



Community Food Project Evaluation Handbook

COMMUNITY FOOD SECURITY COALITION



Sponsor:
USDA Community Food Projects Program

Author:
National Research Center, Inc.

Third Edition, 2006



Community Food Project Evaluation Handbook

COMMUNITY FOOD SECURITY COALITION

Community Food Security Coalition
PO Box 209
Venice, CA 90294
Phone: 310-822-5410
www.foodsecurity.org

Sponsor: USDA Community Food Projects Program

Author: National Research Center, Inc.
3005 30th Street, Boulder, Colorado 80301
Phone: 303-444-7863
nrc@n-r-c.com
www.n-r-c.com

First Edition, 2003
Second Edition, 2004
Third Edition, 2006

Copyright © 2006 by National Research Center, Inc. All rights reserved.

We welcome limited duplication of contents of the Evaluation Handbook for non-profit and educational purposes. Please credit the source in all copies, and if possible, include this page.

The *Community Food Project Evaluation Handbook* was written by National Research Center, Inc. Some content within this handbook was adapted from previously written handbooks authored by NRC including *Outcome Handbook for City of Boulder Human Service Providers* (January 1999) and *After School Initiative Evaluation Handbook* sponsored by The Colorado Trust (2001-2002).

A hearty thanks to the following Community Food Project grantees for the cover photos they shared of their project activities. Starting from the top left:

Photo # 1 Girl living at Re-Vision House Shelter enjoys a yellow tomato

Submitted by Laurell Sims

Re-Vision Urban Farm

Dorchester, MA

www.re-visionfarm.org

Photo #2 Nou Yang, a Hmong Farmer, at Boston Flats

Submitted by Jennifer Hashley

New Entry Sustainable Farming Project

Boston, MA

jennifer.hashley@tufts.edu

<http://www.nesfp.org>

Photo #3 Esperanza Echeverria, a farmer from Guatemala, prepares for the Lewiston Farmers Market

Submitted by Jim Hanna

New American Sustainable Agriculture Project

Coastal Enterprises, Inc.

Lewiston, ME

jimhanna@maine.rr.com

<http://www.ceimaine.org/farm/home.htm>

Photo # 4 Gleaners in a tomato field

Submitted by Shannon Kushnick

The Appalachian Center for Economic Networks (ACENet)

Athens, OH

www.acenetworks.org.

Photo # 5 Grandmother who had a garden plot at Kalpulli Community Garden

Submitted by Raul Aragon

Center for Academic Preparedness

Sylmar Cooperative Food and Farm Project

Northridge, CA

Photo # 6 Interns selling produce grown at the Re-Vision Urban Farm

Submitted by Laurell Sims

Re-Vision Urban Farm

Dorchester, MA

www.re-visionfarm.org

Overview

The Community Food Project Evaluation Handbook (Handbook) is a practical guide to help community food project staff conduct rigorous program evaluation that will develop the kind of information about their programs that will be compelling not only to fellow staff but to funders, participants and community residents. It is full of basic concepts, specific examples and worksheets.

The Handbook is comprehensive, with 10 chapters covering information from planning, to implementation, to using your results for program improvement. While it may seem daunting, the Handbook clearly walks the reader step by step through the evaluation process. The Table of Contents outlines in detail the various sections of the Handbook and can help the reader find a specific topic of interest. A glossary is provided as the final Appendix to include definitions of research and evaluation terms used in the chapters of the Handbook.

We recommend reading through the first four chapters of the Handbook completely. These chapters (Introduction to Evaluation, Mapping Your Project Using a Logic Model, Selecting Your Outcomes and Laying the Evaluation Groundwork) will provide an excellent foundation for beginning your evaluation plan. You may choose to be more selective in reading through the remaining chapters, depending on the resources and timeline you have for engaging in evaluation. For example, if a volunteer from a local university plans to analyze your evaluation data, you may not need to read Chapter 8 (Analyzing Your Data) page by page.

Readers will find new content and improved organization in this third edition of the *Community Food Project Evaluation Handbook*. The Handbook includes new sections in Chapter 7 covering the use of incentives, guidelines for evaluation administrators and culturally responsive data collection. Two new appendices were also added: Tips for Conducting Focus Groups with Teens (Appendix VI) and Common Response Options (Appendix VIII). The Handbook was also reorganized to add clarity and improve the transitioning from one evaluation step to another. Restructuring allowed for the merging of some chapter sections so, although no content was removed, readers will find that this edition contains two fewer chapters than the previous version.

Background

This handbook was originally designed in 2003 as a reference guide for Community Food Projects (CFPs) funded by the United States Department of Agriculture (USDA). A companion *Community Food Project Evaluation Toolkit* (Toolkit) was also developed to provide specific surveys and evaluation templates tailored for the needs of community food projects. Both publications were funded through a CFP Training and Technical Assistance grant awarded to the Community Food Security Coalition (CFSC) to provide evaluation support to other CFP grantees.

Now in their third edition, the Handbook and the Toolkit are available to anyone interested in program evaluation through the CFSC's Evaluation Program or at www.foodsecurity.org (in the Publications section).

National Research Center, Inc. (NRC) of Boulder, Colorado, was contracted by CFSC to provide evaluation training and technical assistance to CFPs as well as to develop a set of common evaluation tools for evaluating CFPs. NRC, in collaboration with the CFSC Evaluation Program, authored the Handbook and Toolkit and worked closely with CFP grantees to pilot test the information and tools developed.

Participation in the evaluation workshops offered by the CFSC and utilization of this handbook and the companion toolkit will help community food project staff to increase their understanding of evaluation, especially as it relates to the impact of their projects. Staff will enhance their ability to perform outcome-based evaluation and to compile and present evaluation findings to project staff, the individuals their project serves, other community member and funders. The Handbook is aimed at building the capacity of community food project staff to conduct their own evaluations leading to stronger programs and greater program sustainability.

A Note to Community Food Project Grantees and Practitioners

Community Food Projects around the country represent some of the most dynamic individuals and organizations working for food security and social change. This was strikingly apparent during the over forty assessment interviews we conducted in 2003 with Community Food Project grantees to learn about your projects' evaluation efforts. The diversity of project goals shared in those interviews points to the complex and interwoven nature of the social, economic and environmental issues most CFPs hope to impact. The criteria USDA uses to determine Community Food Projects awards are ambitious. They look at such issues as:

- Meeting the food needs of low-income people
- Increasing community self-reliance
- Promoting comprehensive responses to food, farm and nutrition issues
- Developing innovative links
- Supporting entrepreneurial development
- Encouraging long-term planning
- Encouraging multi-system, interagency approach
- Achieving project self-sufficiency

Over the past 10 years since the USDA's Community Food Projects Program began, there have been over 200 projects awarded. These projects, along with other community food security advocates and organizations, are working both to impact their communities and to build a broader movement for community food security and positive social change.

Your work as community food project practitioners is important, inspiring and unique.

The USDA's award to the Community Food Security Coalition of a grant for evaluation training and technical assistance is an important contribution to the development of the community food security movement. Evaluating the effectiveness of one's program can be difficult in any setting. It is even more complex when one is dealing with the diverse and multi-faceted nature of Community Food Projects. Yet, the importance of evaluation for Community Food Projects is compelling.

Thorough and thoughtful evaluations of your Community Food Project will help you to document your work, to understand the impacts of that work, and to improve your program's effectiveness. They will also contribute to the community food security movement as a whole, by adding to the growing body of knowledge about what is and isn't working. They will help new and fledging projects to set appropriate goals and to improve their program design. A growing body of evaluation results will also help to

document the multiple benefits of Community Food Projects and to persuade funders and decision-makers to support such projects.

The CFP Evaluation Handbook, the companion CFP Evaluation Toolkit, and our evaluation workshops are designed to build the evaluation capacity of Community Food Projects by focusing on outcome-based evaluation using a logic model.

We realize this is one of a variety of ways to conduct evaluation. We encourage you to enhance these resources by sharing your own insights and feedback and by supplementing them with other evaluation methods and tools appropriate for your community and your needs.

Because we realize the important role of Community Food Project coordinators, we are asking for your feedback on the CFSC Evaluation Program. We hope you will take the time to complete the CFP Evaluation Handbook and Toolkit Evaluation Form found online at www.foodsecurity.org, in the publications section.

We are also interested in your ideas about how to make the Evaluation Program relevant to your resources, needs and activities. Please contact me with your feedback or if you would like to become involved with the CFSC Evaluation Program by pilot testing tools, review training materials or acting as a peer trainer. Your insight on the evaluation methods we develop and use will enhance our collective ability to tell the 'stories' of our projects, of our communities and of the dynamic and growing community food security movement.

Thank you for your important work!

Jeanette Abi-Nader
CFSC Evaluation Program Manager
jeanette@foodsecurity.org

February 2006

COMMUNITY FOOD SECURITY COALITION 

Contents

Chapter 1. Introduction to Evaluation	1
What Is Evaluation?	2
The Difference between Evaluation and Research	3
Evaluation Landscapes	3
Purpose of This Handbook: Outcome Identification and Tracking	7
The Evaluation Process	8
The Culture of Evaluation	9
Chapter 2. Mapping Your Project Using a Logic Model	11
Revisiting Your Project Goals	12
Linking Program Activities to Goals: Introduction to the Logic Model	13
The Basic Logic Model	13
A More Complete Logic Model	18
Steps to Developing a Complete Logic Model	23
Chapter 3. Selecting Your Outcomes	29
Outcomes in Daily Life	30
An Increasing National Emphasis on Outcomes	30
Organizational Benefits of Pursuing Outcome Assessment	31
Lessons from Community-Based Organizations Working on Evaluation	33
Defining “Outcomes”	34
Setting Performance Standards	40
Chapter 4. Laying the Evaluation Groundwork	45
Establishing a Learning Environment	46
Engaging Program Stakeholders in Evaluation	47
Assessing the Data on Hand	50
Developing an Evaluation Plan	60
Chapter 5. Selecting Evaluation Strategies and Study Designs	67
Quantitative and Qualitative Research: A Sibling Rivalry	68
Typical Study Designs	70
Common Evaluation Methods	77
Selecting the Best Methods	79
Comparison of Evaluation Methods	81
Guidelines for Method Selection	83
Culturally Responsive Evaluation Methods	84
Chapter 6. Designing and Customizing Evaluation Tools	87
Searching for Tools	88
Evaluating Existing Tools	90
Developing Your Own Tools	91
Designing Culturally Responsive Evaluation Tools	91
Developing Effective Tools	94
Pilot Testing Your Evaluation Tools	101
Chapter 7. Collecting Data	103
Determining How Many To Include in Your Evaluation	104
The Use of Incentives in Evaluation	106
Developing Data Collection Protocol	107

Guidelines for Evaluation Administrators.....	111
Understanding and Protecting the Rights of Evaluation Participants	113
Collecting Data in a Culturally Responsive Way	115
The Key to Ethical Evaluation	115
Chapter 8. Analyzing Your Data.....	117
Creating an Evaluation Notebook.....	118
Developing an Analysis Plan	118
Quantitative Data Analysis.....	123
Qualitative Data Analysis.....	136
Analysis of Secondary Data.....	139
Chapter 9. Understanding and Using Your Results	141
Understanding Your Data.....	142
Using Your Results for Program Improvement.....	147
Chapter 10. Communicating Results.....	155
Determining Your Audiences.....	156
The Basic Evaluation Report.....	157
Effective Methods for Data Presentation and Dissemination	160
Attributing Causation.....	161
Using Your Results for Fundraising and Grant Writing.....	161

Appendices

Appendix I: Program Evaluation Standards.....	167
Appendix II: Electronic Evaluation Resources	171
Appendix III: Rubric Template.....	173
Appendix IV: Sample Rubrics.....	175
Appendix V: Designing Age-Appropriate Tools for Youth and Children	177
Appendix VI: Tips for Conducting Focus Groups with Teens.....	185
Appendix VII: Designing Age-Appropriate Tools for Older Adults	187
Appendix VIII: Common Response Options	189
Appendix IX: Descriptive Analyses Using Excel.....	191
Appendix X: More Advanced Statistics and Statistical Testing	197
Appendix XI: Sample Executive Summary	199
Appendix XII: Sample Annotated Evaluation Tool	203
Appendix XIII: Communication Materials.....	207
Appendix XIV: Glossary.....	217

Worksheets

Worksheet 1: Developing Your Project’s Basic Logic Model	17
Worksheet 2: The Complete Logic Model.....	25
Worksheet 3: Outcomes, Indicators and Performance Standards	43
Worksheet 4: Data Resource Inventory.....	52
Worksheet 5: Evaluation Plan	62
Worksheet 6: Questions to Help Design a Data Collection Protocol	110
Worksheet 7: Analysis Plan	122
Worksheet 8: Plan to Use Evaluation Results.....	153

Chapter 1.

Introduction to Evaluation

-  What is Evaluation?
-  The Difference Between Evaluation and Research
-  Evaluation Landscapes
-  Purpose of This Handbook: Outcome Identification and Tracking
-  The Evaluation Process
-  The Culture of Evaluation

Chapter 1. Introduction to Evaluation

What Is Evaluation?

The term “to evaluate” means to determine the worth of something. It has a cold ring to it, like a pawnbroker examining a reputed gem or a car dealer deciding what your trade-in is worth. In the early days of modern evaluation, social scientists were expected to make such cold judgments about social programs and policies. Over time, evaluation of community-based service programs has taken on a softer tone. Everyone, from the program staff to the participants and funders, still wants to know what impact a program has, but today evaluation tends to be conducted in partnership with program delivery.

