
 - Data Collection Methods..... 21
 - Validity and Reliability 25
 - Instrument Development and Pre-Testing 26
 - Deciding When and How to Collect Data 26

- PHASE 3: ANALYZE THE DATA 32**
 - Tools for Analysis 32
 - Sample Analysis and Data Displays..... 33

PHASE 4: COMMUNICATE THE RESULTS	37
What to Communicate	37
How to Communicate.....	38
Target Multiple Audiences	38
PHASE 5: REFLECT, LEARN, DO IT AGAIN	39
APPENDICES.....	41
APPENDIX A: Online Resources: Survey and Questionnaire Development	41
APPENDIX B: Glossary	42
APPENDIX C: Sample Program Logic Model	44
APPENDIX D: Outcomes and Indicators Worksheet	45
APPENDIX E: Data Collection Plan Worksheet, Part 1 and 2.....	46
APPENDIX F: Getting the Best Results from Surveys and Interviews	48
APPENDIX G : Technical Assistance Survey Template	51
APPENDIX H: References.....	54

INTRODUCTION

The Compassion Capital Fund (CCF), administered by the U.S. Department of Health and Human Services, provided capacity building grants to expand and strengthen the role of nonprofit organizations in their ability to provide social services to low-income individuals. Between 2002 and 2009, CCF awarded 1,277 grants, and the CCF National Resource Center provided training and technical assistance to all CCF grantees.

Strengthening Nonprofits: A Capacity Builder's Resource Library is born out of the expansive set of resources created by the National Resource Center during that time period, to be shared and to continue the legacy of CCF's capacity building work.

Strengthening Nonprofits: A Capacity Builder's Resource Library contains guidebooks and e-learnings on the following topics:

1. Conducting a Community Assessment
2. Delivering Training and Technical Assistance
3. Designing and Managing a Subaward Program
4. Going Virtual
5. Identifying and Promoting Effective Practices
6. Leading a Nonprofit Organization: Tips and Tools for Executive Directors and Team Leaders
7. Managing Crisis: Risk Management and Crisis Response Planning
8. Managing Public Grants
9. Measuring Outcomes
10. Partnerships: Frameworks for Working Together
11. Sustainability
12. Working with Consultants

Who is the audience for *Strengthening Nonprofits: A Capacity Builder's Resource Library*?

Anyone who is interested in expanding the capacity of nonprofit services in their community – from front line service providers to executives in large intermediary organizations – will benefit from the information contained in this resource library. The National Resource Center originally developed many of these resources for intermediary organizations, organizations that were granted funds by CCF to build the capacity of the faith-based and community-based organizations (FBCOs) they served. As such, the majority of the resources in *Strengthening Nonprofits: A Capacity Builder's Resource Library* support intermediary organizations in their capacity building efforts. However, funders of capacity building programs (Federal program offices and foundations) and the nonprofit community (including FBCOs) at large will also find these resources helpful. In addition, individuals working to build capacity within a program or an organization will be able to use these resources to support their efforts to implement change and make improvements.

The *Measuring Outcomes* guidebook will be helpful to any organization interested in learning more about implementing or improving its procedures for measuring outcomes.

Who developed the *Measuring Outcomes* guidebook?

The guidebook was originally developed by the National Resource Center with assistance from Judy Clegg and Dawn Smart. It was updated in 2010 for the Department of Health and Human Services by the National Resource Center.

OVERVIEW

Although there are many uses for the information generated by **outcome measurement**, organizations often make the effort because they are required to do so. They are asked to be accountable for the use of their grant maker's funds. This includes foundations and grant making organizations such as the United Way, as well as local, state, and Federal government agencies.

In 1993, Congress enacted the Government Performance Results Act (GPRA) to ensure that the Federal government focuses on the performance of programs. Federal departments are now required to assess the effectiveness of their programs. Most programs require grantees to report on activities and outcomes throughout and after their grant, as well as participate in research and evaluation efforts.

Every organization hopes to deliver quality services. Outcome measurement will help you understand whether yours does. With the information you collect, you can determine which activities to continue and build upon and which you may need to change in order to improve the effectiveness of your program.

Why Measure Outcomes?

Outcome measurement has become one of the most important activities that social service organizations undertake. There are several compelling reasons to measure outcomes:

1. **Measuring the effectiveness of an intervention.** How do you know if a program was effective? If a program was not effective, would you want to know so that it could be improved? It is unlikely that we build capacity simply for the sake of building capacity, and because of this, we need to know whether or not our programs are effective, just as the nonprofits we serve need to know if their services are achieving the desired results.
2. **Identifying effective practices.** With the information you collect, you can determine which activities to continue and build upon. Some practices might be modified and replicated for other programs or initiatives based on your results.
3. **Identifying practices that need improvement.** Some activities may need to change in order to improve the effectiveness of your program.
4. **Proving your value to existing and potential funders.** The climate for funding social services and capacity building has undergone some radical transformations in the last few decades. Funders are keenly aware of the need to document the success of programs. Capacity building has lagged somewhat behind the social services field in having established evaluation standards. In 2003, the Alliance for Nonprofit Management reported that the capacity building field still lacked a framework for evaluation. However, this is changing, and it is important for capacity builders to stay ahead of the curve to be able to continue doing their work.
5. **Getting clarity and consensus around the purpose of your program.** Everyone in your organization, from board members to service staff to volunteers, should understand what is going on in your program and what it is intended to achieve. Outcome measurement helps to clarify your understanding of your program.

What is Outcome Measurement?

Accredit, adjudge, analyze, appraise, appreciate, assay, assess, audit, check, classify, consider, critique, determine, estimate, examine, evaluate, find, gauge, grade, inspect, investigate, judge, measure, monitor, rank, rate, referee, report, review, score, scrutinize, study, test, validate, weigh.

There are countless words used in the world of evaluation and a fair amount of confusion and debate about precisely what each word means. This guidebook uses the phrase “outcome measurement” to describe one approach to exploring the impacts or results of a program and to distinguish outcome measurement from more elaborate or complex types of evaluation.

Outcome measurement is “a systematic way to assess the extent to which a program has achieved its intended results.”¹ The main questions addressed in outcome measurement are:

- What has changed in the lives of individuals, families, organizations, or the community as a result of this program?
- Has this program made a difference?
- How are the lives of program participants better as a result of the program?²

During the last thirty years or so, most of the reporting required of groups receiving government or charitable funds has focused on what staff do, how many people they serve, and how many hours of service they deliver. Outcome measurement asks, and attempts to answer, the question, “So what?”

- So what if you provide an organization with ten hours of technical assistance on fundraising techniques? Is the organization better able to raise money? Do they?
- So what if you train an organization on how to develop a strategic planning process? Can the organization effectively perform the steps involved? Do they?
- So what if your staff works with five faith-based or community organizations on developing partnerships? Do those organizations actually follow through and increase their collaboration efforts? Do those efforts result in new partnerships?

Outcome measurement is most commonly used in the nonprofit world; a similar phrase, “**performance measurement**,” is used more often in the business and government arenas. In essence, they mean the same thing.

“**Compliance monitoring**” is another phrase in use today, referring most often to the contractual arrangements made between an organization and its grant maker on the use of funds. Compliance monitoring keeps records on what and how much service a program delivers, the clients it serves, and how much money it expends in relation to what the organization agreed to with the funder. It may, but does not necessarily, include information on the outcomes of the program.

The term “**evaluation**” is used broadly to cover an entire range of activities, including studies where the steps undertaken can specifically and with more certainty show that the results the program achieves are attributable to it and not to other factors. A typical definition of program evaluation is, “the systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of health or social interventions.”³

Evaluation research focuses on “causation,” proving that the activities provided through the program are the reason why change occurred for the people receiving the service. This requires considerably more time and effort, and this is the basis for the distinction in this guidebook between outcome measurement and evaluation:

1 Reisman, Jane and Judith Clegg. *Outcomes for Success!* (Seattle, WA: The Evaluation Forum, 2000), 9.

2 Ibid.

3 Rossi, Peter H., Mark W. Lipsey, and Howard E. Freeman. *Evaluation: A Systematic Approach*, 7th ed. Thousand Oaks, CA: Sage Publications, 2004.

Outcome measurement will explore what your program provides, what its intended impacts are, and whether or not it achieves them. It will *not* prove that the changes that take place are a result of your program.

People often get stuck in the outcome measurement process because of all the terminology. Is it an outcome, an accomplishment, an achievement, or a result? Is it a **goal** or an **objective**? Is it an **indicator** or a performance measure? Some people see goals and objectives as interchangeable; others see outcomes and objectives as the same. To keep it simple, this guidebook uses three terms consistently:

- **Goal** — broad statement of the ultimate aims of a program
- **Outcome** — the changes in the lives of individuals, families, organizations, or the community as a result of the program
- **Indicator** — the specific, measurable information collected to track whether an outcome has actually occurred

You should not let semantics stop you from moving ahead with your outcome measurement work. This guidebook uses the term “outcome measurement,” but it also may use the words “evaluate” or “evaluation” to cover the broad concept of exploring the changes that take place as a result of a program. The glossary in the appendices clarifies many of the terms used frequently in outcome measurement and evaluation work. You likely will need to “translate” some of the terms you encounter.

Some Words of Caution: Limitations of Measuring Outcomes

Program managers should remember that outcome measurement is not in and of itself a meaningful activity. Measurement is simply a means to help programs and nonprofits collect more information in support of continuous improvement of services. Here are some limitations to consider as you implement your outcome measurement plans:

- “Soft outcomes” may be more important than the movement towards metrics allows them to be. Building relationships between people or organizations or within communities is an important result of activities undertaken by many nonprofits, including capacity builders, but is hard to measure. The outcome measures then can be unsatisfactory, either because they are poor substitutes for these soft outcomes or because they ignore them altogether.
- Measurement cannot take the place of judgment and managerial decision making. The analysis and interpretation of data, as well as knowing how and when to share the information you collect, is not something that can be replicated by a statistical analysis tool or a spreadsheet. Managers will still need to apply critical thinking skills to the information gathered in the process.
- In some cases, the outcomes take years, if not decades, to materialize. Once again, the outcomes an organization chooses may simply be proxies for longer-term outcomes. On the other hand, the process of measuring outcomes forces us to consider what changes our activities have the most influence over. This must be balanced with activities that contribute to systemic changes that take years and decades to realize.
- Long-range planning is difficult, and because performance data does not speak to causality, managers are unable to definitively say how the agency’s activities contributed to the improvements. Moreover, outcome measurement is about the past. Decision making (budgets, policy, etc.) is about the future, where environments and other influencing factors may be changing.

Effective Outcome Measurement Practices

This guidebook is not meant to make you an expert in evaluation or research methodology. Rather, it will provide you with the basics you need to conduct outcome measurement. However, it is important to recognize that the outcome measurement work you undertake (and the outcome-related work of faith-based and community organizations who receive your support and assistance) must meet certain standards. Four program evaluation standards have been developed by the Joint Committee on Standards for Educational Evaluation.⁴ They are commonly accepted as guiding principles for any effective evaluation work, including outcome measurement. The standards include:

- **Utility** — the information generated must serve the needs of the intended users
- **Feasibility** — the process undertaken must be realistic, prudent, diplomatic, and frugal
- **Propriety** — the evaluators must behave legally, ethically, and with due regard for the welfare of those involved and affected
- **Accuracy** — the findings must reveal and convey technically accurate information

Without attention to these standards, your work and your results will not be credible or useful and, ultimately, will not help you continually improve your program. Additionally, you may find helpful these promising practices, which have been adapted from the Alliance for Nonprofit Management's survey of evaluation of capacity building:⁵

Begin with the end in mind. Ideally, evaluation planning starts when the planning for the project starts, and before the activities are underway. The entire first phase of outcome measurement in this guidebook discusses outlining your goals and the outcomes your program wishes to achieve.

Involve stakeholders. Involving stakeholders will help you gain buy-in for the process because the organizations and individuals involved will influence the definitions of success. It also helps to ensure that all stakeholders are on the same page when it comes to the purpose of the program.

Align closely with assessment. The assessment not only provides a baseline from which improvement can be measured, but a thorough and accurate assessment ensures that the capacity building activities identified will truly move the organization forward towards the intended outcomes.

Understand the context. Nonprofits operate in many different environments and industries, and so a one-size-fits-all approach to evaluation will likely fail. The evaluation design must take into account the context of the nonprofit. Is it an education program or a hospital? Does it operate in the city or in rural areas? Does it serve immigrants or particular minority communities? All of these contexts can influence what it means to be effective and how you can gather data. Customize the evaluation according to the context in which the organization operates and the outcomes toward which you are striving. Throughout this guidebook, you will find tips for customizing your efforts to the audience with whom you are working.

Use the evaluation for learning. The Alliance report puts it best when it says, "The ultimate purpose of evaluation should be focused on continuous learning and developing practices that move organizations toward greater effectiveness." The last phase of outcome measurement addressed in this guidebook is about reflecting on your outcome measurement process in order to improve your services as well as the outcome measurement process itself.

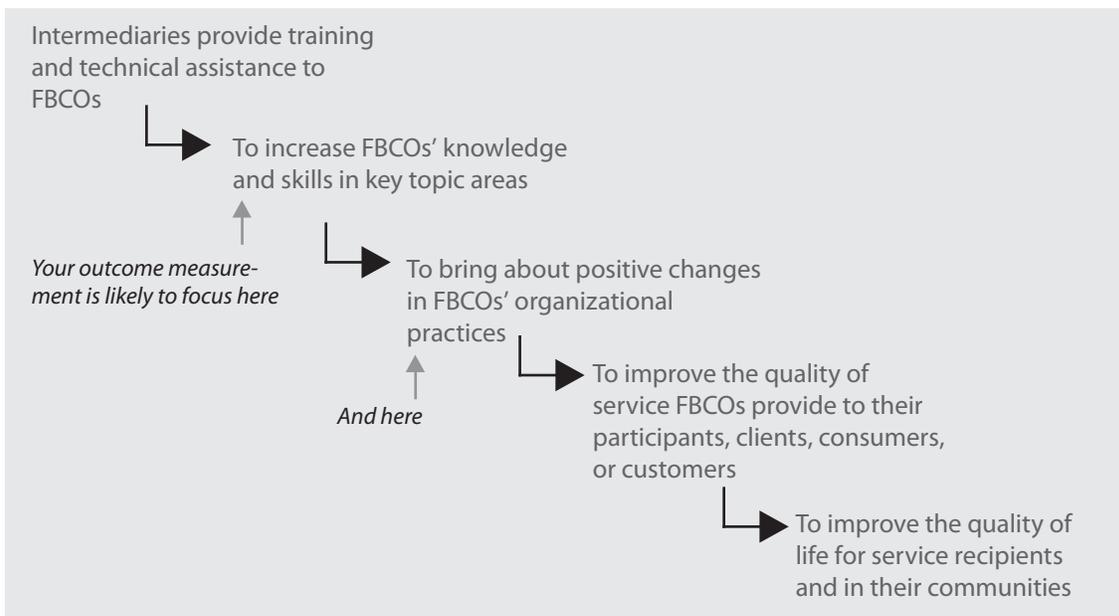
4 Yarbrough, Donald B., Lyn M. Shulha, Rodney K. Hopson, and Flora A. Caruthers. *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users*, 3rd ed. Thousand Oaks, CA: Sage Publications, 2010.

5 Linnell, Deborah. *Evaluation of Capacity Building: Lessons from the Field*. The Alliance for Nonprofit Management. 2003.

An Outcome Framework

This guidebook provides information useful to both capacity builders and the faith-based and community organizations (FBCOs) they serve. The outcome measurement steps are the same, regardless of who conducts the work, and regardless of the type of program. The examples included in the guidebook generally apply to functions performed by both groups, e.g., grants acquisition, grants management, or partnership development.

As an intermediary, your outcomes focus on increasing the capacity of grassroots organizations to deliver high-quality services. Ultimately, these organizations focus on improving the quality of functioning and life for the participants, clients, consumers, or customers they serve and the communities in which they operate. One way to visualize the relationship between you and the grassroots organizations with whom you work is shown below and demonstrates how your success on your outcomes contributes to the success of FBCOs.



GETTING STARTED

Your outcome measurement efforts will benefit greatly from some planning at the front end. Before you jump in, bring together the staff that will be involved in the outcome measurement process and discuss the questions outlined below. The better the planning, the more impact the outcome measurement will have on your organization.

1. Where Should We Focus?

In starting out, it's particularly important not to try to measure too much. Start slowly and learn from your experience. Don't try to perform outcome measurement at the same time. Pick one program or service as a beginning. Questions that will help you figure out where to begin include:

- Is a funder, board member, or staff person requesting that you look at a particular program?
- Do you have a new program with an unproven track record?

- Do you have an existing service with shaky performance?
- Do you have an effective program that you want to document as being so?

2. What Do We Want to Accomplish with Our Outcome Measurement Efforts?

Once you have decided what program to evaluate, explore the purpose for the outcome measurement processes. There are often a variety of reasons underlying the decision to carry out outcome measurement related to your programs. Which options below apply most directly to the program you intend to examine?

- To identify what constitutes success for a program and how it will achieve that success (often as part of a program design process prior to funding and implementation)
- To determine whether the program theory underlying the program is correct (often as part of a program redesign when new research indicates that the assumptions underlying the program cannot be substantiated)
- To describe how you intend to measure a program or service's impacts (often as part of a funding request)
- To report on a program's impact (often as part of a staff or board planning process)
- To determine whether the program has been effective in achieving its intended outcomes (often as part of a funder's request for evidence that the program merits continued funding)
- To determine whether to continue the program and allocate ongoing funding (often as part of internal budget-related priority setting)
- To make improvements in how the program is delivered (often as part of a program-wide or organization-wide quality improvement effort)

3. Who is on Our Outcome Measurement Team?

Once you have decided what you want to learn, it's time to make it operational. You can start by putting together a small team of people to manage the process. As you think about the roles of the team, consider who on your staff possesses the following skills:

- **Project coordination**, including laying out tasks in a sequence, informing other staff of their roles and assignments, providing assistance to people as they complete their parts of the evaluation process, and ensuring that the work is being done
- **Service or program knowledge**, including the ability to identify the relationship between the activities being provided and the intended impacts, and an understanding of the types of outcomes your program could achieve
- **Computer skills**, including expertise in formatting surveys and other data collection **instruments**, creating spreadsheets or databases, and entering data

Like any organizational activity, someone must be in charge of outcome measurement to make sure that it proceeds smoothly and according to a timeline. Among the team members, who should be designated as the team leader? Outcome measurement is more difficult to implement successfully when it is an unbudgeted add-on to a staff person's full-time job. Your outcome measurement work will go much more smoothly if you carve out the time for the lead staff person to manage these efforts. If you start with an outcome measurement focus on one program, ideally the team leader managing your evaluation will spend about one day per week to accomplish the work.

Do We Need to Get Help?

Often, smaller organizations feel that they do not have the staff capacity to implement outcome measurement activities. It is possible to contract out for help. You can hire an individual experienced in evaluation to design and/or implement your outcome measurement efforts. You may have someone on your board with the expertise you need. You may be able to seek assistance from a local university or community college or advertise for a volunteer with evaluation skills. You may not need help for the entire evaluation; perhaps you only require an outside person to assist your staff with some of the more technical aspects of the work, e.g., survey design, data analysis, etc.

CHECKLIST FOR OBTAINING AND USING OUTSIDE HELP

- Do we need specialized expertise to carry out the outcome measurement tasks? What specific expertise do we need?
- How much control should this individual have as compared to our outcome measurement team? How flexible is this person going to be?
- Who has the expertise and is available to help us—universities or other research institutions, consultants, students, or other volunteers?
- Do we know any other organizations who have obtained outside help? Can they provide any recommendations?
- What exactly do we need help doing? What specific tasks should we contract out?

4. How Can We Make This a Participatory Process?

Once you have clearly identified why you are carrying out your outcome measurement process and put a team in charge of carrying it out, you can decide on involving others in the process. Front-line staff may find outcome measurement threatening. Many have participated in outcome measurement as a response to poor performance or as a part of corrective action plans and may feel that their programs are under attack. You can allay people's concerns about outcome measurement by involving them in the process. Outcome measurement results have the greatest likelihood of being used (for the intended purposes you have identified above) if the people most directly affected have meaningful participation in the design and implementation of the process.

When deciding whom to invite into the evaluation process, consider the various tasks and roles needed and the following options:

- Program director
- Front-line staff directly involved in providing the service
- Participants or clients who use the program
- Staff from other programs or organizations working with you
- Volunteer researchers or students from a university, if available

You might also want to include local funders. They can be very helpful in identifying the programs to focus on and in articulating the key questions to answer.

5. Is the Timing Right?

Developing an outcome measurement plan will likely take a number of months. For organizations first learning outcome measurement, it may take as long as three years to develop a comprehensive, smoothly running outcome measurement system. The first year may be a learning year, with the second devoted to testing your systems. It may not be until the third year of data collection that staff sense the results are valid and meaningful. It may take you much less time, but be patient! United Way of America's *Measuring Program Outcomes* manual⁶ includes a section on timing and a template you can use to lay out a timeline for your outcome measurement work.

Timing is also important to funders. Make sure to consider when your funders' reporting cycles are so that you are producing outcome measurement results at a time that aligns with their requests for information about your program's impacts.

There are times when conducting outcome measurement may not be a good idea. If your organization is in crisis, people cannot give their attention to evaluation tasks. If the organization is in severe financial trouble, people must focus on solving that problem. If this is true for you, and if you have any leeway in terms of reporting requirements, consider delaying your outcome measurement work until it can get the attention it deserves.

6. What Resources Will We Need?

Outcome measurement does require resources. The categories below will give you some idea of what to consider when planning your budget:

- Staff time
- Consultants
- Communication (e.g., postage and telephone)
- Supplies and equipment
- Printing

A standard that many organizations use for estimating the cost of evaluation is 5 percent of a program's budget. However, once an outcome measurement system is put into place, it costs less to maintain over time.

Most organizations include the cost of outcome measurement as part of their administrative overhead, but funding to support the work may be available from other sources. Program funders are often willing to support a portion of the outcome measurement cost, particularly if you have built these costs into the funding request at the front end or the funder has placed a high priority on outcome measurement as part of its funding cycle. Some funders specify an amount or percentage of their grants that must be used for evaluation. In addition, organizations wishing to evaluate a program more deeply may seek funding specifically for that purpose.

6 United Way of America. *Measuring Program Outcomes: A Practical Approach*. Alexandria, VA: 1996.

CHECKLIST FOR GETTING STARTED

- What program will you start with?
- Who will be involved in the outcome measurement team for the planning and the implementation?
- Who will manage the process?
- Will you need outside help with your work?
- What additional costs do you anticipate and how can you budget for them?
- What resources are available to conduct outcome measurement?
- What is the timeline for your outcome measurement process?

Next, we will look at the key phases in the outcome measurement process and provide information to help you prepare for and conduct outcome measurement. The phases include:

1. Identify outcomes and develop performance measures
2. Create and implement a data collection plan
3. Analyze the data
4. Communicate the results
5. Reflect, learn, and do it again

PHASE 1: IDENTIFY OUTCOMES AND DEVELOP PERFORMANCE INDICATORS

This phase has two objectives:

1. Establish a shared understanding of what the program or project is and how it is supposed to work by completing a logic model including inputs, activities, outputs, and outcomes.
2. Create a set of measures that correspond to your logic model and can be used to assess the accomplishments of staff and project partners.

About Outcomes and Outcome Statements

Outcomes define what you want a program or service to accomplish. As a capacity building organization, your intended outcomes focus on the impacts or changes that the organizations experience as a result of your help. Sometimes providers of technical assistance believe that outcomes only relate to an “end-user,” e.g., clients, individuals, or families participating in a program provided by an organization with whom you are working. These are outcomes for the programs operated by the FBCOs, *not* for you as the organization providing technical assistance.