Evaluation can be defined in a variety of ways. A definition that is applicable to community-based service programs follows:

Evaluation is the systematic way that data are assembled into a picture of (1) how well an organization is delivering its services and (2) the impact of those services on the target population. ¹

A useful community food project (CFP) evaluation should tell you something about your program and meaningful changes experienced by its participants. It will help you not only understand your successes, but also provide you with information that will help you improve your project to best serve your community. What are the best things that your program has to offer? Is the project meeting its intended goals? Does it make a difference in the lives of the people served? These questions, and others, are evaluation questions.

The concept of evaluation as a field of study began in the 1960's. Economists, psychologists, political scientists and social scientists were welcomed into the Kennedy administration with the hope that their trade “would give cause-and-effect theories for policy making so that statesmen would know which variables to alter in order to affect the desired outcome. And once policies were in operation, it would provide objective evaluation of their effectiveness.”² The deepest roots of evaluation can be linked to the deepest roots of all research, going back to Socrates and the Socratic Method as well as the development of the scientific method. However, during the explosion of social programs in the 1960's, politicians needed to understand which of the many programs made the most sense to fund and which programs were the most cost effective. This drove the modern boom in evaluation studies.

Although often it is funders who require evaluation, there are many ways you, too, can use evaluation results to benefit your program and activities. We will talk about these in more detail in later chapters of The Handbook.

The Difference between Evaluation and Research

While good research skills may be used in designing an evaluation, there are key ways in which evaluation differs from research. Some of these distinctions are shown in the table below.

How Evaluation Differs from Research ³
<ul style="list-style-type: none">• Evaluation is controlled by those involved (the stakeholders) instead of being rigorously designed by an investigator.• The steps of evaluation vary considerably from those of basic research.• Standards of evaluation include usefulness, feasibility, accuracy, and fairness rather than internal and external validity.• Evaluation assesses merit, worth and importance rather than emphasizing associations.• Evaluation is holistic and flexible by design to allow for changes and unexpected circumstances rather than being tightly controlled.• Evaluation is ongoing rather than being limited to a specific timeframe.• The scope of the evaluation is broad, in an attempt to be integrative, rather than narrowly focused.• Judgments from evaluation depend on agreed-upon or specifically stated values of a stakeholder rather than being value-free.• Uses of evaluation data is imperative not just to further knowledge and help improve similar programs through publication, but also to build capacity or improve a program.

Evaluation Landscapes

Evaluators share a language, just like other professionals. You do not need to be a professional evaluator to understand the basics of evaluation nor to participate in a basic evaluation. But you may wish to have some understanding of the terminology and philosophy of the field. In brief, evaluation methods have been developed and refined over the years providing many different kinds of methodologies to serve a variety of functions. “Program evaluation” can legitimately mean many things to many people, as the field covers a wide range of activities and purposes.

Different types of evaluation are defined by the intention of the evaluation or the strategies employed to carry out the evaluation. Many of these various types of evaluation can overlap, creating a complex picture. The following section is provided to help promote understanding of these evaluation distinctions. Rather than thinking of these distinctions as being employed in isolation of the other, it would be more accurate to think of them as continuums. Many evaluations will fall somewhere along the spectrum of the two extremes.



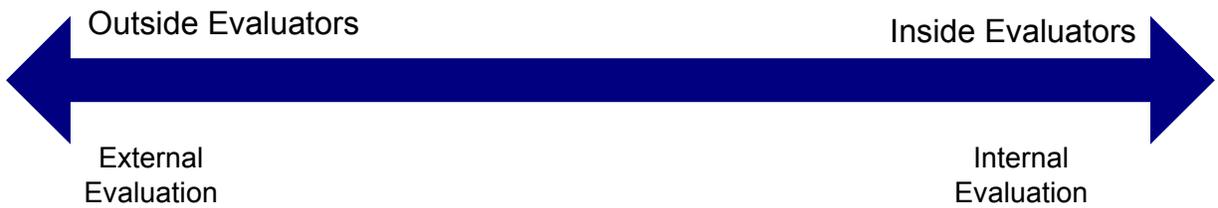
Methodology: Quantitative and Qualitative

Evaluators sometimes use “quantitative” or “qualitative” to describe the methodology used to design the evaluation, collect the data, analyze the data and report on the data. Evaluations that emphasize numbers (from surveys, existing databases and tests) in collecting and summarizing data are quantitative while those that emphasize words (from stories, focused discussions, personal interviews) are qualitative. Many quantitative studies use methodologies from the “hard” sciences fields (i.e., biology, statistics) while many qualitative studies use methodologies from the “soft” science fields (i.e., anthropology, sociology). Both quantitative and qualitative studies must follow rigorous methodologies in order to be legitimate evaluations. These methodologies are discussed in greater detail in Chapter 5.



Purpose: Formative to Summative

Evaluations that describe how a program’s activities might be improved are called formative evaluations. Evaluations that demonstrate what a program has accomplished are called summative evaluations. Formative evaluations ask “What is it?”, “How does it work?”, and “How can it be improved?” They often occur during early stages of a program because they provide feedback and allow for changes in the program. Summative evaluations ask, “Did it work?” They often take place once a program is fully in place.



Position of the Evaluators/Data Collectors: Internal to External

If someone without a vested interest in your program comes into your organization and gathers data about project outcomes, they will be considered an external evaluator. If someone with a vested interest in your program (e.g., the program director, staff, a board member) does the same work, they will be considered an internal evaluator. There are gradients along this dimension; an evaluator hired by the funder who has very little contact with project staff would be “very” external; while a consultant hired by the program to evaluate its impacts might be considered more “internal.” An evaluation designed by an outside consultant but conducted by program staff would be even more “internal.” In all likelihood, an internal evaluator would probably have more information about what are the most important aspects of your program to evaluate and they would probably design a different evaluation. An internal and external evaluator might offer different interpretations of outcome data. Both kinds of evaluation have its pluses and minuses.



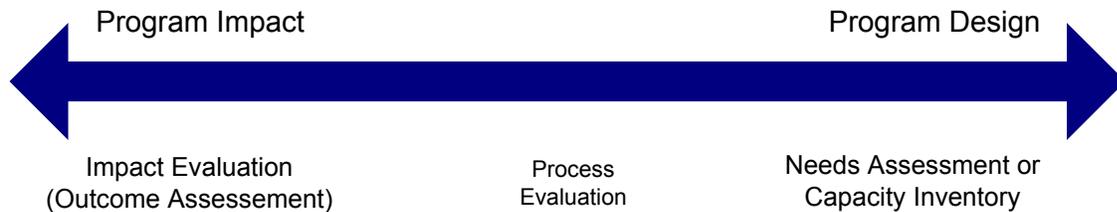
Level of Stakeholder Involvement: Participatory and Empowerment Evaluation

Outcome evaluation may incorporate strategies known as participatory evaluation or empowerment evaluation. Program values often contribute to the decisions CFPs make about the types of evaluation they conduct. An organization that holds a philosophy of grassroots organizing and/or an empowerment approach to community development may be more likely to engage in participatory evaluation.

In participatory evaluation, the purpose, design and implementation of the evaluation is determined by a large circle of “stakeholders” who may include participants, program staff, boards, volunteers, and funders as well as external evaluators. Those involved not only identify the evaluation issues and implement the evaluation, but also decide the action to be taken as a result of the evaluation findings.

Empowerment evaluation can be a form of participatory evaluation and be used as an outcome evaluation strategy. As described by American Evaluation Association, it “has an unambiguous value orientation – it is designed to help people help themselves and improve their programs using a form of self-evaluation and reflection.”⁴

The other end of the stakeholder involvement continuum reflects low stakeholder involvement, evaluation generally conducted by an outside evaluator.

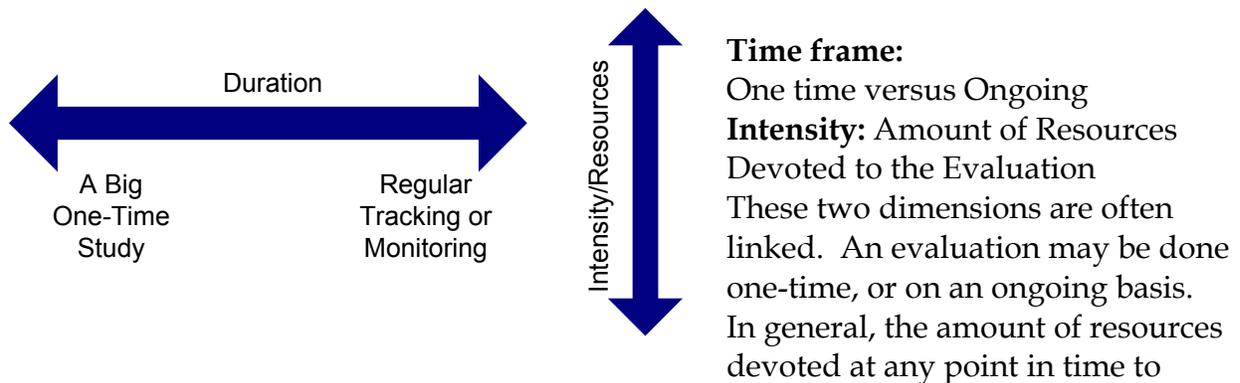


When the Evaluation Takes Place: Needs Assessment, Process Evaluation, Impact Evaluation
A needs assessment may provide a description of a community’s needs and assets. It is done to help plan a project (or perhaps to refine the mission and goals of an existing program). In an asset-based model, this might be characterized as a capacity inventory. The purpose of this type of evaluation is to understand the context of a project, and to help shape the activities that will be undertaken and the goals that are to be met.

The term "process evaluation" refers to a systematic method of assessing how well a program is operating, compared with the manner in which the program was intended to operate. The major emphasis in process evaluation is on documenting and analyzing the way a program has been implemented. Process evaluations often provide information on the number and types of participants served; the number and types of activities offered; and the program activities that lead toward outcomes and future sustainability. The primary purpose is to improve understanding of how a project achieves what it does. Program activities often measured by community-based programs include: networking/partnerships, capacity building, policy/procedure changes and community involvement. For a CFP some processes may describe numbers of volunteers, characteristics of clients, amount and kind of food grown and amount of food distributed.

Impact evaluation (also known as outcome assessment or outcome evaluation) addresses whether a program has successfully achieved its goals and objectives. It focuses on whether the activities of the program had their intended impact. According to the Bureau of Justice Assistance Evaluation Web Site, “Impact evaluations address those issues which are critical to the concerns of decision-makers, funders and the community. In a world where effective use of limited resources is a major concern, it is the impact evaluation which indicates whether the program has a positive impact and

can justify its strategy.”⁵ Outcome and process evaluations are closely related and interconnected because understanding how well a program has been implemented (process evaluation) is vital to interpreting the results of an outcome evaluation. Further, process evaluations can provide information early on for a program and can be used to determine if the outcome evaluation is worth pursuing.



ongoing evaluation is somewhat less than to a one-time study. Ongoing evaluations are often referred to as “monitoring” or “tracking.” However, these are not mutually exclusive. A project or organization may choose to conduct a larger-scale evaluation to answer a “big” question about the program, while monitoring smaller-scale outcomes on an on-going basis. Or, a program may decide that a large number of resources are needed to credibly measure its impact, and that it will do so only every few years, rather than on a continual or annual basis.

Purpose of This Handbook: Outcome Identification and Tracking

Summative? Formative? Quantitative? Qualitative? Internal? External? This handbook centers on the best strategies for demonstrating program effectiveness (by monitoring “outcomes”) rather than emphasizing program descriptions or evaluating implementation strategies. Outcomes are the benefits to participants or the community that are associated with or caused by a program’s activities or the direct products of those activities. The goal of CFSC’s Evaluation Program is to support grantees and others working in community food security in developing effective programs, to highlight their organization’s accomplishments, and to make it easier for them to meet evaluation requirements established by USDA. Building your capacity to continue internal program evaluation and outcome assessment on your own will meet these goals and build your organization’s sustainability. Focused on identification of outcomes and tracking, this handbook is intended to help you generate systematic, rigorous, credible and useful data about the impact of your community food project. We will focus on collecting data that are meaningful; data that programs will decide defines success for them.

The act of identifying outcomes may be as important as the measurement of those outcomes, because it helps program staff to focus on the goals of their programs and the expected impact on participants and the community. Any causal link from program activities to outcomes is one that savvy staff will be expected to propose without relying on data from social research but instead relying on changes in the program model that they will be monitoring over the coming years.

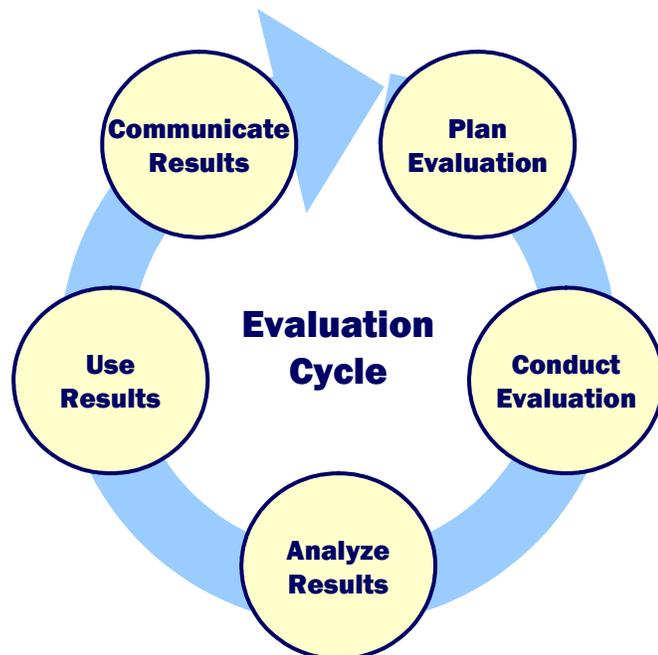
The Evaluation Process

As you think through and design your evaluation system, there are four questions you should keep in mind throughout the process.