In this example, your client is the FBCO and your outcomes relate to increasing that organization’s capacity. Here are two questions to help you think about this important point:

- Who is the recipient of the technical assistance and/or funding I am providing?
- What is the intended impact or change I want this assistance to produce?

Another trap many organizations fall into when identifying outcomes is to describe what *they* have done (i.e., the activities they carried out) rather than the impact of these activities on the *client organization*.

When you review your draft outcomes, ask yourself these two questions:

- Do they focus on my client (the organization) and not the client’s client (the individual, group, community, etc.)?
- Do they describe the intended changes for the client organization, not what I am doing to achieve that impact?

As you become familiar with outcomes, remember that you are not going to measure very many. You’ll want to pick a couple that are the most directly connected to the assistance you are providing. You can use outcome statements to help you develop your outcomes. Writing an outcome statement can take a number of forms—the more straightforward, the better. Here are two typical formats you might use:

TYPE OF CHANGE:	IN WHAT:	FOR WHOM:
Increased	ability to raise funds	FBCOs receiving training

WHO:	CHANGE:	WHAT:
FBCOs receiving training	increase	ability to raise funds

Using Outcome Chains

Outcome chains can help organize your thinking about what you hope to achieve. They require you to put your program theory to work and articulate how your activities will bring about the impacts in the organizations with whom you are working. Outcome chains create a logical progression of the short-term, intermediate, and long-term outcomes that lead to your goals. The outcome framework depicted in the Measuring Outcomes Overview is an example of an outcome chain. Here is another example:

Training and Technical Assistance Leads to:

REACTION:	Feedback	How useful was the assistance provided? <i>(Was the client satisfied with the service?)</i>
↓		
LEARNING:	Changes in knowledge or skill	What gains were made in knowledge or skills? <i>(Did the client learn from it?)</i>
↓		
BEHAVIOR:	Changes in behavior	How has their behavior changed? How have they applied their knowledge? <i>(Did the client use it?)</i>
↓		
RESULT:	Improvements in the organization	What has changed for the organization as a result? <i>(Did it produce return on investment for the client?)</i>
↓		
GOAL ATTAINMENT:	Achievement of organizational goals	What results are evident in the community or individuals? <i>(Did it benefit the community where the client was located?)</i>

Sometimes organizations only measure the first level of outcome in the chain (in this case, by getting feedback on customer satisfaction). At other times, they may measure outcomes at all levels. Your decision regarding the “length” of your outcome chain depends to some degree on the type of and extent of the services you provided. If, for example, you sponsored a brief orientation about board development, you might only seek *reaction* from the participants in the session (the first outcome in the chain). If, however, you provided a series of workshops on board development and people attended them all, you would likely also measure *learning* and *behavior* (the next two levels on the outcome chain). If you additionally provided one-on-one technical assistance on board development, you might measure the *result* in terms of what changed in the organization. You probably will decide not to measure *goal attainment* at first, and instead focus on points earlier in the outcome chain.

Outcome Chain Example

Technical assistance in needs assessment and planning with clients (FBCOs) and their communities leads to:



OUTCOMES CHECKLIST

- Are the outcomes related to the “core business” of your program?
- Is it within your control to influence the outcomes?
- Are your outcomes realistic and attainable? Are your outcomes achievable within the funding and reporting periods?
- Are your outcomes written as change statements—will things increase, decrease, or stay the same?
- Have you moved beyond client satisfaction in your outcomes?
- Is there a logical sequence among your short-term, intermediate, and long-term outcomes?
- Are there any big “leaps” in your outcomes, i.e., gaps in the progression of impacts?

Outcome chains help develop a clearer picture of what you are trying to accomplish in your programs. To get a more complete picture of what outcomes your program’s activities and inputs will lead to, you can build a logic model.

Logic Models

Presented in a clear graphic format in precise language, the **program logic model** is a representation of the linkages between program activities and the changes those activities will produce.⁷ It helps you see whether there is any “logic” to your program—whether the connections between what you do and what you hope to achieve make sense. It provides greater clarity about and helps tell the story of your program.

You can create a logic model to organize a great deal of information about your technical assistance services. As you can see from the example below, a logic model is a great way to “tell your story.” It describes the resources you need, the activities you plan to carry out, the products of those activities, and the resulting impact you intend to achieve. It is often helpful to use the development of a logic model as a program planning or design tool, as it can help you think through what it will take to achieve the outcomes you have in mind.

This connection between the activities you provide and the outcomes you hope to accomplish is known as the program theory; it articulates the assumptions about the ability of certain activities to drive particular changes. Many nonprofit providers, including FBCOs, use logic models as a tool to reflect the program theory underlying their programs. One contribution that outcome measurement can make in your own organization is to demonstrate whether or not the program theory underlying your technical assistance efforts is valid. There is a great deal of research available regarding methods of achieving different types of organizational capacity. By referring to these studies, you can make sure you build a program design based on sound theory.

Here’s an example of a program theory, along with a sample logic model that shows how you might display a technical assistance program:

There is a great deal of research available that describes effective methods for disseminating knowledge and skills. This research reveals the relationship between the complexity of the information being transferred and the transfer method. The findings from the research indicate that the more complex the information being transferred, the more important in-person training and coaching become. In designing effective technical assistance for capacity building, you need to think about how to build such research findings into your practice.

Technical Assistance Program Goal: Increased FBCO Organizational Capacity

INPUTS OR RESOURCES	ACTIVITIES	OUTPUTS	OUTCOMES
Resources available to the program that allow and support provision of technical assistance, including money, staff, volunteers, clients, materials or equipment	The technical assistance methods carried out by staff	The product delivered or unit of technical assistance provided, usually described numerically, such as number of organizations served or number of hours of service	The changes that occur for FBCOs, as a result of the technical assistance provided
<ul style="list-style-type: none"> ▪ Funding ▪ Staff expertise ▪ Training manuals ▪ FBCO clients 	<ul style="list-style-type: none"> ▪ Provide one-on-one financial management coaching ▪ Conduct board development training workshops ▪ Develop content-specific manuals ▪ Provide phone and e-mail follow-up after training 	<ul style="list-style-type: none"> ▪ # of one-on-one coaching sessions ▪ # of hours of coaching ▪ # of training workshops ▪ # of staff trained ▪ # of hours of training ▪ # of manuals ▪ # of follow-up phone calls 	<ul style="list-style-type: none"> ▪ Increased FBCO compliance with financial management requirements ▪ Improved FBCO

7 Reisman, Jane and Judith Clegg. *Outcomes for Success!* (Seattle, WA: The Evaluation Forum, 2000), 28.

Look at the elements that make up the logic model and use the questions below to help you define the elements of your technical assistance program:

- **Inputs/Resources:** What **inputs** or ingredients do you need to operate your program? How many staff? How large a budget? How many FBCOs do you need to have as clients?
- **Activities:** What will you do? What methods will you use to deliver your technical assistance? What content areas will you cover? What level of assistance will you provide? (e.g., six one-on-one coaching sessions for executive directors)
- **Outputs:** What will be the tangible products of your assistance? To how many FBCOs will you provide coaching? How many sets of financial policy guidelines will you produce? How many hours of training will you provide? How many people will you train?
- **Outcomes:** What impact will your technical assistance have on your FBCO clients? What is reasonable to expect in the way of change?

Resist the temptation to create logic models that show a one-to-one relationship between each resource, activity, **output**, and outcome. It may be true that you can create a single output related to each activity, but it generally requires a comprehensive array of resources to deliver the activities, and it may take several or all of the activities to produce the outcomes.

Begin with the outcomes first; it's important to decide what you want to achieve in the way of impacts before you define what and how much of your technical assistance activities will be necessary to accomplish them.

View your program logic models as working documents that are subject to change as you learn more about what works. You'll find the logic model to be a useful program planning tool in incorporating changes and improving your technical assistance program over time.

TIPS FOR CREATING A LOGIC MODEL

Similar to the development of outcome chains, creating logic models works well as a group activity.

- ✓ Take large pieces of newsprint and hang them on the wall.
- ✓ Write down one component of the logic model on each sheet (i.e., resources, activities, outputs, outcomes) in the logic model's order.

LOGIC MODEL AND PROGRAM THEORY CHECKLIST

- What resources are available to conduct outcome measurement?
- What is the timeline for your outcome measurement process?
- Have you included all of the resources you will need? Are the following items listed?
 - ✓ Service providers: staff, volunteers
 - ✓ Program setting: community settings, agency facilities
 - ✓ "Service technologies": curriculum/class material, treatment plans
 - ✓ Funding sources: private or public funding, donations, fee for service
 - ✓ Participants: client organizations
- Are all of the activities included and described in some detail (e.g., number and length of coaching sessions, types of computer training)?
- Have you described an output, product, or unit of service for each activity? (e.g., number of FBCOs trained, number of hours of training each organization received, type of computer assistance, etc.)?
- Have the participants been counted and described in your output column? (e.g., nine FBCOs, fifty-four staff, eighty-eight board members, etc.)?
- Is the intensity of the activities appropriate for the type of organization with whom you are working? Organizations with greater needs require more assistance, e.g., six four-hour coaching sessions on fundraising rather than two four-hour sessions. (You will need to develop your outcomes before you can answer this question; that's one reason to create the outcomes first. If you don't know your clients' needs or abilities at the outset, you may not know the answer to this question and will have to come back to it later.)
- Is the duration of the activities appropriate for the type of organization? Higher-needs organizations will take longer to absorb and implement the improvements you are helping them make, e.g., you may be providing coaching for twelve months rather than six. (Once again, you'll need to define the impacts you hope to make before you can assess how much assistance your clients will require, and you may need to come back to this question if you don't know about your clients' needs or abilities.)
- Are your outcomes directly related to your activities (i.e., is it possible to achieve the results you have listed with the type and amount of activities you are planning to deliver)?
- Do the outcomes address changes in knowledge, perceptions, attitude, skills, or behavior?

The completion of your logic model signifies your success in accomplishing the first phase in conducting outcome measurement. You have articulated what you want to accomplish and how you plan to go about making it happen. Further, your outcome chain describes why you believe your planned technical assistance activities will produce the impacts you intend. A sample logic model to assist you with creating logic models is featured in the appendices of this guidebook.

Developing Performance Indicators

The next step in putting together an effective outcome measurement system focuses on deciding how you are going to make your intended outcomes measurable, i.e., defining a set of performance measures or indicators.

As written, outcomes are usually too broad to enable data collection tools to be developed directly from them. Remember, the purpose of outcomes is to describe the intended impact the client organization experiences. For this reason, indicators are used to serve as a bridge between intended outcomes and the actual data collection process. Indicators enable you to determine whether the FBCOs you work with have, in fact, changed in the specific ways that indicate your intended outcome has occurred.

OUTCOME/INDICATOR EXAMPLE

Outcome:	Indicator:
Increased awareness of basic financial management practice	<ul style="list-style-type: none">• Number and percent of organizations that can articulate the components included in financial management• Number and percent of organizations' boards that perform at least quarterly monitoring of the agency's financial performance• Number and percent of FBCOs that comply with board-approved policies for capital and operating reserves

Indicators must be specific and observable. They answer questions like how you will know when changes have occurred and how you will know when you have achieved the outcomes. Thinking ahead to possible data collection methods will tell you if your indicators are specific enough. Ask questions like these to determine whether your indicators will work:

- ✓ How can I see the change? (*Through what kind of observation?*)
- ✓ How can I hear the change? (*Through interviews? Focus groups?*)
- ✓ How can I read the change? (*Through surveys? In records?*)

For example, look at this outcome:

- Outcome:** FBCOs have increased fund development capabilities and opportunities.
- How will you know whether an organization has increased its fund development capabilities? What will you look for? What questions will you ask? When you think about becoming better at fund development, what comes to mind?
 - Similarly, what will you include as a measure for increased fund development opportunities? What is a fund development opportunity? A phone call to a potential donor? A fundraising event? Subscribing to a philanthropy magazine? How will you decide whether the organization is pursuing more fundraising than it used to?
 - Overall, what will you look at to see whether your technical assistance has helped the organization bring in additional funds through its fund development efforts?

In order to serve effectively as a bridge to data collection, indicators must be *specific* items of information that describe *observable*, measurable characteristics or changes in corresponding outcomes. Indicators must be measures that can be seen, heard, counted, reported, or enumerated using some type of data collection method.

Tips for Creating Appropriate Indicators

- **Define the characteristics of an outcome as a way of identifying possible indicators.** For example, the characteristics of “increased fund development capabilities” could include the amount of money obtained in additional support and the diversity of those funding sources. The specific measures related to these two characteristics for “increased fundraising ability” might be 1) the number and percent of FBCOs who raise additional funds this year as compared to last year and 2) the number and percent of FBCOs who show an increase in the number of sources of funding for their programs.

- **Use “if...then statements” to identify indicators.** Look for indicators that are *indicative* of an outcome rather than a predictor or a result of an outcome. If the relationship between an outcome and its proposed indicator sounds like an “if...then statement,” then it is probably not the right indicator for that outcome. For example, “if” an organization attends grant writing training, “then” it is more likely to bring in additional grant funding. In this example, attending grant writing training is not an indicator for increased fund development capabilities, but may rather be a predictor of increased success. A more indicative indicator of increased grant funding would be “the number and percent of organizations whose budgets show an increase in the number of grants and/or an increase in the amount of support from grants.”
- **Apply the “means that” rule in assessing your indicators.** In theory, the accomplishment of an indicator “means that” you have achieved an outcome. For example, if an organization has completed, submitted, and obtained approval for 501(c)(3) status (indicator), it “means that” it has a stronger organizational structure (outcome). In contrast, an organization having an expanded service area for more clients (indicator) does not mean the organization has improved its coordination and/or collaboration in service with others (outcome).
- **Develop one to three indicators per outcome.** One to three indicators is usually a realistic number for each outcome you have identified. Some straightforward outcomes can be quantified easily through the use of only one indicator. Other more complex outcomes will necessitate two or three indicators.
- **Distill to the fewest outcomes possible.** As you look at what indicators you need to adequately describe your progress in achieving your intended outcomes, it’s important to use the fewest number of outcomes possible. It takes time and money to gather and analyze the data for each one. What’s important is not quantity but quality. What’s the best way to see, hear, or read about the change?
- **Take into account the feasibility of collecting data for the measurement.** Select the indicator that is most feasible for staff to measure within the time and financial resources available to you.
- **Identify the most useful indicators.** Select the indicator that is most useful to you and gives you the most useful information about the outcome.

Performance Targets

When developing indicators, you may wonder about setting a specific desired level of achievement. This is called a “**target**.” In addition, you may be interested in comparing your performance on the indicator to other organizations or time periods. This approach is known as a “**comparative standard**.” Here is how outcomes, targets, and comparative standards fit together in outcome measurement:

Outcomes statements — describe an increase, decrease, maintaining, etc. of a behavior, skill, knowledge, etc. *Example: Improve fund development capabilities*

Target statements — specify a particular level of achievement related to the outcome. *Example: Raise \$20,000 in foundation funds during 2004*

Comparative statements — include comparative targets, usually related to standards, other time periods, or organizations. *Example: Increase the organization’s funding from non-governmental sources by 30% as compared to last year*

Targets and comparisons will be useful to you over time as you gain a better sense of where your client organizations are at the beginning of your capacity building process. It’s important to establish a **baseline** for the indicators you plan to measure over time. A baseline is information collected to give you a starting

place—something to measure against in the future. A baseline will allow you to set reasonable targets and measure the FBCOs’ progress against them. For organizations that have not collected data on indicators in the past, the first round of data collection will establish the baseline.

Similarly, comparisons are useful when you want to measure your performance against something—either another time period, another organization, or an established set of standards. Although there are comparative standards in some fields (e.g., health—where infant mortality rates and other morbidity rates have been established nationally), there are few standards in human services that organizations can use for comparison purposes. Some national human service organizations are beginning to establish standards specific to their own programs. More likely, however, you will use your previous performance as a standard, comparing your success last year to this year, or this year to next year.

At the outset, it’s important to stick to the basics of outcome measurement—defining outcomes and the indicators that make them measurable. You’ll have time to move on to targets and comparisons later. A worksheet to assist you with identifying indicators is included in the appendices.

PERFORMANCE INDICATORS CHECKLIST

- Are the outcomes related to the “core business” of your program?
- Do your indicators make sense in relation to the outcomes they are intended to measure?
- Are your indicators directly related to the outcome? Do they define the outcome?
- Are your indicators specific?
- Are your indicators measurable or observable? Can they be seen (i.e., observed behavior), heard (i.e., participant interview), or read (i.e., client records)?
- Is it reasonable that you can collect data on the indicators?
- Is it likely within your resources to collect data?

PHASE 2: CREATE AND IMPLEMENT A DATA COLLECTION PLAN

This section discusses how to plan for and implement the collection of data. You will consider what data collection methods are most appropriate and get the best data, and how to manage and ensure the integrity of the data you collect.

Data Collection Methods

Your data collection will include attention to all the elements of your logic model—what resources you had available, what activities you actually provided, how many of each output you delivered, and to what degree you accomplished your outcomes. For all of these elements, except the outcome and indicator data, your organization’s records should provide the information you need. However, in collecting indicator data, you have other options. You are likely to use one or more of four methods for collecting data—surveys, interviews, observation, and record or document review. In selecting the best method for data collection, you will need to consider the advantages and disadvantages regarding:

- Type of information needed — some outcomes and indicators lend themselves to one particular method; in other cases, data could be collected in more than one way
- **Validity** and **reliability** — some methods generate more accurate and consistent information than others
- Resources available —including staff availability and expertise, time, and money
- Cultural appropriateness — using the broadest definition of culture to ensure that the methods fit language, norms, and values of the individuals and groups from whom you are collecting data

Surveys

Surveys are standardized written instruments that can be administered by mail, e-mail, or in person. The primary advantage of surveys is their cost in relation to the amount of data you can collect. Surveying generally is considered efficient because you can include large numbers of people at a relatively low cost.

There are two key disadvantages to surveys. If a survey is conducted by mail, response rates can be very low, jeopardizing the validity of the data collected. There are mechanisms to increase response rates, but they will add to the cost of the survey. Later in this section, we include some tips for boosting response rates. Written surveys also provide no opportunity for respondents to clarify a confusing question, again creating a problem with the validity of responses. Good pre-testing of a survey instrument can reduce the likelihood that problems will arise.

Examples:

- Survey to track grassroots organizations' use of and satisfaction with technical assistance services you provide
- Survey of all organizations receiving technical assistance to learn about changes in their fundraising tactics and the results of their efforts to raise more money

Interviews and Focus Groups

Interviews also use standardized instruments, but they are conducted person-to-person either in person or over the telephone. In fact, an interview may use the same instrument created for a written survey, although interviewing generally provides the advantage of having the opportunity to explore questions more deeply than with a written survey. You can ask more complex questions in an interview, since you have the opportunity to clarify any confusion. You also can ask respondents to elaborate on their answers, eliciting more in-depth information than a survey can provide.

The primary disadvantage of interviews is their cost. It takes considerably more time and, therefore, more money to conduct telephone and in-person interviews. This usually means that you are able to collect information from far fewer people. The reliability of interviews can also be problematic if interviewers are not well-trained, since they may ask questions in different ways or otherwise bias the responses.

Examples:

- Interviews at different grassroots organizations to learn about the way in which they are applying new knowledge of partnership development
- Interviews with individuals within an organization to explore their perceptions of changes in capacity and ability to deliver services

A focus group is a particular type of interview conducted with a small group of people to obtain information in a defined area of interest related to your outcomes and indicators. While interviews with individuals are

meant to solicit data without any influence or bias from the interviewer or other individual, focus groups are designed to allow the participants to discuss the questions and share their opinions. This means that people can influence one another in the process, stimulating memory or debate on an issue. The advantage of focus groups lies in the richness of the information generated. The disadvantage is that you can rarely generalize or apply the findings to your entire population of participants or clients. Focus groups are often used prior to creating a survey to test concepts and wording of questions. Following a written survey, focus groups are used to explore specific questions or issues more thoroughly.

Examples:

- Structured meetings with staff in a faith-based or community organization to learn more about their grants management practices, what worked during the year, and what did not
- Structured meetings with staff from several organizations to explore their use of computer technology for tracking financial data

Observation

Observations can be conducted of individual behaviors or interactions among individuals, of events, or of physical conditions within a site or facility. They require well-trained observers and detailed guidelines about whom or what to observe, when and for how long, and by what method of recording. The primary advantage of observation is its validity. When done well, observation is considered a strong data collection method because it generates first-hand, unbiased information by individuals who have been trained on what to look for and how to record it. Observation does, however, require time for development of the observation tool, training of the observers, and the data collection, making it a more costly data collection method than some of the others.

Examples:

- Observation of individuals participating in training to track the development of their skill in the training topic
- Observation of community meetings sponsored by grassroots organizations to learn about their partnership-building techniques and collaborative behavior

Record Review

Record or document review involves the systematic collection of needed data from internal, organizational records or official records collected by other groups or institutions. Internal records available to a capacity builder might include financial documents, monthly reports, activity logs, purchase orders, etc. The advantage of using records from your organization is the ease of data collection. The data is there and no additional effort needs to be made to collect it—that is if the specific data you need is actually available and up to date.

If the data is available, record review is a very economical and efficient data collection method. If the data you need is not available, it is likely well worth the time to make improvements to your data management system so that you can rely on internal record review for your outcome measurement work. Just a few changes to an existing form can turn it into a useful data collection tool. And just a small amount of staff training can increase the validity and reliability of internally generated data.

Examples:

- Sign-in logs from a series of workshops to track attendance in training, measuring consistency of attendance as an indicator of organizational commitment to learning
- Feedback forms completed by workshop participants to learn about satisfaction with training provided

Official records can include Federal, state, or local government sources such as the U.S. Census, health departments, law enforcement, school records, assessor data, etc. If the data available in official records are relevant to your outcomes and the indicators are accessible, then official record review is a very low-cost method for collecting data; but the questions about relevance and accessibility are major ones.

Checklist for Selecting Data Collection Methods⁸

This checklist can help you decide which data collection methods are most appropriate for your outcome measurement.

SURVEYS		
1. Do I need data from the perspective of the participant, client, beneficiary, or customer?	Yes	No
2. Do I have a systematic way to get it from these individuals?	Yes	No
3. Do I need data that are standardized so that statistical comparisons can be made? (For example, will I need to report percents or other statistics?)	Yes	No
4. Will participants be able to understand the survey questions? (Consider age, cultural backgrounds, etc.)	Yes	No
5. Do participants have the necessary knowledge or awareness to <i>accurately</i> answer questions about the outcomes?	Yes	No
<i>If you answered Yes to questions 1 through 5, surveys may be appropriate for collecting data on your outcomes and indicators.</i>		

INTERVIEWS		
6. Are more in-depth answers necessary to adequately measure the indicators or to get information on what is needed or what should change?	Yes	No
7. Will it be necessary for someone to personally ask participants questions (either on the phone or in person) in order to collect the information related to this outcome? (Consider age, cultural background, as well as state of mind or receptivity of participants.)	Yes	No
<i>If you answered Yes to questions 6 and 7, interviews may be appropriate for collecting data on your outcomes and indicators.</i>		

8 Reisman, Jane and Judith Clegg. *Outcomes for Success!* (Seattle, WA: The Evaluation Forum, 2000), 94–97.

OBSERVATION		
8. Is it difficult to accurately measure the indicators by asking people questions about opinions or perceptions?	Yes	No
9. Can this outcome or indicator be assessed accurately by someone trained to observe it in action (i.e., can something actually be observed)?	Yes	No
10. Do you have the staff resources for someone to observe events, conditions, interactions, or behaviors?	Yes	No
<i>If you answered Yes to questions 8, 9, and 10, observation may be appropriate for collecting data on your outcomes and indicators.</i>		

INTERNAL RECORD REVIEW		
11. Do you have individualized records, reports, logs, or other systematic ways that you track things in your program or services?	Yes	No
12. If an information system exists, are the data consistently entered into it in a timely way?	Yes	No
13. If a system exists, can information be extracted from it easily?	Yes	No
<i>If you answered Yes to questions 11, 12, and 13, internal record review may be appropriate for collecting data on your outcomes and indicators.</i>		

OFFICIAL RECORD REVIEW		
14. Do official records exist that track the data you need on your outcomes and indicators?	Yes	No
15. Are the data accessible to you (i.e., will it be possible to get the cooperation of outside agencies or institutions in order to get access to official records)?	Yes	No
<i>If you answered Yes to questions 14 and 15, internal record review may be appropriate for collecting data on your outcomes and indicators.</i>		

Validity and Reliability

Validity and reliability are two critical concepts in implementing effective outcome measurement systems. Validity is the accuracy of the information generated. The validity of a data collection instrument is how well it measures what it is supposed to measure. Putting in the time to create good instruments, carefully considering what is being measured, and pre-testing the instruments will help increase their validity.