Is the evaluation ⁶ . . .	
Useful?	Will the amount and type of information you collect meet the needs of those who intend to use the evaluation findings?
Feasible?	Will the evaluation be practical, doable and realistic?
Accurate?	Will the evaluation findings be correct?
Fair?	Will the evaluation be conducted with awareness of the rights of the people involved in the program and will it be fundamentally unbiased?

While you may not be able to give an emphatic “yes” to each question in every situation, you will want to always strive to come as close to a firm “yes” as possible. Accurate measurement of some of your outcomes may not be possible because of cost or complexity, but the measures you do choose should not compromise the rights of those being evaluated. Likewise, an evaluation is not worth doing if the results will not be used or your methods leave readers suspicious of your objectivity. Appendix I contains thirty standards that can be used as guiding principles as an evaluation plan is developed. These principles are organized around the four questions asked above.

There are five major steps in program evaluation: planning evaluation, conducting evaluation, analyzing results, using results for program improvement and communicating results to program stakeholders, community residents and potential funders.



This handbook dedicates a chapter or two to each step.

Chapter 2:	Mapping Your Project Using a Logic Model
Chapter 3:	Selecting Your Outcomes
Chapter 4:	Laying the Evaluation Groundwork
Chapter 5:	Selecting Evaluation Strategies and Study Designs
Chapter 6:	Designing and Customizing Evaluation Tools
Chapter 7:	Collecting Data
Chapter 8:	Analyzing Your Data
Chapter 9:	Understanding and Using Your Results
Chapters 10:	Communicating Results

The Culture of Evaluation

Evaluators are bound by a culture similar to members of other professions such as education and human services. The many descriptions of evaluation provided in this chapter should make it clear that professional evaluators have a specific way of looking at the world. One might even say they walk to the beat of a different drum.

Many evaluators believe that reality can be known and that it can be measured. Many evaluators believe that there are factors within program staff control that, if altered, may change the effects of programs on participants. Evaluators are often skeptical, tending to trust information most when it is collected with a plan that is replicable by those who love the program, those who hate the program or those who do not know the program at all. Evaluators tend to consider the world of community-based service programs as operating in a mostly causal fashion, with if-then consequences. This perspective is reflected in the logic model approach that is at the heart of chapters 2 and 3 of the Handbook. For more insight into the culture of evaluation, Appendix II includes a list of electronic evaluation resources that can be accessed on the Internet. These resources will help you learn even more about the world of evaluation and how evaluation strategies can work best for measuring your program's impact.

By reading through this handbook and participating in one of the CFSC evaluation workshops, you will become familiar with basic evaluation vocabulary and basic evaluation practices; you will begin a journey into the culture of evaluation. How far you go on your journey, or how exotic your journey becomes, will depend in large part on your enthusiasm and your willingness to immerse yourself in a new culture. We are here to guide you on the start of this adventure.

Working with an Outside Evaluator

Some CFPs may be fortunate enough to obtain help with their evaluation from individuals outside their organization. These outside or external evaluators may be paid contractors or volunteers. Regardless, there are advantages and disadvantages to working with outside evaluators:

Advantages

- Bring research knowledge and skills
- Lend credibility
- Create less burden on staff

Disadvantages

- Have less knowledge of program
- Are less participatory

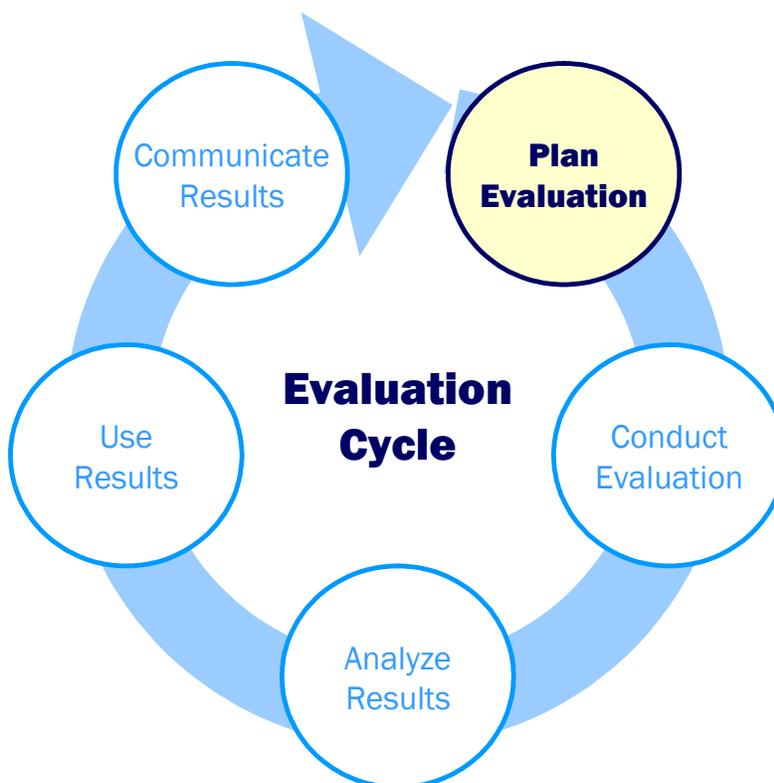
Organizations may have the opportunity to choose an external evaluator to work with and others may have the outsider chosen for them (due to a funder's requirements). If you have the choice, here are some things to look for in an external evaluator.

An outside evaluator should:

- Understand you and whom you understand (the community you serve)
 - Be able to work with diverse groups
 - Be able to explain the benefits and limitations of various evaluation methods
 - Be a good writer and clear speaker
 - Be willing to challenge the usefulness of evaluation activities
 - Be able to get the right descriptive statistics out of a computer
 - Be willing to take pieces of the project and allow you to handle some of the work in-house
 - Be someone who shares your evaluation philosophy
 - Understand the difference between research and evaluation
-

Chapter 2. Mapping Your Project Using a Logic Model

- 🌸 Revisiting Project Goals
- 🌸 Linking Program Activities to Goals:
Introduction to the Logic Model
- 🌸 The Basic Logic Model
- 🌸 A More Complete Logic Model
- 🌸 Steps to Developing a Complete Logic
Model



Chapter 2. Mapping Your Project Using a Logic Model

The evaluation planning stage ideally coincides with project planning. It begins at the conception phase when developing a new project, applying for a new grant or revising a program’s action plan. With evaluation, however, it is best to “begin with the end in mind.” What do you want your end result to be? In this chapter, we will first consider a project’s overarching mission and goals as these should always be in the forefront of planning efforts. They will form the backbone of your evaluation system. Then with these goals in mind we will embark on the logic modeling process, a way of mapping your project from start to finish, showing how these goals will be achieved.

Revisiting Your Project Goals

It will be impossible to identify meaningful outcomes in the absence of clearly stated goals that are understood by everyone connected with and, in fact, interested in your program. Whether your community food project is new or has been in existence for many years, taking a careful look at your goals is an essential first step in identifying the outcomes you would like to monitor. Keeping in mind the goals of your CFP, you might ask yourselves questions such as these:

Questions to Revisit Your Goals

- Are our goals consistent with and supportive of our mission?
 - Are our goals realistic in view of the resources we have to address community needs?
 - Are our goals broad enough to be useful and motivating to staff?
 - Are our goals focused and specific enough to be translated into measurable outcomes?
 - Are our goals reflective of the diverse needs of the various cultural groups served by our programs?
 - Are our goals designed to develop and foster authentic partnerships with growers, recipients and/or other stakeholders?
 - Do our goals include strategies to sustain our project, retain staff, find replacement funding and maximize other available resources?
-

Goals that do not meet your underlying criteria should be revised or replaced. As you consider the activities of your project, and the resources you have devoted to achieving your goals, you may find you either wish to revise the goals of your program to be in closer alignment with the services you actually provide, or you may wish to redesign your project so that goals important to your project can be attained.

Linking Program Activities to Goals: Introduction to the Logic Model

A logic model is a tool often used to tie a program to its evaluation. A logic model is a picture showing what you hope to achieve and how you plan to do it. It is comprised of “if-then” statements that describe a program’s theory of change, showing how day-to-day activities connect to the outcomes the program is trying to achieve. Similar to a flowchart, the logic model shows how program activities and outcomes connect with one another.⁷ The logic model has been likened to “a roadmap of your program highlighting how it is expected to work”⁸ or “the basis for telling a convincing story of a human service program’s expected performance.”⁹

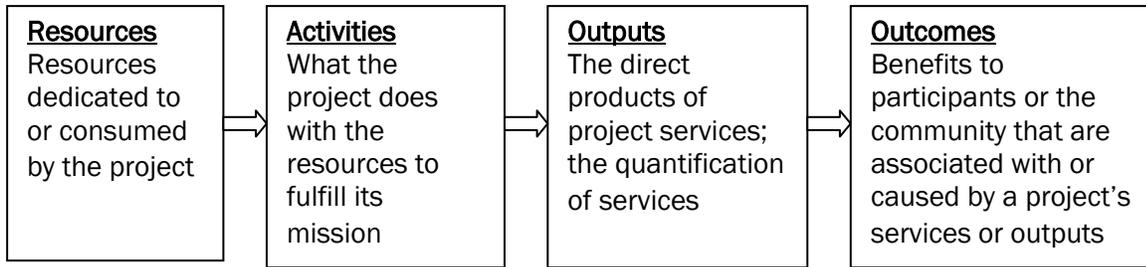
The logic model and its precursors have been used to understand the relationship between activities and results for the past two decades. Although the logic model was originally developed and used by evaluators, it has gained recent popularity for its use by program managers, program staff and funders. Some of the advantages of using a logic model are presented in the box below:

Benefits of Using a Logic Model

- It builds a common understanding of the program and expectations for its resources, activities and results, thus is good for sharing ideas, identifying assumptions, team building and communication.
 - It is helpful for program design or improvement, identifying activities that are critical to goal attainment, redundant or have inconsistent or implausible linkages among program elements.
 - It points to “a balanced set of key performance measurement points and evaluation issues, thus improves data collection and usefulness.”¹⁰
 - It ensures that a program’s process is not overlooked in an evaluation. The model makes it easier to look at both program process and outcomes.
 - It enhances the process of learning through evaluation. “As data are collected, the logic model can be used to put the data in perspective, examine the theory that underlies the program and make program mid-course corrections if needed.”¹¹
-

The Basic Logic Model

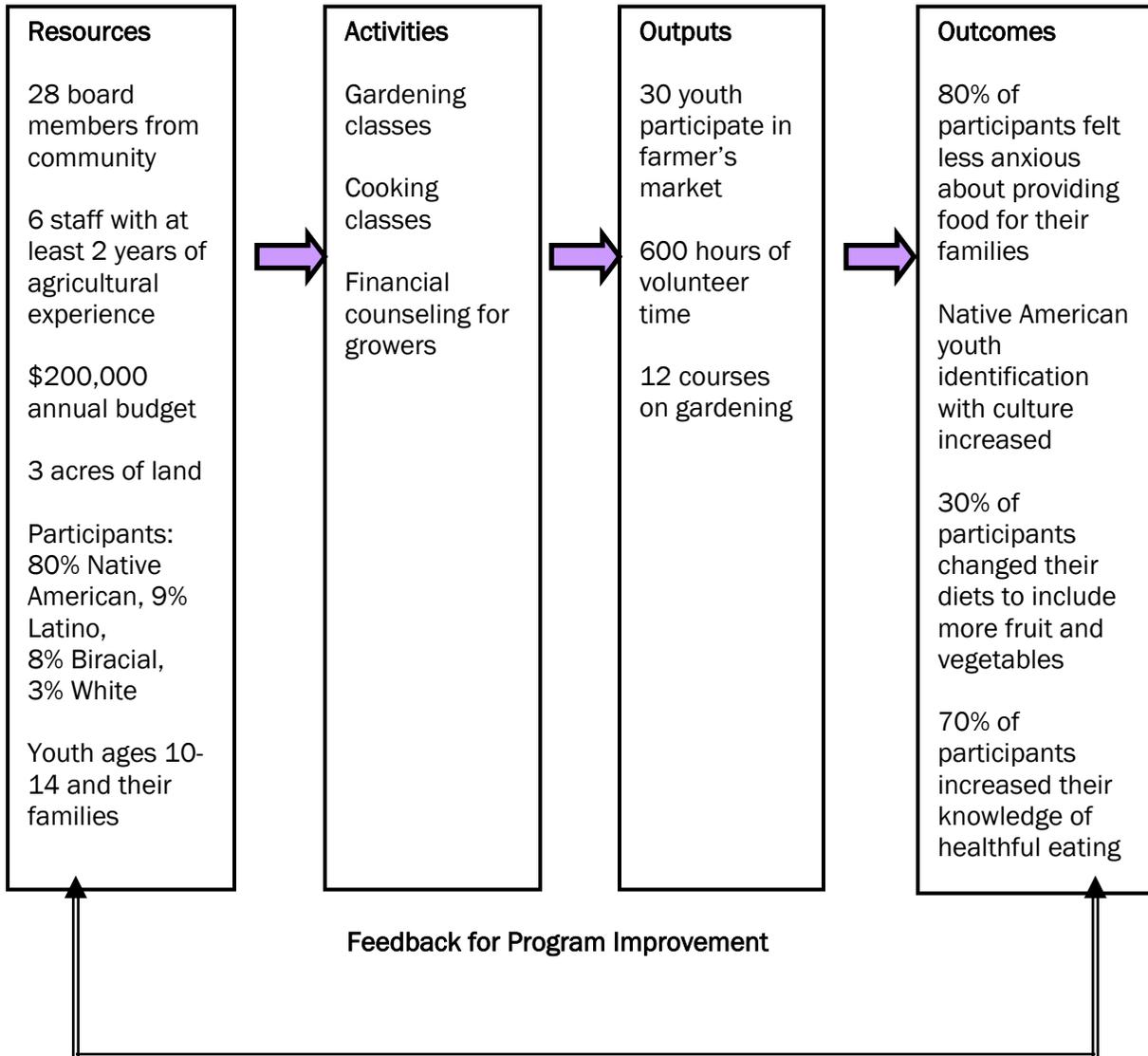
There is no one prescribed way to map a logic model; the specific elements of the model may vary in terms of complexity and the language used to define concepts. For community food projects, we have decided to begin with a basic logic model, one similar to that proposed by United Way of America.¹² The basic logic model is comprised of four components: resources (or inputs), activities, outputs and outcomes.¹³ The basic logic model, along with definitions, is described in the figure on the following page:



Examples of each component are presented in the table below. The next page displays a basic logic model from a fictitious community food project.