Reliability refers to consistency, or the extent to which data are reproducible. Do items or questions on a survey, for example, repeatedly produce the same response regardless of when the survey is administered or

whether the respondents are men or women? Bias in the data collection instrument is a primary threat to reliability and can be reduced by repeated testing and revision of the instrument.

You cannot have a valid instrument if it is not reliable. However, you can have a reliable instrument that is not valid. Think of shooting arrows at a target. Reliability is getting the arrows to land in about the same place each time you shoot. You can do this without hitting the bull's-eye. Validity is getting the arrow to land on the bull's-eye. Lots of arrows landing on the bull's-eye means you have both reliability and validity. Pre-testing is critical to ensuring that your data collection instrument will produce valid and reliable results.

Instrument Development and Pre-Testing

This guidebook does not include information on how to create the instruments you'll need for data collection, although there are many resources available to help with this task (see the appendices). The appendices also feature a checklist you can use to assess survey and interview instruments and two examples of surveys related to organizational capacity building.

You need to pre-test each data collection instrument you create. To pre-test, you try out your tools with a small group of people who are representative of those you intend to include in your data collection. For a survey or interview, you must include people who are like those who will complete the survey or interview during your actual data collection. For observation, you must try out your form with individuals who will complete it during data collection. For record review, you must test the record extraction form with the individuals who will do the work to pull information from your records.

During the pre-test phase, you want people's help in checking the wording of questions, the content, the clarity of instructions, and the layout and format. You'll also want to know how long it took and whether any problems arose. Most importantly, you want to see if you get the responses you expect and the information you need.

Deciding When and How to Collect Data

Once you have identified the data collection methods you intend to use, you need to decide when you will collect the data and how often. Then consider the procedures you need to put in place to ensure that your outcome measurement system is sustainable and produces quality data.

Frequency and Scheduling of Data Collection

The following table describes the five approaches or designs you are likely to use for your data collection.

Data Collection Designs

TYPE OF DESIGN	COLLECTION FREQUENCY	EXAMPLE
Post-only Measures	Data are collected once — at the end of the program, service, or activity	<i>Level of participant knowledge on a survey after a training workshop</i>
Pre/Post Measures	Data are collected twice — at the beginning, to establish a baseline, as well as at the end of the program	<i>Comparison of an organization's documented fundraising success before and after receiving technical assistance</i>
Time Series	Data are collected a number of times — during an ongoing program and in follow-up	<i>Monthly observations of an organization's collaboration meetings to track changes in partnership development and communication</i>
Measures with a Comparison Group	Data are collected from two groups — one group that receives the intervention and one that doesn't	<i>Comparison of data on skill development from individuals who participated in training and those who have not yet taken your workshop</i>
Measures with a Comparative Standard	Data are collected once — at the end of the program, service, or activity, and are compared with a standard	<i>Comparison of this year's data on organization's success in fundraising as compared to last year's data</i>

The first three approaches in the table are fairly self-explanatory; the last two approaches need a little discussion. “Comparison groups” can be very useful in demonstrating the success of your intervention. The main question is whether you can find a group of people or organizations that is just like the group with whom you are working. In order to provide a valid comparison, the two groups must have the same general characteristics. A similar group may be difficult to find. However, if you are working with different groups at different times, and the groups are similar, this approach may work for you. Or if, as in the example of using a comparison group shown above, you have people on a waiting list for training who are similar to the people who have trained already, they may make a good comparison group. Or perhaps you might compare FBCOs receiving subawards with those who applied but did not receive them.

“Comparative standards” are standards against which you can measure yourself. There are standards of success in some fields, e.g., health mortality and morbidity rates, student achievement scores, teen birth rates. For intermediaries, however, there are unlikely to be many standards regarding your program outcomes or indicators. But you could compare your results for one time period to an earlier one, as shown in the example of using comparative standards in the table above. You would collect data for the first time period as your baseline and use it as your standard in the future.

CHECKLIST ON DATA COLLECTION DESIGN

When considering which approach is best, these questions may help you make a decision:

- When would change occur—immediately, gradually, or over a long period of time?
- Are there milestones that can be measured along the way to the outcome you are trying to achieve?
- What is the frequency of contact with the organizations with which you are working—once, weekly, monthly, or at some other interval?
- When will data be available? (When are some of your internal records available?)
- Are there any groups that might serve as comparison groups?
- Do you have baseline data you can use as a standard?
- Are you accountable for short-term outcomes or longer-term impacts?

Data Collection Procedures

To finish your planning work, a number of procedural issues must be considered. You and your team will need to think about the following questions as you formulate and implement your data collection plan:

1. **Who will collect the data?** How will you recruit these data collectors? What steps will they need to take to collect the data? How will you train them? Finally, who will be responsible for monitoring the data collection process to ensure that you are getting what you need? It's important to answer each of these questions during your planning. You don't want to discover halfway through the process that your three-month follow-up surveys were not mailed out because you didn't identify who would do so!
2. **Have you prepared your clients for data collection?** Do they know that you will be collecting data, either at the time of service or in follow-up? Do they understand why it is important to you and how you intend to use the data? Organizations often have outcome reporting requirements themselves, so they are usually responsive if they have been alerted to your needs ahead of time. Advising them in advance about your data collection plans will help increase their willingness to participate during implementation.
3. **How will you ensure confidentiality and informed consent?** Anonymous and confidential do not mean the same thing. **Anonymous** means you do not know who provided the responses. **Confidential** means you know or can find out who provided the responses, but you are committed to keeping the information to yourself.

You must ensure that you protect the confidentiality of any individual's data or comment. It is easy to make your surveys anonymous, but if you want to track people over time, you'll likely need to attach ID numbers to each person from whom you collect, keeping a list of the names and numbers in a locked file.

It is important to inform people that you are measuring your program's outcomes and may use data they provide in some way. You must let them know that their participation is voluntary and explain how you will maintain the confidentiality of their data.⁹

9 A sample consent form can be found in United Way of America's *Measuring Outcomes: A Practical Approach* <http://www.liveunited.org/outcomes/resources/mpo>.

4. **How will you ensure quality of data?** As data is collected and entered into the storage mechanism, checking for errors and data quality is an important step that is easily overlooked. Build in time to review data and follow up about discrepancies in your overall timeline; the more data you collect, the more time you will need to assure its quality. Here are some typical strategies for assuring the quality of data:
- **Double entry.** This entails setting up a system to collect data twice, and then comparing for discrepancies. This can be costly and time-consuming, but it is the most thorough method for quality control.
 - **Spot checking.** This entails reviewing a random sample of data and comparing it to the source document for discrepancies or other anomalies. If discrepancies are found, the first step is to identify any patterns (data entered during a particular time period or by a specific staff person; data associated with a particular beneficiary organization; a specific type of data that is incorrect across many records, e.g., if all data for additional persons served at an organization was formatted as a percentage instead of as a whole number). The capacity builder may need to review all the data entered, especially if there is no discernible pattern to the errors.
 - **Sorting data** to find missing, high, or low values. If you are using a database or spreadsheet function, identifying outliers—those pieces of data at either extreme—is very easy, whether through the use of formulas or sorting functions.
 - **Use automation**, such as drop-down menus. Automating a data collection provides a uniform way to report information and makes sorting and analyzing data much easier. For example, organizations reporting the number of additional persons served will all use the same language to report the outcome, whereas without such automation the language could vary significantly from report to report. Additionally, more sophisticated forms can pre-populate performance goals from an existing database, which reduces data entry errors made by those filling out the forms.
 - **Format database** to accept only numbers. Whether organizations are filling out forms directly or your staff is entering data from a handwritten form, formatting your data fields to accept only numbers reduces errors related to typos.
 - **Review data for feasibility.** This strategy requires that a staff person who is familiar with the organization’s capacity building interventions and who has a good eye for detail review the data collected and identify anomalies. Some of these anomalies may not appear with general sorting.
 - **Discuss data discrepancies** with the organization. If discrepancies remain unexplained after implementing any of these quality assurance mechanisms, take the data back to the organization for discussion and clarification.
5. **How can you boost response rates?** Low response rates can threaten an outcome measurement effort. Following up with organizations that are no longer actively participating in a capacity building program is especially difficult. Listed here are some strategies to increase response rates:
- **Tie data collection to project milestones.** Throughout the course of the capacity building relationship, it is relatively simple to require organizations to report desired data. For example, an evaluation could be due as a requirement to moving on to the next phase of the project, such as releasing funds for a capacity building project or approving a consultant to begin work. However, once the organization exits the capacity building program, the capacity builder loses this leverage.
 - **Conduct an exit interview.** Once the entire engagement is complete, the capacity builder can schedule an exit interview, on the phone or in person, to collect data. Participation in this interview can be mandated in a memorandum of understanding. The organization will need to know what data to prepare for the interview. The advantage of this method is that an exit interview is close

enough to the intervention that the organization may still be invested in maintaining its relationship with the capacity builder and follow through on the commitment. However, the organization may not have realized all its possible outcomes, and therefore the data may not capture some of the ripple effects, where outcomes are realized after the data has been collected.

- **Stay in touch.** By holding monthly meetings or conference calls with organizations after they exit the program, the capacity builder can maintain more informal connections. The organizations have access to advice and support and may be more likely to participate in a follow-up data collection effort. Establishing a community of practice among organizations so that they have even more reason to be in touch with each other and your organization is one way to implement this strategy.
- **Provide the outcome data to the organization.** Offer organizations a short summary report card of the data you collect from them and demonstrate how it can be used as a marketing tool. This summary can prove invaluable to a program, and it may increase the number of responses you get to your data surveys. If you can use the merging functions available in most productivity software like Microsoft Word and Outlook, generating report cards for tens or even hundreds of organizations may take just a few hours.
- **Offer multiple collection methods.** Be available to complete the survey on the phone with the organization. Be available to go to the organization's headquarters and conduct the survey in person. Be prepared to offer language translation services if necessary, offer the survey electronically (at least as an attachment to an e-mail), or mail the survey with a stamped envelope. Response rates will increase the easier it is for an individual to complete the survey, and preferences will vary across individuals.
- **Be culturally competent.** Capacity builders may take great steps to ensure that training and technical assistance is culturally appropriate, and this should be done for data collection efforts as well. Moreover, if you are engaging a third party to collect data—a consultant or a team of interns, for example—remember that being a third party means that they have not had the benefit of getting to know an organization and its staff through the course of the capacity building engagement. Language barriers, cultural differences, and individual preferences can influence whether you are likely to get a response.
- **Introduce your external data collectors.** If you are working with those third parties, introduce them to the organizations you are working with early on. If maintaining a relationship helps improve response rates, then so too will the lack of a relationship hurt response rates. As a caveat, be sure to maintain confidentiality about the results, especially if the third party is collecting direct feedback about your services.