Examples of Basic Elements of the Logic Model			
Resources	Activities	Outputs	Outcomes
<ul style="list-style-type: none"> • Money • Staff • Volunteers • Equipment • Supplies 	<ul style="list-style-type: none"> • Mentoring • Technical assistance • Education • Nutrition counseling • Skill building activities • Policy advocacy • Provision of food 	<ul style="list-style-type: none"> • Hours of service delivered • Number of participants • Amount of materials distributed • Number of policies initiated • Number of organizations recruited • Pounds of food distributed 	<ul style="list-style-type: none"> • Increased knowledge • Changes in attitudes and values • Increased skills • Modified behavior • Improved condition • Altered physical and social environments

Sample Basic Logic Model



The Steps to Developing a Basic Logic Model

There is no right or wrong way to begin developing a logic model, the sequence of the steps presented below is only a suggestion. A worksheet on the following page is provided to help you develop your projects' basic logic model.

While we recommend starting with the end in mind – the goals of your program – information about constructing and choosing outcome measures is provided in the next chapter. As you read through these steps, and begin filling in the worksheet, you might want to leave the outcome square blank, or fill it in with your current ideas, and revise them after going through Chapter 3.

Step 1. Establish your **outcomes**. Begin with one of your project's goals. Translate this goal into one or more outcomes using Worksheet #3 from Chapter 3. Remember that outcomes are specific changes in project participants' behaviors, knowledge, skills, status and level of functioning¹⁴ directly resulting from a project's activities. Place these outcomes in the Outcomes column of the table on Worksheet #1. If you have completed Worksheet #3 and have decided on your outcome indicators and performance standards, you can also add these to the box.

Step 2. Enter your **resources**. Resources are those items dedicated to or consumed by the project (e.g., staff, facilities, funding, equipment, etc.). Place all the resources associated with your goal in the Resources column of the table of Worksheet #1.

Step 3. Enter your **activities**. Activities are what the program does with the resources to fulfill its mission. They are processes, tools, events, technology and actions¹⁵ used to directly serve your participants. Place the activities for your goal in the Activities column of the table of Worksheet #1.

Step 4. Enter your anticipated **outputs**. Outputs are the direct products of program activities; they are the quantification of activities (e.g., number of participants served, number of hours of service provided, etc.). Place the outputs associated with your goal in the Outputs column of the table of Worksheet #1.

Step 5. Repeat steps 1-4 for each of your program goals.

Worksheet 1: Developing Your Project's Basic Logic Model

Resources	Activities	Outputs	Outcomes*

** see Chapter 3 for more information*

A More Complete Logic Model

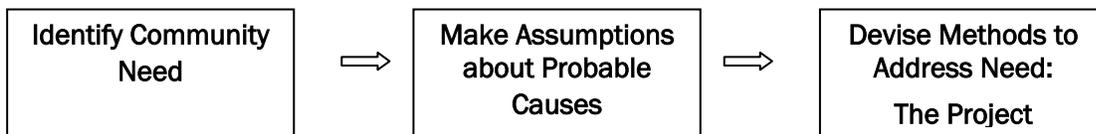
The basic logic model is a useful yet simple depiction of a project's operations and results. However, painting a more complete picture of your project may prove useful to better understand your evaluation results and more readily decide how to improve your project. The additional information and modifiers are useful in developing a more complete logic model and include providing more information on the project's planning phase, the community context in which the project operates, the population served and more detail on its outputs and outcomes. Adding components is not necessary to finish a logic model for your project because the basic model will suffice.

Enhancing the Logic Model

The basic logic model describes a number of important components of a program: resources, activities and results. However, an enhanced logic model can take into account other areas of the evaluation cycle. We recommend adding more information on your program-planning phase and on using results or the feedback phase in your project. (Additional information on using results is provided in Chapter 9 of this handbook.)

Program Planning

Although the basic logic model begins with resources and activities, there is an implicit social or community need¹⁶ underpinning the model as well as assumptions about factors causing the need and the actions that should be taken to strengthen the community.



Community Needs Statement: Understanding your community or target populations' underlying needs is critical in both framing and defining the type of program to deliver. Needs may also be defined in terms of "assets to be strengthened" rather than focusing on a "problem" or "deficits."^{17,18} The Economic Research Service of the U.S. Department of Agriculture has developed a set of tools to help communities assess their food security needs in its Community Food Security Assessment Toolkit (website: <http://www.ers.usda.gov/publications/efan02013/>).

Assumed Causes: Once a need is identified, the factors contributing to the need must be identified. These assumptions may be based on experience and/or research.

Method to Address Need: The program is born in this step of the process. Program mission and goals are determined which will lead toward meeting the identified need by addressing the causes.

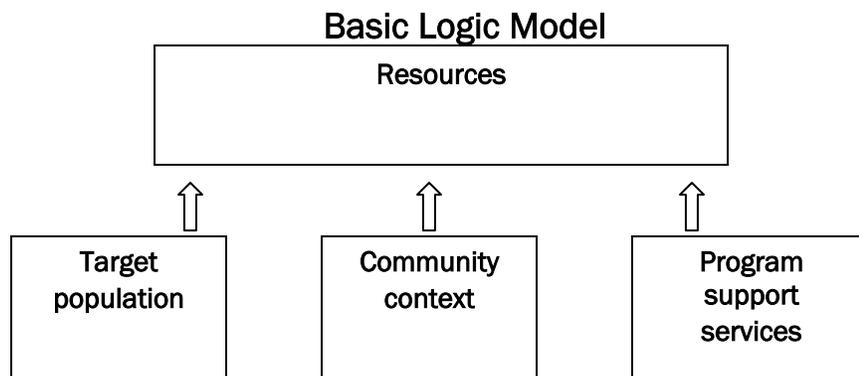
Using Results

An integral part of the original systems model on which the logic model is based is the feedback phase. Taking the evaluation results and reintroducing the information into the system to improve performance is the basis for learning and improving. Until one full evaluation cycle is complete, a program may not know the types of feedback or the place in the model where the data could be reintroduced. However, an arrow specifying this repetitive nature of learning should be added to the model to signify and remind readers of this important step. In Chapter 9 of this handbook, we will discuss how to add this information to the logic model and how to use it for program improvement.

Modifiers to the Basic Logic Model

Other types of information that may help further clarify the four components of the basic logic model are presented below. Again, these are components we feel might be useful to your project. You may select which, if any, you would like to add to your model.

Modifying Resources



Target Population: The population you serve or intend to serve will influence not only the project’s inputs, but also its activities, outputs and outcomes.¹⁹ Changes in population characteristics may influence project attendance, attitudes and behaviors.

Community Context: Defining the community context means identifying the conditions or events in the project, community or target population that may limit or expand the extent to which the project actually achieves its desired outputs and outcomes. The community context may also describe community resources and the regulations and policies that govern the service delivery.

Program Support Activities: Beyond activities provided directly for your participants, your project will engage in activities that supply the infrastructure necessary to provide quality activities. These might include building partnerships, promoting cultural competency, building capacity, board development, promoting sustainability and performing evaluation. Examples of these types of activities might be: weekly staff meetings, oversight by a multi-cultural board, evaluation team meetings, etc.

Modifying Outputs

Quality Measures or Outputs: As part of the Total Quality Management (TQM) movement organizations are urged not only to assess the amounts of their activities but the quality with which their activities are provided. According to TQM theory, productivity is increased when high quality program activities are provided and decreased when low quality programs are provided.²⁰ Examples of quality outputs: the proportion of individuals repeating enrollment, the proportion of participants that complete a specific course and participant satisfaction (one of our favorites).

Modifying Outcomes

Most of the literature on logic models breaks the outcomes category into two or three types of outcomes usually based on temporal relationships (e.g., short-term, immediate, intermediate, long-term, impact, etc.) We prefer the definitions used by the Harvard Family Research Program:²¹

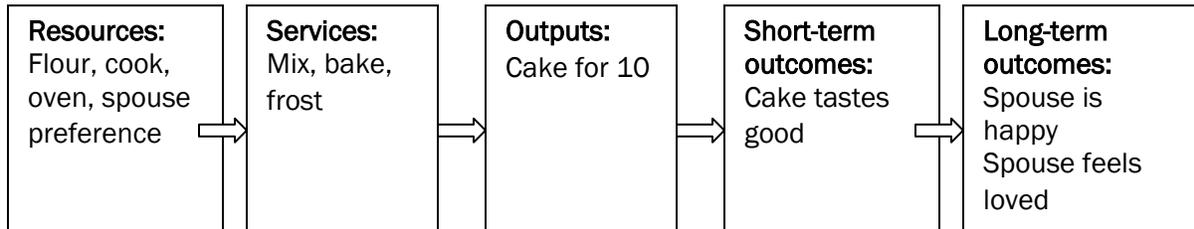


Short-term outcomes: The direct result of your program activities. They indicate a measurable change, and the language used often starts with “to increase” or “to decrease.”

Long-term outcomes: Changes in individual or group behavior or community conditions that a program hopes to achieve over time. Short-term outcomes contribute to the achievement of long-term outcomes, but other factors may contribute as well. It is important to remember, however, that programs typically are accountable for demonstrating success or progress in achieving long-term outcomes. As a result, they should be measurable and as specific as possible.

Outcomes are not intrinsically short-term or long-term. A short-term outcome for one program may be the long-term outcome for another. The major distinction is the sequence or order: the short-term outcome always precedes the long-term outcome.²² For example, changes in the larger community (e.g., shifting community food security) is often more appropriate as a long-term goal of your project rather than a short-term goal (e.g., shifting the food security of individual households).

Just to add some sweetness to the discussion, following is an example of a logic model for making a cake for a spouse’s birthday that illustrates the difference between short-term and long-term outcomes.

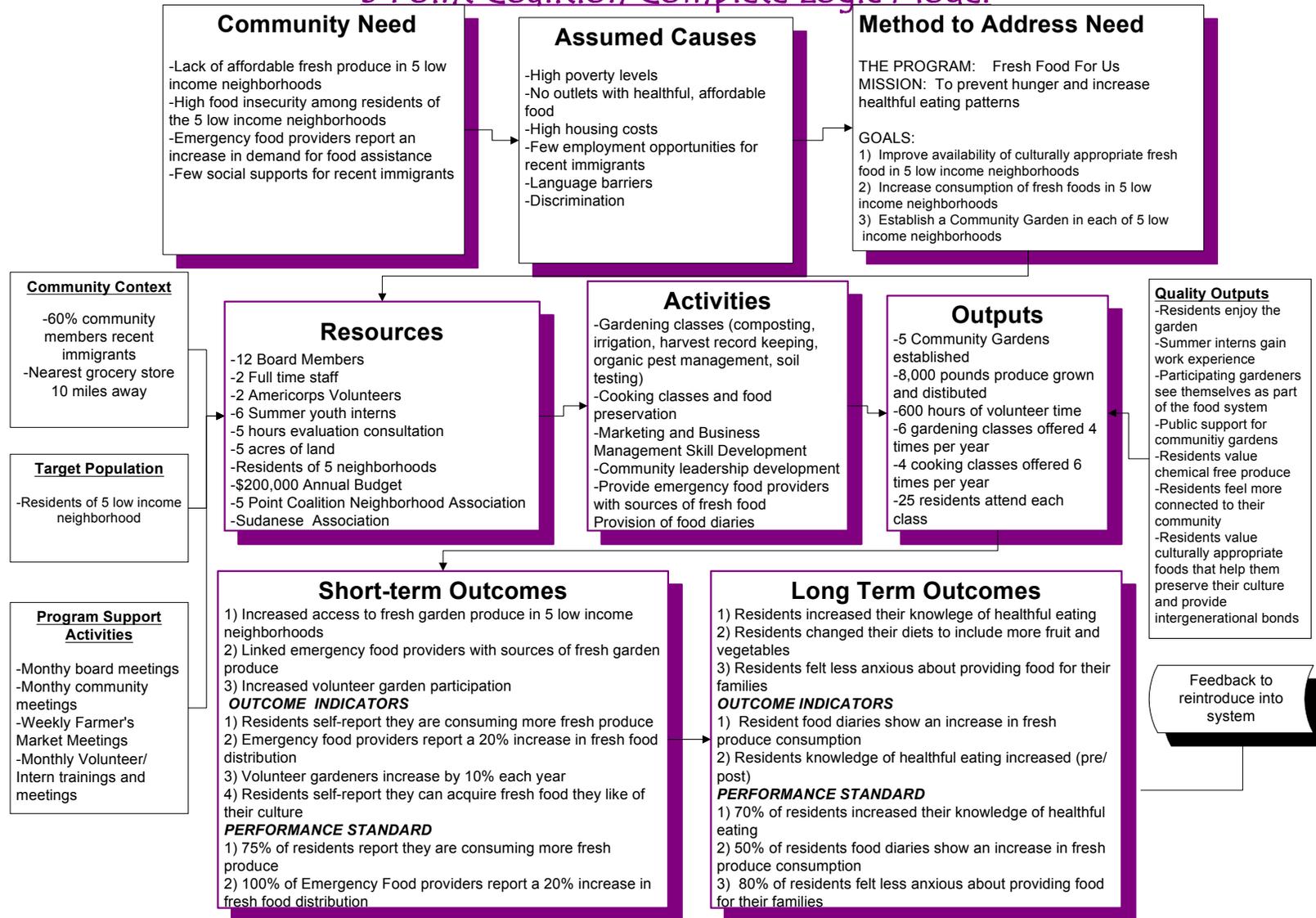


Examples of short-term and long-term outcomes in community food projects are shown below:

Examples of Short-term and Long-term Outcomes		
Program Focus	Short-term Outcome	Long-term Outcome
Increased economic security for local growers	Increased market for produce	Farm is still in business 24 months later
Increased access to healthy foods at school	Students report eating more salads at lunch	Students report eating 5 fruits and vegetables a day

An example of a more Complete Logic Model is presented on the following page. The model is based on a fictitious community food project and provides an example of each component that comprises this fuller program model.

5 Point Coalition Complete Logic Model



Steps to Developing a Complete Logic Model

As with the basic logic model, there is no right or wrong way to begin developing a complete logic model, the sequence of the steps presented below is only a suggestion. The worksheets on the following pages are provided to help you with your project's complete logic model.

Step 1. Specify the community need your project is designed to address.

Step 2. Specify the assumptions of cause you have made about the factors that probably influence the community need.

Step 3. Specify your method to address the need by describing your project, its mission and its goals.

Step 4. Establish your outcomes. Begin with your project's goals. Translate these goals into outcomes using Worksheet #3 in Chapter 3, if you have not already done so. Remember that outcomes are specific changes in program participants' behaviors, knowledge, skills, status and level of functioning²³ directly resulting from a program's activities. Categorize each of these outcomes into "short-term outcomes" or "long-term outcomes." Also, if you have completed Worksheet #1 and have decided on your outcome indicators and performance standards, you can also add these to the box.

Step 5. Enter your resources. Resources are those items dedicated to or consumed by the program (e.g., staff, facilities, funding, equipment, etc.).

Step 6. Specify the community context in which your project operates. Describe the conditions or events in the project, community or target population that may limit or expand the extent to which the project actually achieves its desired results. These items may also describe community resources, and the regulations and policies that govern the service operations.

Step 7. Specify the target population. List characteristics of your participants that may impact either the activities provided or the outcomes achieved. Common characteristics may be age, ethnicity and measures of socio-economic status.

Step 8. Specify the program support activities. These are the administrative activities provided to govern the project operations and ensure the core components of the CFP are addressed.

Step 9. Enter your activities. Activities are what the project does with the resources to fulfill its mission. They are processes, tools, events, technology and actions provided directly to the participants you serve.

Step 10. Enter your anticipated outputs. Outputs are the direct products of program activities; they are the quantification of activities (e.g., number of participants served, number of hours of service provided, etc.).

Step 11. Specify your anticipated quality outputs. Quality outputs are those measures that not only assess the amount of activities provided but the quality of the activities provided (e.g., participant satisfaction).

Worksheet 2: The Complete Logic Model

Program Planning

Statement of Community Need (Step 1):	Assumed/Researched Causes (Step 2):	Methods to Address Need (Step 3):

(Page 1 of 4 for Worksheet 2)

Program Implementation

Resources (Step 5) (The following sheet provides room for resource modifiers.)	Activities (Step 9)	Outputs (Step 10)
		Quality Outputs (Step 8)

(Page 2 of 4 for Worksheet 2)

Resource Modifiers

Resources		
Community Context (Step 6)	Target Population (Step 7)	Program Support Activities (Step 8)

(Page 3 of 4 for Worksheet 2)

Program Evaluation

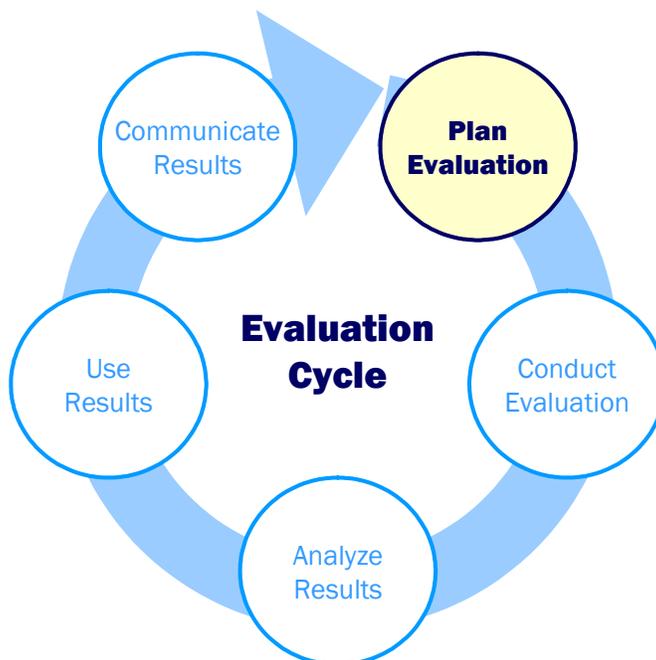
Note: You can choose to enter outcome indicator and performance standard information if you already know how these outcomes will be measured (Worksheet #3, Chapter 3).

Short-term Outcomes (Step 4 continued)	Long-term Outcomes (Step 4 Continued)

(Page 4 of 4 for Worksheet 2)

Chapter 3. Selecting Your Outcomes

- 🌸 From Program Goals to Outcome Measures
- 🌸 Outcomes in Daily Life
- 🌸 An Increasing National Emphasis on Outcomes
- 🌸 Organizational Benefits of Pursuing Outcomes Assessment
- 🌸 Lessons from Community-Based Organization Evaluations
- 🌸 Defining “Outcomes”
- 🌸 Setting Performance Standards



Chapter 3. Selecting Your Outcomes

Outcomes in Daily Life

Everyday we rely on simple measures to help us sort through the increasing mounds of information that come at us in magazines, newspapers, radio, television and the Internet. These important outcomes or “indicators of life” summarize complex messages, in shorthand, numbers or symbols. For example, many people rely on the Dow Jones Industrial Average to indicate the health of the economy. It used to be considered a measure of “standard of living” until broader thinkers successfully argued that money is not everything. It is unlikely that anyone you know, knows how the Dow Jones is in fact calculated, nor what the values of the Dow Jones mean. We all understand that 9,000 are better than 8,000, but 9,000 what? It is instructive to note how important this summary indicator has become in spite of its mystery. It is an important, if cryptic, indicator of the health of the U.S. economy. There are many examples of quantitative measures (that is, numeric ratings) we encounter frequently:

- Hotel quality and expense are summarized as quantitative measures in a series of stars from 1 (least expensive but probably has running water somewhere in the building) to 5 (you cannot afford it).
- Movie quality is summarized using a star (1-4) or thumbs up (0-2) system.
- Car safety, cost, comfort and reliability are summarized by a circle system (in Consumer Reports) where circles are empty, half filled or opaque.
- Americans seem riveted by rankings that indicate the best colleges, hospitals and places to live.
- The colored bands on meteorological maps of North America tell us how mild we can expect the weather to be.

An Increasing National Emphasis on Outcomes

The demand for participant outcomes is not unique to non-profit organizations and community-based programs. In government there is renewed interest in performance measures. In education the emphasis is on student and district assessment and in medicine a variety of organizations are generating indicators of good health care, which include a new emphasis on the patient’s quality of life.

These enthusiasms for measuring outcomes have long been ignited by the understanding that because of limited resources “not everything can be done [so] there must be a basis for deciding which things are worth doing.”²⁴ There has long been a growing mistrust of government and, now, thanks to computer technology, a new capacity to gather and understand vast amounts of data creating hope that complex measurement can be practicable. The G.W. Bush administration believes that accountability is everything (at least in education) and the Clinton administration

emphasized outcomes in its Government Performance and Results Act. Outcomes not only focus service providers on the results of their own work, they provide a credible source of information for funders who expect recipients of tax dollars or donations to be accountable to their investors.

Organizational Benefits of Pursuing Outcome Assessment

The process of determining program effectiveness can prove to be an invaluable tool in community-based service provision. The information collected will not only help managers maintain and increase funding for the program, but will provide staff with the feedback necessary to do what they do best: improve the lives of those they serve by improving their program. Outcome assessment can benefit management; demonstrate accountability; offer individual and organizational learning; and provide the foundation for a program's continued and increased sustainability.

Outcome Assessment for Management

Staff and weekly team meetings tend to be packed with problems to solve – how to keep communication active among different organizations' programs serving the same participant population, how to be certain that procedures for activities are being followed, how to help a particularly shy or isolated participant. What evaluation can do for managers and staff of a CFP is offer a different point of view.

Too often there is little time in the busy day of program managers to step back from the fray and view the program from a more dispassionate perspective, to ask, "Is this working?" "How well is this working?" "What needs to get better?" And even with time allotted for these sobering questions, there is often too little information by which to formulate a valid answer.

A good evaluation system will accomplish the following:

- Help staff answer these fundamental questions and free managers from the last minute rush to compile the evidence requested in grant applications from a growing variety of funders.
- Provide clarity to staff about what data they are or are not responsible for and give staff feedback about the success of their work, such that they would otherwise only get from hit or miss contact with individuals participating in activities.
- Reduce redundant data collection, produce meaningful measures of success and aim only at the essential information needed for management and funding.

While an evaluation system can serve as the instrument panel to tell managers where things are working and where they are not, it cannot tell managers exactly how to improve their programs when the warning lights flash. Skilled managers, like skilled

captains, must determine how to right the ship when turbulence hits. Getting the program moving forward smoothly requires the best thinking from all staff.

Furthermore, the first years of evaluation will not bear as much fruit for program management as can be expected when the system matures. As more data are collected, trends in areas of program success or failure will become apparent and characteristics of participants can be used to predict participant success or the need for greater staff resources.

Evaluation for Accountability

For many program directors, evaluation has less obvious value as a management tool and far clearer value as a tool for compliance with funders who want staff to be accountable for monies granted. The thoughts of one group in Virginia²⁵ about the value of what it calls “Results-based Accountability” are instructive for community-based service programs:

- Management by outcomes allows funders to focus on the “what” while providers focus on the “how.”
- Outcomes assure both funders and the public that their investments are paying off.
- Agreement on desired results across many organizations helps create a community-wide “culture of responsibility” for its members.
- A staff focus on desired results reduces the chance that many resources will be devoted to activities that do not contribute to participant or community improvement.
- A focus on results often forces the question of whether outcome expectations must be scaled back or project activities and investments must be increased because outcomes assessment may expose the fact that program providers are asked to accomplish massive tasks with inadequate resources.

Evaluation for Learning

Despite the fact that accountability represents a strong motivation for participating in evaluation, the intention of most funders of community-based organizations is to offer help to improve participant and community outcomes and improve opportunities to further sustainability. Technical assistance and evaluation training provided by CFSC are intended to enhance the capacity of community food projects to develop the systems they will find useful for improved program management, understanding program impacts and conducting effective fund-raising. Information derived from CFP programs’ evaluations should be used for learning more than judging.

The learning model of evaluation proposes that outcome measures are the culmination of a process in which program resources are expended to deliver services to

individuals, families or communities. (See Chapter 2 for more information on the learning model.) For service providers, outcomes play a special role. Rather than being an end result, outcomes are an integral part of a feedback loop. Monitoring outcomes should always result in review and evaluation of the program's goals, the adequacy of its resources and the effectiveness of its activities. The time and energy you spend in monitoring outcomes will be worthwhile only if you take all that you have learned about your program – where it works and where it falls short – and feed it back in to program improvement.

Evaluation for Sustainability

Evaluation results that show a program is indeed accomplishing or making progress toward its goals can be used to secure additional funding. The possibility that a funder or potential donor will be motivated to fund a program based on outcomes often creates the greatest motivation for staff to participate in outcome assessment. By devoting the resources necessary to learn the methods and vocabulary of outcome assessment, program staff invest in a service (identifying outcomes) that can show a great return for their program financially.

Programmatically, evaluation data can be used to strengthen a program, to make it more successful thus creating greater impact on the community they serve. As staff make changes in their programs based on participant outcomes, they create stronger, more effective programs. As the program becomes more efficient and more effective in fulfilling its goals, outcomes continue to improve; communities are better served. Additionally, staff that are aware of their program's goals, outcomes and participants can create innovative and engaging ways to change their programs increasing their own investment in the program.

Lessons from Community-Based Organizations Working on Evaluation

The United Way Outcomes Project in Milwaukee provides useful guides on what the likely and early benefits and barriers are to switching to more rigorous assessment. Human service agencies that had been involved with outcomes assessment for a year convened in Milwaukee to discuss outcome assessment's pros and cons. Some of their conclusions follow:²⁶

Cautions

- Shifting to outcome-based reporting is a process. Each organization should be met at its current capacity.
- It is difficult to switch from a mindset that values effort to one that values results.
- Building organizational capacity takes time.