Finally, when you put together your outcomes, indicators, data collection methods, frequency and schedule of data collection, as well as procedures, you will have created your plan for outcome measurement. Worksheets to assist you with your evaluation planning are in the appendices of this guidebook. The bottom line is that how you collect your data and your attention to those you're collecting it from will ensure that you have optimum quality data to work with in the analysis phase.

PLANNING CHECKLIST ¹⁰

- Is the data collection method (survey, interview, observation, internal records, official records) appropriate for the outcomes and indicators?
- Does the plan rely on external sources of data or require collaboration with other agencies? If so, will it be possible to get the data?
- Have the tools/questions been pre-tested?
- Are the measurement approaches and tools culturally relevant?
- Can the approach/tools likely be implemented with available program resources?
- Does the data collection schedule include time points that follow completion of services or program activities?
- Does the frequency of data collection match time points when realistic progress can be expected from participants?
- Does the frequency of data collection match the level of contact with participants (e.g., greater frequency with greater level of contact)?
- Are the roles and responsibilities clear for all the staff involved in collecting data?
- Is there a staff person (or staff team) responsible for managing and monitoring the process to assure the work is completed on time and is of high quality?

SAMPLE DATA COLLECTION PLAN FOR OUTCOME MEASURES

OUTCOME	INDICATOR	WHAT DATA COLLECTION METHOD?	WHEN WILL DATA BE COLLECTED?	WHO WILL COLLECT DATA?	WHAT WILL THEY DO?	HOW WILL DATA COLLECTED BE MONITORED?
Increased ability of FBCOs to raise funds for program services	Number and percent of organizations who put new fundraising practices in place	Survey of organizations receiving fundraising technical assistance	Once — three months after completion of technical assistance services	Technical assistance staff	<ul style="list-style-type: none"> ▪ Mail survey to organizations three months after the completion of technical assistance services ▪ Send reminder postcards to organizations who do not return the survey within two weeks 	Program manager will track completion of services, mailing of surveys, and response rate to the survey and notify staff if additional follow-up is needed

¹⁰ Reisman, Jane and Judith Clegg. *Outcomes for Success!* Seattle, WA: The Evaluation Forum, 2000.

PHASE 3: ANALYZE THE DATA

Once data is collected, the next step is to analyze the data. Data analysis is more than figuring out ways to make beautiful pie charts and other graphics. It is about looking at the information you have collected and asking yourself what it all means. This is a prime example for when data does not substitute for judgment or managerial decision making. Once you have the data, it is up to you to make use of it to inform decisions about your programs.

Moreover, analysis of data can be as simple or as complex as the tools you decide to use to apply to the data. You can simply count, sort, and order the pieces of data. You can perform statistical tests to determine the relationship between two sets of data, or use the information to find patterns that allow you to predict future behavior. Whichever tools you use for analysis, be sure to continue to ask yourself, “To what end?” And make sure that the analysis is used in service of the mission of your program.

Tools for Analysis

Common Descriptive Statistics. It is likely that most analysis will use measures of total, arithmetic mean, and standard deviation. However, it is useful to be familiar with all of the descriptive statistics.

MEASURE	DESCRIPTION	CONSIDERATIONS
TOTAL	Self-explanatory	The total by itself is less useful than if used by comparison to show increases, e.g., the total number of organizations served, the total number of youth participating in a program, or the total amount of additional funds raised by an organization.
ARITHMETIC MEAN	The arithmetic mean is the sum of the observations divided by the number of observations. It is the most common statistic of central tendency, and when someone says simply “the mean” or “the average,” this is what they mean.	This measurement is easily skewed by outrageous outliers. For example, if you are reporting the average additional funds that organizations raised, if one organization was able to raise many times the amount of others, the average goes up significantly.
MEDIAN	The mean is found by sorting all the data from lowest to highest, and taking the value of the number in the middle. If there is an even number of observations, the median is the average of the two numbers in the middle.	If the distribution of data is very skewed, the median is a more useful tool to indicate the central tendency because it is less influenced by outliers.
MODE	The mode is the common value in the data set.	Mode is particularly useful when you have data that is grouped into a small number of classes, e.g., the type of organization you are serving, or in what county the organization operates. The mode would simply be the type of organization you serve most frequently, or the county where the largest number of organizations operate.

MEASURE	DESCRIPTION	CONSIDERATIONS
STANDARD DEVIATION	A measure of variability or dispersion of a data set.	A low standard deviation indicates that the data points tend to be very close to the same value (the mean), while high standard deviation indicates that the data are spread out over a large range of values. For example, if in looking at the amount of money each organization devoted to hiring consultants, the range of dollars spent was \$40-\$200, the standard deviation would be low compared to a range of \$20-\$300,000.
RATIO	A ratio is an expression that compares quantities relative to each other.	Often, ratios are expressed as an “input” relative to an “output” to describe the cost—financial or otherwise—that the activity, output, or outcome requires.

Inferential Statistics. Where descriptive statistics help you understand the data as you have it, inferential statistics help you draw conclusions that may be more widely applicable beyond the specific data set with which you are working. Essentially, inferential statistics allow you to “infer” additional information. Two common operations are correlation and regression.

“Correlation” uses statistical formulas to calculate the relationship between two variables. For example, two variables might be the number of hours of technical assistance an organization receives and the capacity index score at the beginning or end of the intervention, or even the difference between the two. You might expect that organizations with lower capacity index scores required more technical assistance, or that those receiving more hours of technical assistance saw more increases in capacity index scores. The statistical methods of calculating correlative relationships help define the degree of interdependence between the two variables.

“Regression” is the process of plotting the two variables on a graph, and then finding a line that “best fits” the trends in the data. It helps predict what values you might expect to see in one variable given another variable. For a sample regression, refer to Sample Measure 3 in the next section.

For both correlation and regression, it is important to understand that the equations do not definitively provide evidence of a cause-and-effect relationship. For example, the number of hours of technical assistance may increase as the capacity index score decreases (a negative relationship). One possible explanation would be that your capacity building program design requires you to spend more time with organizations with lower capacity scores. On the other hand, more hours of technical assistance may be associated with organizations with lower capacity scores because those organizations exhibited higher levels of readiness for change.

Sample Analysis and Data Displays

This section contains samples of analysis you might find in a real capacity builder’s evaluation reports.

Sample Measure 1: Increase in knowledge of best practices in managing Federal grants

The goal of this capacity builder’s training was to equip FBCOs with the knowledge, skills, and tools to successfully manage Federal grants. The capacity builder measured the outcome in two ways. First, in a post-training survey, participants were asked to agree or disagree with the statement, “I gained new knowledge about managing Federal grants.” These outcomes were self-reported, and if the capacity builder wanted to

get more objective assessments about whether the organization did, in fact, learn something about managing Federal grants, he/she might have administered a brief survey or quiz prior to and after the workshop.

Let us suppose that 54 percent of recipients self-reported that they gained new knowledge on the topic. The analysis would ask, “Is this good? What does this tell us about the training, and possibly about the participants?” On one hand, this tells us that nearly half of the participants did not learn anything new from the training. You might conclude that your training participants have all the knowledge they need of grants management and schedule other training topics. But it is possible that your staff observes something very different—perhaps that most of the participating organizations are not tracking employee time in accordance with Federal rules and regulations, or retaining records. Now you have a different situation, where your training contained very important information that the participants simply did not receive. Your next step, rather than to move onto other topics, might be to find other ways to deliver the necessary information.

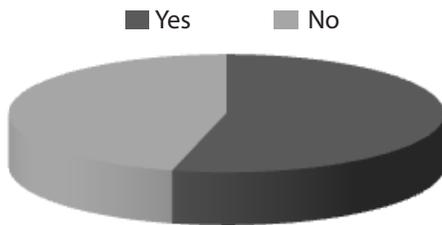


Figure 1. The data is displayed in a simple pie chart to illustrate the percentage of training participants that gained new knowledge. Pictures often provide a compelling supplement to text descriptions of outcomes.

Sample Measure 2: Organizations implement management best practices

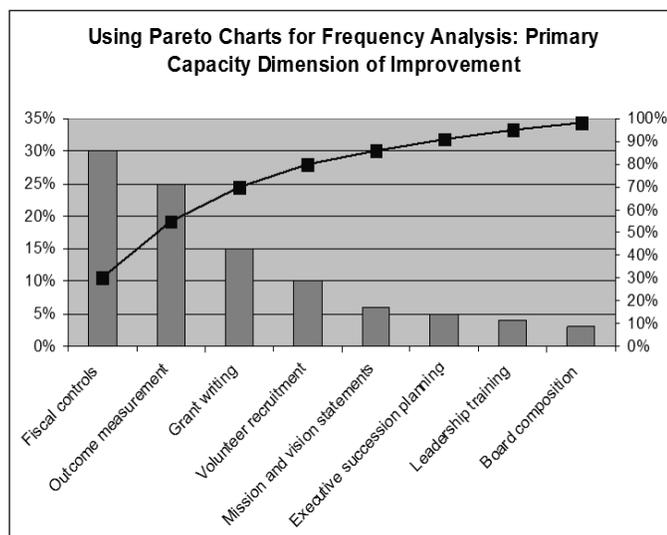
One capacity builder uses index scores for a capacity assessment (see earlier section, “Data Collection Methods”) and tracks the progress of each organization, as well as a cohort of organizations, across several categories of assessment. Here is a sample of the data, which analyzes the average and median index scores of all organizations in a cohort for each category. The assessment asks whether the organization has key practices in place for eleven areas of categories. The results are tabulated into scores as the percentage of “Yes” answers. The index scores are tracked across time by category and in aggregate for the organization and the cohort.

INDEX AREA	TIME 1 (BASELINE)	TIME 2 (END YEAR 1)	TIME 3 (END YEAR 2)
Mission-based Leadership			
Average	59	59	79
Median	60	60	100
Financial Management			
Average	64	58	63
Median	68	59	64
Information Technology			
Average	65	77	83
Median	78	78	89
Networking, Partnerships, and Alliances			
Average	64	76	80
Median	60	80	80
Fundraising and Resource Development			
Average	53	60	67
Median	57	68	75

Observations of this date show that, overall, capacity index scores were increasing across the cohort, as expected by the program. However, in the area of financial management, the scores actually dropped. According to this capacity builder, this was because, after the initial assessment, the participating organizations increased their knowledge about the topic and then answered the questions on the assessment differently, realizing that they did not have as many effective financial management practices in place as they initially thought. What could have been interpreted as a “negative” outcome is instead explained as a possible “positive” outcome.

The capacity builder might also be interested in what capacity dimensions saw the most frequent improvements overall. This might help him or her define future capacity building. The program manager might start by identifying the single biggest area of improvement for each organization:

Capacity Dimension	Percent of Organizations
Fiscal controls	30%
Outcome measurement	25%
Grant writing	15%
Volunteer recruitment	10%
Mission and vision statements	6%
Executive succession planning	5%
Leadership training	4%
Board composition	3%



Sample Measure 3: Hours of technical assistance per organization

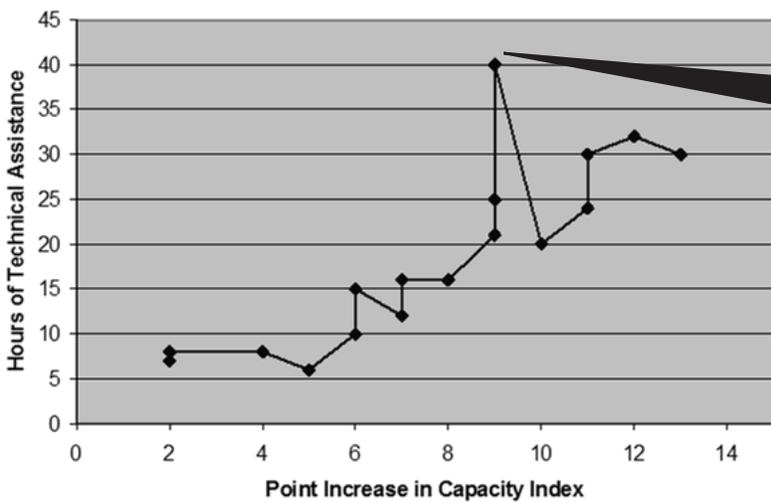
Let’s suppose that the individual capacity index scores for the previous example were compared with the hours of technical assistance each organization received and were then plotted. Why would this be of interest? In many programs, a central theory of capacity building is that more technical assistance yields better results for capacity building. If an organization is measuring capacity through the index scores and measuring the amount of technical assistance in hours, it makes sense to explore the relationship between those sets of data to see what information is available to managers. Contracted evaluators may perform rigorous statistical analysis, including correlation and regression analysis as discussed in the previous section; however, you will see in the chart below that even non-statistically savvy managers can examine the relationships between two variables.

The data set might look something like the following:

Organization	Increase in Capacity Index Score	Hours of Technical Assistance
1	2	7
2	2	8
3	4	8
4	5	6
5	6	10
6	7	15
7	7	12
8	7	16
9	8	16
10	9	16
11	9	21
12	9	25
13	9	40
14	10	30
15	11	24
16	11	30
17	12	32
18	13	30

You can see that the table compares the increase in capacity index scores, rather than just the scores themselves. In this example, the program theory specifically relates improved capacity to hours of technical assistance. The improved capacity is measured not just by the capacity index score, but by an increase in the scores between the baseline assessment prior to the intervention and the second score at the conclusion of the intervention. Additionally, the table sorts the data by the increase of capacity scores.

Figure 2 shows what the data would look like, as displayed in a two-variable plot:



Data point for further analysis: Why is this data point so far outside the trends? What circumstances might account for this variation?

Figure 2. The two-variable plot shows two sets of data. In this chart, each point represents an organization and is plotted to show the hours of technical assistance received and the point increase in capacity index score.

When the data is displayed this way, a statistician might overlay a line of regression—the line that represents the “best fit” of the trend shown. The regression line shows an approximate relationship between the two variables being plotted. In this case, it would show what kind of increase in capacity index scores you might expect given a certain amount of technical assistance. But beware not to fall into the “causation” trap. The regression line, or any trend you might observe, does not indicate that the increase in capacity index scores is caused by hours of technical assistance. It simply shows a positive relationship. To illustrate this example, let’s consider how a program manager might use this data:

- **What does the positive relationship mean?** That is, what does it mean that, in general, greater increases in capacity index score are associated with more hours of technical assistance? It could mean that investing more time in technical assistance will generate a greater increase in capacity index because the organization is able to learn more and implement more best practices. It could also mean that the rigorous process this program uses to assess readiness for change seems to work. The capacity builder assesses which organizations are ready to make the greatest changes and provides more in-depth technical assistance accordingly. This might then affirm the organization’s process for assessing readiness for change and creating technical assistance plans.
- **What about the outlier?** How can we explain one organization that received forty hours of technical assistance but did not realize the same gains in the capacity index score as other organizations in the chart? Perhaps the organization was not as ready for change as the capacity builder initially thought. That data point might tell the program manager that the initial assessment for this organization was flawed, prompting a review of the assessment process to see if there are any lessons learned for future assessments. Or, upon closer examination, the data point may indicate that this particular organization had a sticky governance issue that took additional hours to untangle before a best practice could be implemented. The capacity index in this example is based on the number of best practices the organization implements, so that if an organization spends significantly more time on a practice than other organizations, it may seem as though the gains were not as significant as those of other organizations. Another way to look at this conclusion is that the capacity index measures a breadth of issues, but does not always capture the depth. Once again, this is where a program manager’s judgment is required to distill what can be learned from the situation.

PHASE 4: COMMUNICATE THE RESULTS

Phases 4 and 5 ask capacity builders to incorporate their results into the lifeblood of their program. While Phase 5 will turn the focus inward and invite the capacity builder to use the information to improve its own practices, Phase 4 focuses externally, where the capacity builder communicates the results. This process focuses on marshalling support from various external stakeholders and generating enthusiasm about the results of the program. If the capacity builder has dedicated marketers or fundraisers, they should be leading or advising on the plans you make.

When considering how to use the results externally, create a plan that incorporates 1) what to communicate, 2) how to communicate it, and 3) the target audience for each communication.

What to Communicate

The following suggestions are good ways to communicate your findings. You will determine the right mix of these six forms of communication:

1. **Issue a formal report.** At the conclusion of the project, complete a full report of your evaluation efforts. Describe your desired outcomes and your logic model, the data collection plan, your results, and any recommendations or actions you have taken or plan to take as a result.
2. **Present case studies or stories of impact.** These may be part of your formal report, but they can also be teased out and made available as a marketing or teaching tool. Focus on a single organization and the results that organization achieved through your capacity building intervention.
3. **Develop press releases.** Draft a press release highlighting the strongest results you have discovered and distribute to local newspapers, columnists, bloggers, community e-mail lists, and neighborhood organizations.
4. **Create snapshots or postcards.** Distill your key results into a short list, and turn that short list into a snappy display for use in print or online. Printed, laminated cards make great promotional materials and can be handed out at community meetings, mailed to constituents, and displayed on websites.
5. **Incorporate visual aids.** Whenever possible, reinforce the numerical results with pictures—graphs or charts. You can also include photos.
6. **Produce a promotional video.** Record interviews with organizational leaders discussing how the capacity building help they received improved their organization. Leaders of the capacity builder can discuss the program's goals and results. The video can be loaded for free on YouTube or hosted on the organization's website.

How to Communicate

Having produced great materials from the suggestions above, it is important for you to put them to good use. The following are tried and true methods of connecting external stakeholders to the materials that communicate evaluation results.

1. **Enhance your web presence.** Snapshots, case studies, stories, visual aids, and even your final report should be made available on your program's website, Facebook page, or any other online communities you are part of, including a nonprofit management association or any associations of management-support organizations.
2. **Invite the media.** Send press releases to local media outlets and invite them to tour your site. Newspapers that cover local events may be interested in reporting on the results you have achieved. Invite members of the media, including local bloggers or activists, to interview leadership and take a tour of your facilities or programs.
3. **Give presentations.** Invite stakeholders (board, partners, funders) for a meeting. Large organizations may be able to call a press conference. Go to your funders or stakeholders. If your organization is invited to give a speech or present information at a forum, be sure to include the key results in the introduction of your organization. Note that these presentations can be more affordable and reach a wider geographic area if done virtually (via a webinar), using a product like GoToWebinar or Adobe Connect.
4. **Identify a program champion.** Identify a charismatic leader from an organization that has benefited greatly from your capacity building program and invite them to be a partner with you to help get the word out about the great results they have experienced in the program. This individual can participate in panels and forums, write letters on behalf of your organization, and participate in other outreach activities. Remember to recognize the leader for his or her efforts in the most personally meaningful way.

Target Multiple Audiences

It will be important to include many audiences when discussing the results of your outcome measurement. Some possible audiences include:

1. **Potential and current funders.** This is most likely the first audience that capacity builders will think to target, and rightly so; communicating the effectiveness of your services is key to sustainability.
2. **Partners.** Include all the organizations and individuals who contribute to your program's operations.
3. **Your client organizations.** Throughout the process, we have encouraged capacity builders to work closely with the organizations that are part of the capacity building engagement. This should not stop when it comes to the distribution of results. In addition to any results for that particular organization, provide the overall results of the program to give the organization an idea of how it compares to others. This builds goodwill, enhances your reputation across the nonprofit sector, and can help generate buzz that is sure to get back to funders and create more demand for your services.
4. **Organizations you did not serve.** Seeing the results of another organization can be a powerful marketing pitch and can generate more interest in your programs. You may consider engaging individuals from organizations you once served to help deliver the message to the organizations that you have not yet served via community panels, webinars, etc.
5. **Community leaders.** Think about who benefits from the success of your capacity building program. Stronger nonprofits mean better results for the individuals being served by those organizations. Therefore, include local politicians (mayors, county executives, congressmen and women, state legislators) and institutions that benefit from the stronger organizations, such as school systems or corrections officials.
6. **Management and staff inside your organization.** Phase 5 discusses in greater detail how to take the results and use them to examine internal performance and processes.

PHASE 5: REFLECT, LEARN, DO IT AGAIN

The Alliance for Nonprofit Management reported that a best practice in evaluating capacity building was to approach the entire evaluation and outcome measurement process as a learning organization. This happens throughout the process and culminates with a critical and thorough examination of what worked and what needs to be improved in the future. Each phase and the steps therein should be reviewed while asking the following questions:

1. **Logic model.** Is your program logic model accurate to what your program actually works toward? Are there inputs, outputs, or outcomes that are missing?
2. **Outcome measures.** Were the outcome measures actually feasible? Did they capture the data you intended to about the outcomes?
3. **Data collection plan.** Were the tools functional? Did the individuals collecting data or filling them out understand what was asked and provide the relevant data? How can they be improved? Was data collected at the right time? Were you able to perform quality control for the data?
4. **Data analysis.** Who performed the analysis? Was the information provided useful? Was it surprising? Why or why not? Was it easy to understand?
5. **Communication.** Who did the results go to, and in what form? What could be done better and what went well?

This is also a good time to look at the results. While your results are not likely to identify a statistical cause-and-effect relationship, if your outcomes are unexpected—for better or for worse—it is common sense to examine the practices. Why was this capacity building strategy put into place over another? Was it effective? What problems were encountered and what were the resolutions? You may find that your goals or

outcomes need to be tweaked or perhaps rewritten altogether. The bottom line is that you should be sure to build in some time to learn from the results and reflect on the process and implementation of the outcome measurement plan. This will allow you to achieve better results in the future.

APPENDICES

APPENDIX A

Online Resources: Survey and Questionnaire Development

A Brief Guide to Questionnaire Development <http://ericae.net/ft/tamu/vpiques3.htm>

Covers such topics as the advantages and disadvantages of open-ended and objective questions, sample size, and questionnaire design and format.

Basics of Developing Questionnaires <http://www.managementhelp.org/evaluatn/questnrs.htm>

Provides a brief checklist of issues to consider when designing a questionnaire, including how the questionnaire is worded, consent and confidentiality issues, and communicating the purpose of the questionnaire.

Other Online Resources

American Evaluation Association <http://www.eval.org>

Includes links to resources and tools for evaluation.

Center for Program Evaluation and Performance Measurement, Bureau of Justice Assistance
<http://www.ojp.usdoj.gov/BJA/evaluation>

Provides a “roadmap” to evaluation and includes useful tools and links to evaluation resources.

Evaluation Working Group, Centers for Disease Control and Prevention <http://www.cdc.gov/eval/resources.htm>

Contains many links related to evaluation, logic models, tools, and other evaluation-related resources.

Outcome Measurement Resource Network, United Way of America <http://www.liveunited.org/outcomes>

Offers information, downloadable documents, and links to resources related to the identification and measurement of program- and community-level outcomes.

Outcome Indicators Project, Urban Institute and The Center for What Works <http://www.urban.org/center/cnpl/projects/outcomeindicators.cfm>

Provides a framework for tracking nonprofit performance; suggests candidate outcomes and outcome indicators to assist nonprofit organizations that seek to develop new outcome monitoring processes or improve their existing systems.

Taking Stock: A Practical Guide to Evaluating Your Own Programs

http://www.horizon-research.com/reports/1997/taking_stock.php

Offers an online guide to program evaluation.

APPENDIX B

Glossary

Activities – The methods of service delivery carried out by staff.

Anonymous data – Information where you do *not* know who provided the responses. (Compare with “confidential data” below.)

Baseline – Data gathered to provide a comparison for assessing program changes or impact.

Comparative standard – Data used as a comparison or a standard of achievement for a specific indicator or outcome.

Compliance monitoring – Tracking and reporting information on what and how much service a program delivers, the clients it serves, how much money it expends, and, possibly, the outcomes it achieved, in relation to what an organization has agreed upon, generally referring to contractual arrangements made between an organization and its grant maker on the use of funds.