- Imposing outcomes from above is unproductive. Outcomes development should be a collaborative process.
- Continued training is needed especially when staff turn over is high.
- Assumptions and findings must be revisited.

Barriers

- There may be resistance to change and some fear about the kinds of judgments that will be based on outcomes.
- Lack of time or resources from the existing pressures already heaped on organizations is a constant challenge.
- Further “train the trainer” workshops are needed so that they can go back to their own organization and train their staff.
- Agencies need for more technical assistance to reduce redundant measurement, develop better data collection tools, increase response rate, analyze and interpret data, etc.
- There is a need for more involvement of program staff, volunteers and boards of directors in development of outcomes.

The same organizations also acknowledged the benefit of outcome assessment. They agreed that the very process of identifying outcomes has resulted in programs that are better focused on obtaining results and that the outcomes approach provided the tools to set priorities, focus and challenge assumptions.

There is no denying that the road to outcome assessment is not as smooth or direct as many non-profit organizations and community-based service providers would prefer; however, the trip, once made, is well worth the effort.

Defining “Outcomes”

Many use the word “goals” and “outcomes” interchangeably. This can be true if your goals are both specific and measurable because an outcome is measurable by definition. Outcomes tell us how our participants and communities will be different as a result of our activities. Outcomes tell us if we are successful in achieving our goals.²⁷

The outcome question:

If we are successful in what we are doing, what change in our participants and communities can we expect to achieve and detect? In what ways will they be different than before?

Measuring Outcomes at the Individual Level

Typically, outcomes are specific changes in individual’s knowledge, skills, status and level of functioning²⁸ directly resulting from a program’s services. Outcomes can also

relate to changes at the neighborhood or community level, although it is the individuals within these communities where changes are generally more measurable.

Not all individual outcomes are necessarily behaviors. Changes in knowledge and attitudes are also worthy outcomes and may be more appropriate given the services provided by your program.

Examples of Individual Outcomes		
Knowledge	Attitudes	Behaviors
Knows the health risks of eating high calorie, low nutrition foods	Report that they feel more confident they can eat a more healthful diet	Changes diet to eat fewer high calorie, low nutrition foods
Knows more about own cultural tradition	Feels more connected to own culture	Participates in more traditions and practices of native culture
Knows which plants are native to region	Reports that growing native plants is a good farming principle	Grows more native plants

The three outcomes in the first column of the above table report on an individual’s knowledge. The three in the second column report participant attitudes and the last three are measures of participant behavior. Not only are all three categories of outcomes reasonable assessments of program effectiveness, there should be a connection between characteristics of programs and the kinds of outcomes they are expected to achieve. Implicit in all of these outcomes (and others) is the conviction that these changes in individuals help to improve their quality of life.

If we think how difficult it is to create enduring change in people, a model that suggests how change occurs may describe a person’s knowledge as the area of easiest change; the next more difficult change is a change in attitude (or feeling) and the most difficult enduring change would be a change in behavior. (For more information on the “stages of change” continuum of motivational readiness, see Prochaska, et al.)²⁹

Understanding a theory of how people make behavioral change will help you identify reasonable program outcomes. This notion, which suggests that knowledge precedes feeling, a precursor of behavior, can help guide programs away from anticipating profound changes in participant behavior after a low-intensity exposure. An example would be expecting participants to become sold on significant dietary modifications only after one or two educational seminars about the dangers of obesity. Such a seminar is unlikely to bring the intensity and duration of exposure we suspect is required to change long-term, eating behavior. But this type of experience might, in fact, be able to inform, so that an outcome measure of participant’s enhanced knowledge about the consequences of dietary modifications would be reasonable.

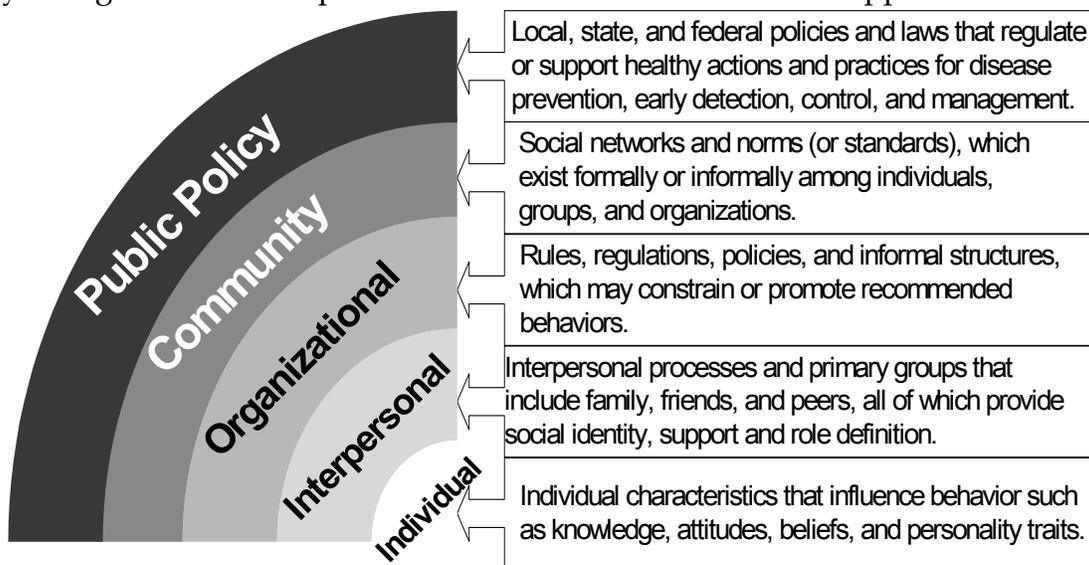
On the other side of the coin, a three-year intensive program designed to help promote life skills in urban youth should not merely settle for having an impact on a participant’s knowledge about appropriate gardening techniques. This knowledge change may be a short-term outcome that predicts the program’s long-term impact, but ultimately (and arguably after 3 year) it is reasonable to expect not only that participants’ knowledge will change but also they may show some behavioral changes as well.

Measuring Outcomes at the Community Level

Many community food projects work to affect changes not only on the individual level but also at the institutional, community and public policy levels. The broader goals of community food security address the “underlying social, economic and institutional factors within a community that affect the quantity and quality of available food and its affordability.”³⁰



A model of community change commonly embraced by public health advocates is the Socio-ecological Model, developed by McLeroy, et al. (1998).³¹ In this model, health and social issues are addressed through a combination of the efforts at all levels--individual, interpersonal, organizational, community, and public policy. This model focuses attention on the environment acknowledging that, although individuals have control over certain aspects of their behavior, many barriers and supports are not under their control. Environmental and policy changes are often required to remove barriers and create supports.



Although a common goal of food security program is to change the life conditions of a larger community, services are generally applied to specific individuals within a

community. Unless a program significantly impacts each member of the community, it is likely that community level measures will weaken the way program impacts are reported because “non-participants” will be thrown into the mix with others receiving your service.

To determine what outcomes or indicators to measure in your program evaluation, you will need to decide at what level(s) your program operates – the individual, community or both. Although the information presented in this handbook might be appropriate for evaluations at multiple levels, it has a stronger focus on measuring the success of CFPs at the individual or family level (e.g., the farmer, the family living in public housing, the youth, the school-age child, etc.). It is in compiling these individual measures of success, that a program can present a case for the impacts it has had on people’s lives.

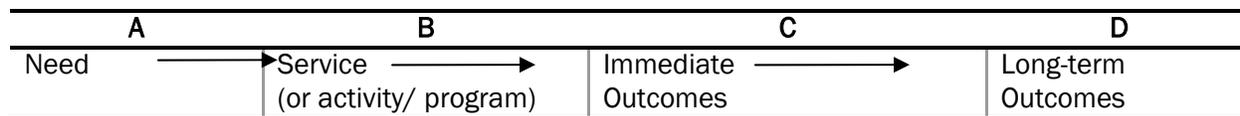
It is important to note that even if your goal is aimed at changes in neighborhoods, it is often more accurately measured through the individuals within the community of interest.³² It is often hard to achieve change at the community level without significant resources and a number of years of service delivery under your belt. It is often more appropriate as a long-term goal of your program rather than a short-term goal.³³

The Economic Research Service of the U.S. Department of Agriculture has developed a set of tools to help communities assess their level of food security in its Community Food Security Assessment Toolkit (website: <http://www.ers.usda.gov/publications/efan02013/>). This toolkit provides a wealth of information on methods to measure community food security and larger community indicators of program success.³⁴

The Use of Outputs as Outcomes in Community Food Security

There are a number of arguments that funders and stakeholders find compelling as reasons to stop short of measuring outcomes and simply use outputs as measures of success for some types of programs or program goals. As noted in the previous chapter, outputs are the direct products of program activities (e.g. number of participants served) and outcomes are specific changes or benefits to participants or the community that directly result from a project’s activities (e.g. increased knowledge or skills). Arguments for simply measuring outputs rather than measuring both outputs and outcomes fall into the following categories: 1) the link between the output and outcome is very well established, or 2) the cost to collect the outcome data is prohibitive given the funding level of the program.

The simple diagram (on the following page) demonstrates a measurement process similar to that of a logic model measuring program impact, such that a CFP might construct to demonstrate how their organization effects participant improvement.

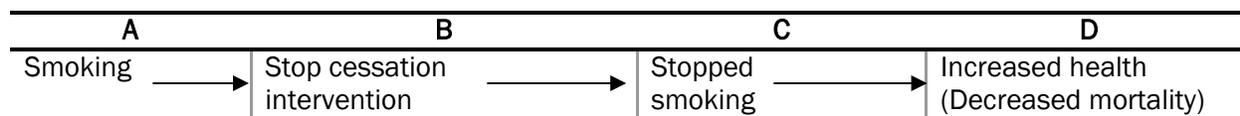


This simplified logic model starts with a community need that leads to the development of a program or service. That program will have short-term and long-term outcomes. Sometimes the provision of the services offered by the program is so obviously linked to intended outcomes that the mere provision of the service is proof enough that good will result. The example below, taken from everyday life, makes the point that some services need no measures of outcomes to prove their value.



We are cold. We wear a warm coat that keeps us warm in the short run. In the long run: being warm and comfortable (compared to being continually cold) keeps us healthy and content. Such a logic model may not suffice as proof that a program has had the intended impact any more than Einstein’s thought experiments provided the evidence needed to demonstrate the truth of the theory of relativity. But the logic model helps us understand where to look for impact of our services. Some logic models are more credible than others as substitutes for proof that an organization has a positive impact on its participants. For example, a program that feeds low-income families can reasonably claim that the link between food and health is so obvious that it need demonstrate only that food was provided and not that health was achieved. This is important because it costs much more to measure participant health than the pounds of food distributed.

Common sense is not the only criterion by which we might judge the adequacy of a claim of success based on the logical link of a service to an outcome, or the link between a short-term outcome and a long-term outcome. More often we will rely on research to establish this link. Smoking cessation interventions may provide a good example.

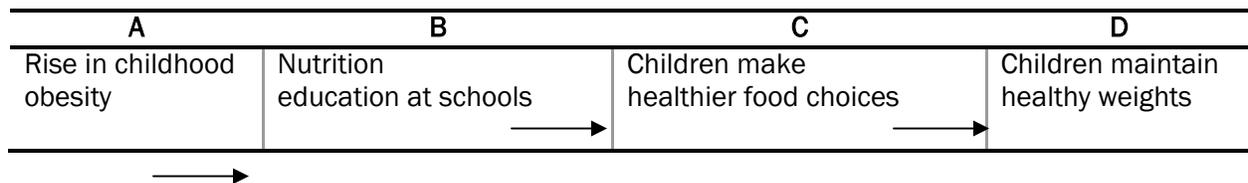


We do not require a smoking cessation clinic to prove that once their participants have stopped smoking their health status improves. Instead, we accept the evidence generated by large (expensive) research studies that show that those who have stopped smoking are healthier (e.g., die less often from lung cancer or heart disease) compared to those who have continued to smoke.

Another example of an instance where an “output” might be considered sufficient is in the case of policy advocacy. Mandated policies and programs that force environmental and organizational changes may not need to go the extra step of measuring outcomes at the individual level.

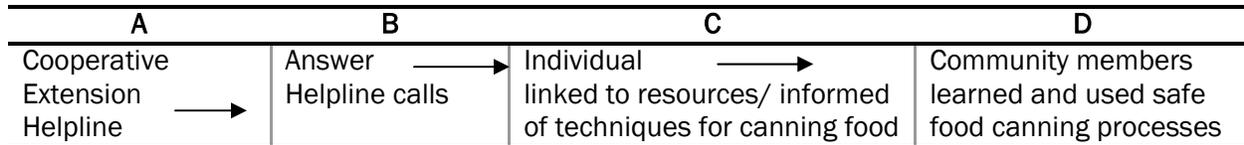


There is a greater burden placed on the organization required to demonstrate its outcomes than the organization that may report only its outputs. In health care, large well-controlled research studies have provided strong evidence that health screenings have important short-term and long-term positive outcomes for patients and there is relatively little deviation across programs in how these screenings occur. Such definitive studies and rigorous service procedures are less the norm for other community-based programs. Consequently, the link between service provision and positive impact on participants (as shown in the example below) is established by neither profound common sense nor strong research.



The research may be mixed on the impact of programs designed to educate kids about nutrition. And even if there were high quality definitive research findings, the specific nutrition program offered may or may not produce the effects seen in a controlled national research study. Consequently, it may not be enough to have an education program report that it distributed 5,000 brochures or gave presentations to 2,000 area high school students. Such outputs may or may not have led to the desired outcomes – healthy weights of kids. While we may have justification for asking programs to produce outcomes as well as outputs, we must remember that to produce credible outcome data, a program will have to devote meaningful resources to the endeavor. Imagine what it takes to track service recipients to discover if they have achieved a healthy weight one year later – let alone for many years down the road.

Another rationale for allowing the use of outputs rather than requiring outcomes is based more on practicality than scientific evidence: when the time and resources needed to establish the proof of effectiveness is greater than the time and resources necessary to provide the service, it makes sense to accept a lower level of evidence. For example:



In the case where the Helpline provider may only spend minutes with a caller, then requiring a follow-up to determine if resources or techniques were actually used and sustained might be a poor use of program resources. In the case where little energy is expended by an organization to serve each participant, large expenditures to produce evidence of effectiveness should not be required. A 10-minute phone call is not enough of a treatment to produce the profound outcome demanded. In epidemiology, the phrase “dose-response” is used to describe this relationship. A dose (or service) that is too small is not expected to result in a noticeable response (or outcome).

Thus, CFPs are encouraged to choose appropriate outcomes that gauge the impact of their programs. However, it is understandable that there may be cases where a measure that would normally be labeled as output can stand in the place of an outcome. This may be especially true for food banks. In other cases, a short-term outcome may be sufficient, as the link to the long-term outcome is understood.

Setting Performance Standards

An outcome can consist of two components – an indicator or measurement and a performance standard.

An indicator is the specific information that will determine how well the program is doing at meeting its outcome goal. It is what is measured by a questionnaire, participant interview, staff observation; test scores, presence or absence of a particular behavior or event, etc. (e.g., improved gardening skills). Specifying an outcome indicator consists of:

- 1) The specific observable, measurable characteristic or change that will represent achievement of the outcome; and
- 2) The specific statistic(s) (e.g., the number, percent, average rating) the program will use to calculate and summarize its level of achievement.

A performance standard is the level or amount of change that is expected to be achieved in an indicator (e.g., increase in the number of vegetables eaten). Performance standards have more to do with motivation than with measurement. A good evaluation does not necessarily require performance standards, but meaningful performance standards can help keep a project focused on its key intended impacts. If program staff think of performance standards as personal challenges or as indications of good work done for participants or the community, achieving those standards can become reasons to celebrate. Bear in mind that neither the celebration nor the motivation can occur

unless staff review the data from the evaluation. It is these data that staff have agreed to collect and it is from these data that staff can learn how well they are achieving the outcomes that all agreed should be accomplished.

How to Set Standards

How much change is reasonable to expect? As with the selection of indicators, you do not want to over promise or aim too low. Learning about what other programs similar to yours have achieved will be helpful in setting realistic performance standards for each of your outcome indicators. Obtaining information on the status of your service population nationwide, in communities where programs like yours do not exist, or in your own community prior to the inception of your program may also help you to select an appropriate target. This information describes the “natural” status of your community when they have no access to the kind of assistance you offer, so it provides a minimum performance standard that you should expect to exceed.

For example, your program helps local farmers bring fresh produce into high school. You also fund significant education efforts in these schools about nutrition and the importance of the 5-A-Day Program. You know that in your state, about 15% of youth eat 5 fruits or vegetables per day. As a starting point, you might expect that you are close to the state average for youth of 15%. Given the types and intensity of the services you offer and the outcomes that similar programs have achieved, how much should you expect to increase the average number of fruits and vegetable above this benchmark? Here is where educated guessing comes in. Using what you know about your program and the high school you serve, you should choose a benchmark significantly higher than the pre-program average of 15% but not so much higher that you could never achieve it.

In the following table are a few specific examples of community food project outcomes and accompanying indicators and performance standards.

Sample Goals, Outcomes, Indicators and Performance Standards			
Program Goal	Outcome	Indicator or Measure	Performance Standard
To increase gardening skills for participants	Increased knowledge of gardening practices	Knowledge score on garden skill inventory	On average, a 50% increase in knowledge in a pre/post test
Latino residents will report greater satisfaction with produce offered at neighborhood farmer's market	Satisfaction with food selection	Average rating on market intercept satisfaction rating scale	At least 80% of Latino shoppers will report being "very satisfied" with the variety of food choices provided at market
To increase civic responsibility	Community service	Number of hours spent in community service	90% of youth will participate in at least 20 hours of community service per semester
To increase the availability of affordable, fresh produce	Access to fresh produce	Number of families reporting they have increased access to fresh produce	50% of the families participating in the co-op will have greater access to fresh produce within 6 months of joining
To increase collaboration of food-related community organizations	Increased collaboration	Number of organizations participating in food network	8 community-based organizations will attend monthly network meetings

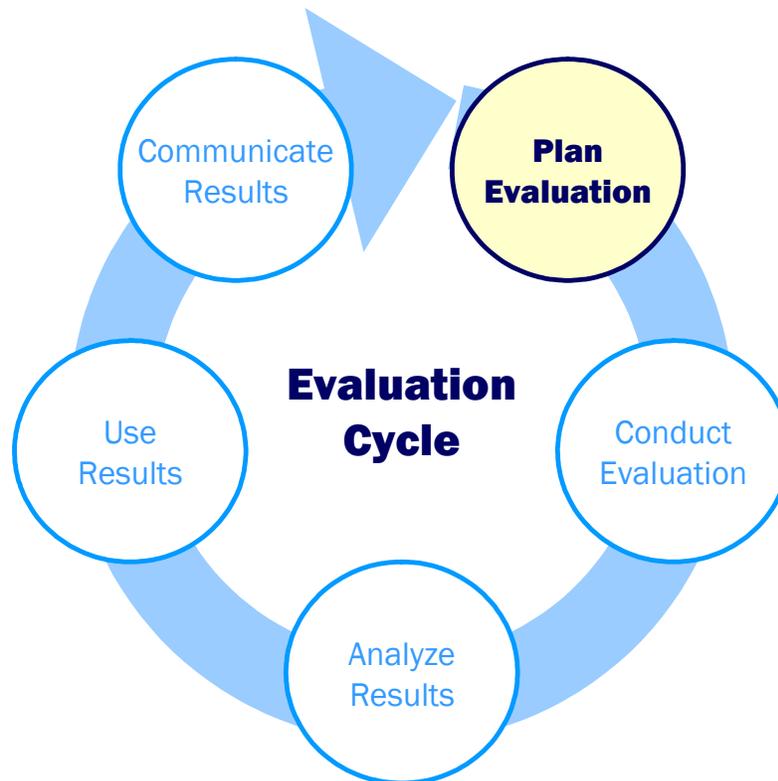
Worksheet 3: Outcomes, Indicators and Performance Standards

Note: You may not know what types of indicators and performance standards might be best suited for your program. You can use this worksheet first to brainstorm and then later complete it in a more formal way. For examples of outcomes, indicators and performance standards, see the text in Chapter 3.

Outcome	Indicator or Measure	Performance Standard
1)		
2)		
3)		
4)		
5)		

Chapter 4. Laying the Evaluation Groundwork

- 🌸 Establishing a Learning Environment
- 🌸 Engaging Program Stakeholders in Evaluation
- 🌸 Assessing the Data on Hand
- 🌸 Developing an Evaluation Plan



Chapter 4. Laying the Evaluation Groundwork

In the beginning phase of program development, investing time up-front to design your evaluation will ensure it focuses on the relevant questions, results in reliable data and provides meaningful answers. Although non-profit agencies and community-based programs rarely have a shortage of things to do, evaluation cannot be done in the bits of down time pieced together in a short amount of time. A system developed as an afterthought will demonstrate, over time, that “something” can be worse than “nothing.” Therefore it is essential that programs perform some evaluation groundwork before embarking on the actual evaluation. In this chapter, readers will be asked to consider laying the groundwork for evaluation by establishing a learning environment, engaging stakeholders, assessing the data they may already have on hand and considering development of an evaluation plan.

Establishing a Learning Environment

Organization leaders and their staff are often hesitant to engage in outcome assessment because they fear negative results. It is normal for even the most competent staff members to be somewhat threatened by evaluation. Creating a “learning” environment, rather than a “judging” environment can help diminish these fears. This learning environment is best integrated when programs are first conceived. The distinction between the two environments is made clear in the following figure:

The Differences Between Judging and Learning Environments ³⁵	
Judging	Learning
Makes Judgments	Changes Behavior
Documents Error	Uses Mistakes
Focuses on Problems	Focuses on Opportunities
Affects Programs	Affects People
Is Reactive	Is Proactive
Is Adversarial	Is Cooperative
Is for Funder	Is For Staff

Only when outcomes are measured in a learning environment will services to participants truly improve. To create a learning environment in your organization, a number of steps can be taken:

- Consider how to bring all different levels of staff, volunteers, participants and board members into program planning and evaluation planning discussions. Together create what The James Irvine Foundation calls a “culture of inquiry”³⁶ by openly discussing your mission, values, the assumptions you make about your programs and the hopes and dreams you have for achieving success.
- Involve staff and other stakeholders in evaluation and assessment tool design. Engaging staff members in development of the measurement

- program will increase staff ownership of the data. Not only will they be more motivated to contribute to the assessment process because they understand its purpose, but they will also have better knowledge of how the system works to ensure accurate, consistent data collection. Further, when staff feel as though the system is designed to provide them feedback, the likelihood of the results being used to improve activities increases.
- Use the first round of results as “pilot-test” data. These will be used to help modify the system, and will not be released to outside agencies unless staff agree that the data can be disseminated. This will give staff a chance to work on any program element before being held accountable by the public.
 - “Walk the talk.” Do not just provide lip service to the idea of a learning environment -- actually promote it. If a staff member receiving a negative service evaluation is fired due to the ratings, the trust of staff will be shaken and the quest for the collection of honest evaluations will be defeated.
 - Do not punish or blame others. To get the full benefit from your evaluation you must be able to learn from the results and link them back to your program. Fear of punishment and blame will threaten staff performance rather than motivate staff improvement. Develop a common practice for sharing outcomes that informs staff but does not threaten a particular staff or program.

Staff and volunteers in all levels of your program may be impacted by evaluation either because their work is being examined or they are asked to help collect the data. For that reason setting up the right environment will be important. The James Irvine Foundation explains that a project’s success with evaluation is often determined by whether it is able to “create a culture that value(s) the process of self-evaluation.” Self-evaluation may require a shift in organizational thinking, in mindset, in your everyday norms and practices.

Engaging Program Stakeholders in Evaluation

It is important to include a program’s major stakeholders when planning an evaluation. It is these people with vested interest that will ensure the future usefulness of the study. To involve stakeholders, it may be helpful to first create a list of people or groups who would use the program or evaluation results. It can include groups of people currently at the table and some who are not. Stakeholders may fall into one of the following categories:

- **Implementers:** people or groups involved in program operations
- **Partners:** people or groups who actively support the program
- **Participants:** people or groups who are served or affected by the program
- **Decision makers:** people or groups in a position to do or decide something about the program

Your evaluation stakeholders will be a subset of all program stakeholders. Examples of possible stakeholders for CFPs are shown below.

Examples of Stakeholders for Community Food Projects³⁷

Community sector

- Target audience members
- Neighborhood associations and residents (as well as community)
- Farmers/growers
- Food retailers/distributors
- Churches and other religious institutions
- Youth groups/organizations (as well as youth)
- Senior groups/organizations (as well as seniors)
- Funders

Government sector

- National, state and local elected officials
- USDA
- Regional or local planning commissions
- State, county or city departments of education, economic development, community development, agriculture, children and families, tourism, etc.
- Block grant agencies
- Law enforcement agencies
- Public housing communities

Health sector

- Wellness councils or coalitions
- Physicians in private practice
- Insurance companies
- Public health departments (nutrition education departments, environmental health, adolescent health, etc.)
- National and state nursing and medical associations
- National and state health education associations

Education sector

- Universities, colleges, Technical schools
 - Cooperative Extension Service
 - State & local chapters of professional teacher/administrator associations
 - Public and private primary and high schools (as well as students)
 - Business sector
 - Chambers of commerce
-

Once you have identified your stakeholders you should engage, as possible, representatives of each group to answer questions such as

- What is important about this program?
- What kind of results would you like to see from this program?
- How will you use the evaluation results?³⁸

The Importance of Stakeholders in Culturally Responsive Evaluation

It is essential for programs to engage with community stakeholders when working with culturally diverse groups. In order to be culturally responsive, it is important to work along side stakeholders who are from, represent, closely understand and are connected with the population being evaluated. These cultural stakeholders may include representatives from community agencies, organizations providing support services, culturally diverse professional evaluators, formal and informal community leaders or other persons of trust (e.g., tribal leader, business owner or a pastor) and the participants themselves. “By being as inclusive as possible in the evaluation process, an evaluator can be respectful of the community and its multiple cultural perspectives.”³⁹

The help of these cultural stakeholders should be enlisted from beginning to end of an evaluation. They may provide advice on the evaluation plan and help guide efforts throughout the evaluation process: aiding in selection of the appropriate evaluation methodology; developing evaluation questions; guiding and possibly assisting in evaluation participant recruitment, advising on administration logistics, consulting on interpretation of evaluation results and helping channel communication of the results with the target community. Further, cultural stakeholders may serve as advisors or brokers with certain populations to help an evaluation team grow their respect, knowledge, awareness, sensitivity and appreciation for the diverse group being evaluated.

When considering how to include cultural stakeholders on your evaluation team, consider a broad definition of the special populations your program might impact. Cultural diversity may include, but not be limited to a person’s race or ethnicity; native language; sexual orientation or identity; socioeconomic status; physical, mental or emotional health or abilities; age; geographic region; or spiritual beliefs.