Confidential data – Information where you do know, or can find out, who provided the responses but keep the information private.

Data – Information collected in a systematic way that is used to draw conclusions about a program or its outcomes.

Evaluation – The systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of health or social interventions.

Goal – A broad statement of the ultimate aims of a program, generally beyond the ability of one organization to achieve on its own.

Indicator – The specific, measurable information that will be collected to track success of an outcome. Another commonly used phrase is “performance measure.”

Inputs – The resources available to a program that allow and support service delivery, e.g., money, staff, volunteers, materials, or equipment.

Instrument – A tool used to collect data, including survey questionnaires, interview guides, observational checklists, and written record extraction forms.

Objective – A specific, measurable accomplishment within a specified time frame.

Outcome – The changes in the lives of individuals, families, organizations, or the community as a result of a program; the benefit for participants during or after their involvement in a program; or the impact of a program on the people it serves.

Outcome evaluation – Systematic examination of the impact of a program and what resulted for the participants, clients, consumers, or customers. Another commonly used phrase is “summative evaluation.” (Also see “process evaluation.”)

Outcome measurement – A systematic way to assess the extent to which a program has achieved its intended results; generally used in the nonprofit world.

Output – The product delivered or the unit of service provided by a program, usually described numerically, such as number of people served or number of hours of service delivered.

Performance measurement – Similar to outcome measurement but generally used in business and government arenas.

Process evaluation – Systematic examination of the degree to which a program is operating as intended, looking at what services it provides, how they are provided, who receives the services, and how much service is delivered. Another commonly used phrase is “formative evaluation.” (Also see “outcome evaluation.”)

Program logic model – A representation of the linkages between the inputs to a program, the resources available to it, and the activities carried out, and the outputs and outcomes those resources and activities are believed to produce.

Qualitative data – Descriptive or subjective information provided in narrative terms.

Quantitative data – Numerical information gathered in a structured way.

Reliability – The extent to which data collected are reproducible or repeatable.

Target – The specific level of achievement for an indicator or outcome.

Validity – The accuracy of information collected.

Sources for Glossary

Development Leadership Network. *Success Measures Guidebook*. Boston, MA: 2000.

Reisman, Jane and Judith Clegg. *Outcomes for Success!* Seattle, WA: The Evaluation Forum, 2000.

Rossi, Peter H., Mark W. Lipsey, and Howard E. Freeman. *Evaluation: A Systematic Approach*, 7th ed. Thousand Oaks, CA: Sage Publications, 2004.

Sullivan, Cris M. *Outcome Evaluation Strategies for Domestic Violence Programs: A Practical Guide*. Harrisburg, PA: Pennsylvania Coalition Against Domestic Violence, 1998.

United Way of America. *Measuring Program Outcomes: A Practical Approach*. Alexandria, VA: 1996.

APPENDIX C

Sample Program Logic Model

Using this worksheet, you can create a logic model for your program with an outcome chain of short-term, intermediate, and longer-term outcomes.

Agency:		Program Name:		Program Goal:	
INPUTS OR RESOURCES	ACTIVITIES	OUTPUTS	SHORT-TERM OUTCOMES	INTERMEDIATE OUTCOMES	LONG-TERM OUTCOMES
Resources available to the program that allow and support provision of technical assistance, including money, staff, volunteers, clients, materials or equipment	The technical assistance methods carried out by staff	The product delivered or unit of technical assistance provided, usually described numerically, such as number of organizations served or number of hours of service	The first changes that occur for individuals, families, organizations, or the community as a result of the program	The subsequent benefit for people during or after their involvement in the program	The eventual impact on individuals, families, organizations, or the community for which the program is accountable

APPENDIX D

Outcomes and Indicators Worksheet

Transfer the outcomes from your logic model into the first column of this worksheet and then use the remaining columns to identify up to three potential indicators for each outcome.

Agency:		Program Name:	
OUTCOMES	INDICATORS	INDICATORS	INDICATORS
Select the most meaningful, relevant, important outcomes from your logic model and write them here	For each outcome, identify the specific, measurable information that will be collected to track success toward that outcome	Is there another possible indicator?	Another?

APPENDIX E

Data Collection Plan Worksheet, Part 1

Select outcomes and indicators from the previous worksheet and identify for each one the data collection methods you will most likely use and the schedule for your data collection.

Agency:		Program Name:	
OUTCOME	INDICATOR	WHAT DATA COLLECTION METHOD?	WHEN WILL DATA BE COLLECTED?
Select the outcomes that will be measured and write them below	Identify which indicator(s) will be measured for the outcome and write them below	Identify which type of data collection method (survey, interview, observation, or record review) will be conducted	Describe the timing and frequency of data collection

Data Collection Plan Worksheet, Part 2

Transfer the information on data collection methods from the previous worksheet to the first column in this worksheet and identify for each one the data collection procedures you believe will help manage your data collection process.

Agency:		Program Name:	
WHAT DATA COLLECTION METHOD?	WHO WILL COLLECT DATA?	WHAT WILL THEY DO?	HOW WILL DATA COLLECTED BE MONITORED?
Write the data collection methods from Part 1 of the Data Collection Plan below	Identify who will be responsible for collecting the data	Describe the steps that will be taken to implement the data collection	Identify who will monitor the data collection process for quality and consistency and how they will do so

APPENDIX F

Getting the Best Results from Surveys and Interviews

Adapted from the Survey Workbook, The Evaluation Forum, Seattle, WA, 2002.

PURPOSE

- Do you know what you want to learn from the survey or interview?
- Have you thought about why you need the information?
- Have you identified who you need to ask to get the information?
- Have you thought about the different audiences for the information?
- Have you thought about how you will use the information?
- Have you identified the objectives of your data collection and/or research questions?

DESIGN AND METHOD

Reasons for using written surveys:

- Are resources or volunteer time limited?
- Do you need to collect data from a sizable number of people?
- Is interviewer bias a concern?
- Is privacy necessary?
- Are quantifiable rankings or priorities necessary?
- If it is to be done through e-mail or the Internet, do respondents have access to those technologies?

Reasons for using interviews:

- Do you have the time necessary to conduct interviews?
- Can you find and train the interviewers?
- Is education or literacy an issue for participants that make interviews more appropriate?
- Are questions likely to be complex?
- Will you need visual aids?
- Is it important that every question be answered?
- Is it difficult to know what the range of responses might be?
- Do you have the resources necessary to conduct interviews?

LANGUAGE

- Is each question meaningful?
- Is the language simple, clear, and concrete?
- Does it use language familiar to the participants?
- Are the instructions understandable?
- Are checklist choices relevant?
- Are the scales simple and understandable?
- Do the scale choices have clear distinctions between them?
- Do the scales have a clear midpoint, if appropriate?
- Is any of the wording judgmental or leading to a desired answer?
- Are time frames specific?
- Does each question include just one thought?

FORMAT

- Is it easy to read and follow?
- Is it as short as possible?
- Are the checklist boxes easy to select?
- Are the scales easy to read and apply?
- Do participants prefer numeric scales or wording scales?
- Do the contingency questions guide the reader appropriately?
- Is the typeface large enough for participants to read?

CULTURAL APPROPRIATENESS

- Is the method you have chosen appropriate for your client group?
- Are the types of questions appropriate for your client group?
- Is the content of the questions acceptable?
- Can the questions be translated easily, if needed?
- How does the meaning of questions change in translation?

PRE-TEST

- Has the instrument been pre-tested?
- Did participants interpret the meaning of the questions correctly?
- Were participants able to follow the format?
- Did any questions appear to make respondents uncomfortable?
- Did any questions consistently go unanswered?
- Did the instrument appear appropriate for the client group?
- Did you get back the information you thought you should?

RESPONSE RATES

- Will the survey be announced ahead of time?
- Will a cover letter be included with the mailing?
- Are respondent addresses up to date and accurate?
- Is the return address included on the survey?
- Have you provided a stamped and addressed return envelope?
- Can you do multiple mailings?
- For phone interviews, has a concise introduction been written?
- Are phone lists up to date and accurate?
- Are several phone attempts made on different days and times?
- Are holidays and other events avoided?

APPENDIX G

Technical Assistance Survey Template

Adapted from materials developed for AIDS Housing of Washington’s technical assistance program by Clegg & Associates, Inc., Seattle, WA (2003).

Please check the appropriate box in response to each statement below, followed by any comments you may have. If you are unfamiliar with a particular area, select “Not Sure.” Thank you!

How satisfied are you with the following aspects of our services?

	Extremely Dissatisfied	Dissatisfied	Satisfied	Extremely Satisfied	Not Sure
Knowledge demonstrated by staff					
Responsiveness of staff to your needs					
Usefulness of information/ assistance provided					
Overall quality of the services provided					

Comments:

How can we improve our service delivery to you?

To what degree do you feel you have improved your knowledge or skills in the following areas as a result of your work with us?

<i>[Complete one line for each topic or skill area addressed by the training and/or technical assistance you provided]</i>	No Improvement Whatsoever	Minimal Improvement	Moderate Improvement	Large Improvement	Not Sure

Comments:

To what degree do you feel your organization has improved its knowledge or skills in the following areas as a result of your work with us?

<i>[Complete one line for each topic or skill area addressed by the training and/or technical assistance you provided]</i>	No Increase Whatsoever	Minimal Increase	Moderate Increase	Large Increase	Not Sure

Comments:

To what degree do you feel your organization has improved its capacity in the following areas as a result of your work with us?

[Complete one line for each topic or skill area addressed by the training and/or technical assistance you provided]

	No Increase Whatsoever	Minimal Increase	Moderate Increase	Large Increase	Not Sure

Comments:

What do you feel is the most valuable result of your work with us?

What additional assistance do you think would help your organization right now?

Any other comments or suggestions?

APPENDIX H

References

- Babbie, Earl. *The Practice of Social Research*, 12th ed. Belmont, CA: Wadsworth Publishing, 2010.
- Bolan, Marc, Kimberly Francis, and Jane Reisman. *How to Manage and Analyze Data for Outcome-Based Evaluation*. Seattle, WA: The Evaluation Forum, 2000.
- Development Leadership Network. *Success Measures Guidebook*. Boston, MA: 2000.
- Fink, Arlene, ed. *The Survey Kit*, 2nd ed. Thousand Oaks, CA: Sage Publications, 2002.
- Herman, Joan L., ed. *Program Evaluation Kit*. Thousand Oaks, CA: Sage Publications, 1987.
- Morgan, David L. and Richard A. Krueger. *The Focus Group Kit*. Thousand Oaks, CA: Sage Publications, 1997.
- Office of Management and Budget, Executive Office of the President. *The President's Management Agenda*. Washington, DC: 2002. <http://www.whitehouse.gov/omb/budget/fy2002/mgmt.pdf>.
- Patton, Michael Quinn. *Utilization-Focused Evaluation*, 4th ed. Thousand Oaks, CA: Sage Publications, 2008.
- Pratt, Clara C., Aphra Katzev, Rachel Ozretich, Tammy Henderson, and William McGuigan. *Measuring Outcomes for Oregon's Children, Youth, and Families: Building Results III*. Salem, OR: Oregon Commission on Children and Families, 1998.
- Reisman, Jane and Judith Clegg. *Outcomes for Success!* Seattle, WA: The Evaluation Forum, 2000.
- Rossi, Peter H., Mark W. Lipsey, and Howard E. Freeman. *Evaluation: A Systematic Approach*, 7th ed. Thousand Oaks, CA: Sage Publications, 2004.
- United Way of America. *Measuring Program Outcomes: A Practical Approach*. Alexandria, VA: 1996.
- Yarbrough, Donald B., Lyn M. Shulha, Rodney K. Hopson, and Flora A. Caruthers. *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users*, 3rd ed. Thousand Oaks, CA: Sage Publications, 2010.

This work was created by the Compassion Capital Fund National Resource Center, operated by Dare Mighty Things, Inc., in the performance of Health and Human Services Contract Number HHSP23320082912YC.