Assessing the Data on Hand

Once you have developed your logic model and have identified the outcomes you plan to monitor, at least to start, you will need to inventory the resources required to collect, process and analyze the information and communicate the results. These resources include:

- Data
- Staff / Volunteers
- Technology

Data Inventory

A critical step in developing your evaluation is identifying all the information you currently collect that is relevant to your service delivery and your desired outcomes.

This collection typically includes:

- Information that can be used to measure outcomes (e.g., registration forms, satisfaction surveys, etc.), and
- Information about program operations, much of which may be related to outcomes. This information will increase your understanding of what makes your program effective and for whom it is most effective (e.g., resources, outputs, etc.).

First, compile all the forms you now use to collect data on your participants (e.g., registration forms, satisfaction questionnaires, skills inventories, needs assessments, service history, etc.). For each form, identify the following information using Worksheet #4 (an example is shown on the next page):

- The types of data collection forms already in place
- How much information is required on the form (i.e., the number of items in the form and how much time it takes to collect them)
- The type of information collected
- The staff responsible for administering or recording and collecting the form
- Who is eligible to complete the form
- When and how it is administered

The table on the following page demonstrates a convenient way of organizing this information. We have filled in the table with some information from a fictionalized community garden. A blank form for your use follows as Worksheet #4. When the form is as complete as possible, distribute it to all staff members to make sure the information is accurate and nothing has been omitted.

Example of Data Currently Collected by Community Garden				
Data Collection Form	# of Items	Type of Information Collected	Who Administers	Methods & Timing of Administration
Program Intake Form	68	Contact information: address, phone, SS#, work phone #, emergency contact information, doctor's name and phone # Participant characteristics: age, sex, ethnicity/race, household income, free lunch eligibility, family composition	Assistant to Director	Completed by participant at enrollment
Daily Roster	N/A	Daily counts of participants at gardens	Garden leader	Sign-in sheet completed by participants
Participant Satisfaction Survey	26	Participants' ratings of satisfaction with program, staff and the activities offered	Garden leader	Pen and paper administration during the last week of the program
Volunteer Focus Groups	N/A	Volunteers' satisfaction with program environment, perceptions of impact on participants, and personal benefits of program	Assistant to Director	Two focus groups of 8 volunteers are held annually

Worksheet 4: Data Resource Inventory

Data Collection Form	# of Items	Type of Information Collected	Who Administers	Methods & Timing of Administration
1)				
2)				
3)				
4)				
5)				
6)				

When you review the status of your data collection (perhaps at a staff meeting devoted to this purpose), address the following questions:

Review of Your Current Data Collection System

- Is redundant information being collected?
 - What data collection tools do you feel are working well, in terms of ease of administration, response rate and accuracy of the information collected?
 - Do any tools need to be revised or replaced?
 - Is the amount of information manageable, given your resources for data collection and processing?
 - Prioritize the information you are currently collecting. What is most useful? Next most useful? And so on. Ask yourselves:
 - *“How will we use each type of information to assess program effectiveness?”*
 - *“What’s it good for? What will we do with it?”*
 - *“Will we be willing and able to make program changes based on what we learn from this information?”*
 - Is data that was originally required by funders still required?
 - What are the gaps in the information you are now collecting?
-

Each question is repeated below with an explanation of its relevance and an example.

- Is redundant information being collected?

In most cases, redundancy is unintentional, wasteful and should be eliminated. In some cases, though, it will be useful, because program staff and participants may provide different levels of information. You will need to decide which areas of redundancy should be eliminated and which should be retained.

For example, a participant involved in a round of intake interviews with staff may feel more comfortable discussing sensitive topics such as household finances with program staff than a volunteer. Others may feel equally comfortable (or uncomfortable) with staff and volunteers alike. For sensitive topics that have a specific and essential purpose, such as designing individualized program plans, you may want to consider collecting the same information by different staff or from different people interacting with the participant. As a general rule, though, you should review all data collection forms with the intent of eliminating redundancies.

- What data collection tools do you feel are working well, in terms of ease of administration, response rate and accuracy of the information collected? Do any tools need to be revised or replaced?

It is not uncommon for tools developed years ago to continue to be administered even though current staff does not find the data useful or relevant. In some instances, staff question the validity of the data collected, but have not had the time or resources to determine if there are better ways to ask the questions.

Another common problem is the lack of consistent administration of data collection tools. When determining if a tool is useful, make sure to explore how systematically data have been collected. This may explain why the data are not meaningful.

For example, if a program does not consistently collect pre-test data on the first day of programming but sometimes administers it a month into the program, comparison to post-test results will underestimate the program's impact on participants because pre-test results occur after participants already have participated in some activities.

Looking at how consistently a tool has been applied can also provide insight on the burden of its administration. Data that are rarely collected on time, or as intended, should be scrutinized to determine if there is a better method of administration.

- Is the amount of information manageable given your resources for data collection and processing?

There is a tendency to collect far more information than will ever be used. Different staff members may collect unnecessarily redundant information; data collection tools may be too long and complicated; or you may be collecting information that is not very useful.

- Prioritize the information you are currently collecting.

Ask yourselves:

- "How will we use each type of information and which is the most useful?"
- "Will we be willing and able to make program changes based on what we learn from this information?"

All data included in your evaluation system should have a well-defined purpose, either as an outcome measure, an output measure or to increase your understanding of how your program works (e.g., participant's gender, race/ethnicity, length of program participation or number of hours of each kind of activity). You may decide that you no longer need to collect certain types of information. Or you may decide to continue to collect the information, augment it and include all of it in your evaluation system.

For example, a vocational training program has, for a number of years, collected staff ratings of participants' progress made in knowledge and skill attainment. However, the program is now beginning to collect information from the participants using pre-post knowledge and skill tests. As a result, the program director may choose to be more selective in the information they collect from staff.

- Is data that was originally required by funders still required?

We are aware of an instance where program staff continued to collect information required at one time by a funder only to discover much later that the data were not needed. Staff had been complying with what they believed remained a requirement and they were mailing data to the State each month. Someone eventually told them that the program had been dissolved.

- What are the gaps in the information you are now collecting?

Do you have accurate and complete information on all the outcomes you want to monitor? What else do you need to know about your participants and their experience with your project in order to increase your understanding of which participants benefit the most from your project? Which service components yield the most positive outcomes? What is the optimal length of time for project activities? As with the information you are already collecting, you will need to set priorities for the new information you would like to collect. What will be a significant enhancement to your assessment of outcomes? What would be nice to know but can wait? Funding priorities, feedback from participants and stakeholders about the most/least beneficial aspects of your project, cost of data collection and information about the effectiveness of similar projects should help you to make these decisions. Even if you do not have the resources to collect the data at present, you may want to create a “wish list” of these types of information to add to your evaluation system as it matures.

Staff/Volunteer Inventory

Building and maintaining an evaluation system consists of much more than indicators and data collection forms. It also requires people to perform specific tasks such as developing and/or revising evaluation tools, data collection, processing, and analysis and report preparation. In order for these tasks to be performed effectively, staff and volunteers must understand the value of outcome assessment and the importance of their roles in the assessment project; feel that each is the appropriate individual for the task; and have the flexibility to fit assessment activities into their regular work flow.

A fictional example of a staff inventory follows. A blank form is included as Worksheet #5 to help you to organize the information necessary to assess staff capabilities and availability to implement evaluation activities.

Staffing Resources

For many programs, the question of where the staff time to support evaluation will come from is not an easy one to answer. Since the primary business of any community-based service organization is to provide service, dollars for administrative costs may be limited. Staff may already work more hours than they are paid for. In addition, staff have a variety of skills. They have been hired for their expertise with activities you provide. They generally are not evaluators and although non-profit staff members are

well known for juggling many different roles, neither the interest nor ability to efficiently manage an evaluation may be present in the beginning.

If the amount of person power is a concern, consider utilizing your volunteers or high school student interns who may be able to benefit professionally or educationally from participating in the process. If data entry is needed, consider cooperating with an agency that works to develop such skills in their clients. Finally, local college and university students and staff may lend a hand with your evaluation. Consider departments such as Public Health, Geography, Urban Planning, Nutrition, Anthropology, Sociology, Landscape Architecture, etc.

Example of Completed Staff Inventory Form

Task	Task Description	Staff member, # of Staff Hours Currently Spent on Task	Staff Member/Estimated # of Additional Staff Hours that may be Spent
Project management	Develop project schedule & work plan Coordinate task assignments Review work Communicate results to internal & external audiences	Jose 20 hours annually	Claudia 20 hours annually
Evaluation methods	Develop/revise data collection protocols Design sampling procedures as needed Design plan for outcome analysis	No one (using forms and methods developed 5 years ago)	Jose and Bill will develop (20 hours)
Database management	Design & modify database Produce/update database documentation Review database to insure data entry is correct and current	Bill 5 hours monthly	Tina 5 hours monthly
Data collection	Select participants to complete forms Arrange for participants to complete forms	Tina and Otis 10 hours monthly	Tina and Otis 5 hours monthly
Data cleaning and coding	Correct errors in completed forms Develop & assign ID # and other code #'s Organize completed forms for data entry	Tina 5 hours monthly	Tina 0 to 10 hours monthly
Data entry	Enter data into database	Tina 3 hours monthly	Tina 0 to 3 hours monthly
Analysis	Perform statistical analysis of data collected	No one	Otis (with help from colleague/volunteer) 10 hours monthly
Report preparation	Write outcome assessment and interim reports appropriate for different audiences	Jose 20 hours annually	None needed

Staff Inventory Form

Task	Task Description	Staff member, # of Staff Hours Currently Spent on Task	Staff Member/Estimated # of Additional Staff Hours that may be Spent*
Project management	Develop project schedule & work plan Coordinate task assignments Review work Communicate results to internal & external audiences		
Evaluation methods	Develop/revise data collection protocols Design sampling procedures as needed Design plan for analysis		
Data base management	Design & modify database Produce/update database documentation Review database to insure data entry is correct and current		
Data collection	Select participants to complete forms Arrange for participants to complete forms		
Data cleaning and coding	Correct errors in completed forms Develop & assign ID # and other code #'s Organize completed forms for data entry		
Data entry	Enter data into database		
Analysis	Perform statistical analysis of data collected		
Report preparation	Write reports appropriate for different audiences		

Technology Inventory

For most organizations, a computer and the appropriate software will be essential ingredients in building an evaluation system. Programs will differ widely in the extent to which that and other technologies are used for storing and analyzing information.

The smallest programs (i.e., those with fewer than approximately twenty (20) participants and a data set with only a few variables or only qualitative data), may not require computerization for effective evaluation. However, the majority of programs will benefit dramatically from the use of computer technology. Some of the benefits your program will experience by using computers for evaluation are listed in the figure below:

Benefits of Using Technology in Evaluation

Although more staff time is required upfront to program databases or learn new software, the amount of time required to analyze data and create reports will decrease significantly. The staff time spent for set up will likely be saved in the first year of data analysis and reporting.

Data are more easily accessed. Reports to staff and funders can be easily generated at any time during the year, using up-to-date information for new participants. If a new question involving the data arises, the answer can often be determined through a few keystrokes. (If data are tabulated using a paper and pen, simple questions can take hours to answer).

Evaluation data can be easily linked to information already stored in participant databases (e.g., socio-demographic characteristics, hours of service provided, etc.) to help understand the kinds of participants for whom the program works best.

Computer-aided analysis typically is more accurate than that derived through hand calculations. Human error is common in hand tallies.

Statistical tests can be easily employed to report “significant” participant progress.

Your staff will attain additional software skills that may be useful for purposes other than evaluation.

Your program may already have the hardware and software to support your evaluation. In this case, there are two final thoughts to consider.

- 1) The main barrier to technological advancement may not be acquiring the hardware or software at all, but your lack of access to technical information and support. This is not an uncommon situation considering that you may be relying on a variety of donated parts and programs. Consider purchasing documentation, taking a class or contracting for technical support. Identify and try out your options before you really need them. As you develop your evaluation system, plan a strategy for the inevitable moment when you and your computer are at odds.

- 2) A final barrier to consider is the product of combining your staff and your technology. This combination could be as smooth as slipping your hand into a cashmere glove or as prickly as a bull ride in a cactus patch. Computer phobia is a common malady among all variety of workers. The best way to overcome this anxiety is to set aside time to play with the hardware and software using “pretend” files. Don’t worry about messing up the computer. Software will only lock up temporarily and not break anything permanently. Or staff may use a video game to learn the use of the mouse and pull-down menus. Ask a teen who is proficient on the computer to teach you.

Chapter 8 of this handbook gives more detailed information data analysis techniques.

Developing an Evaluation Plan

An important step in the process of evaluation is developing the evaluation plan. This is a written plan that links your goals, outcomes and performance standards, the evaluation strategies and analysis techniques to be used. It is generally recommended that the plan be developed before finalizing the evaluation design and tools to ensure that all of the key pieces of information are collected. It will help you to focus on the primary goals of the evaluation and resist the temptation to collect data less critical to the program.

A simple evaluation plan can be created by elaborating on your outcome worksheet (see Chapter 3, Worksheet #3). In that worksheet, you identified your goals, the effects that could be measured as outcomes, the “indicators” for those outcomes and the performance standards for those indicators. (Refer to Chapter 3 and the Glossary if these terms are unfamiliar to you.) In addition to these elements, an evaluation plan usually contains the sources of data that will be used to answer those questions. The evaluation plan should tie together with your program’s logic model. An example of such an evaluation plan is presented along with a worksheet on the following pages.

Sample Evaluation Plan: Neighborhood Garden Project

Program Goal	Indicators (Outputs and Outcomes)	Data Sources	Performance Standard
<i>To increase gardening skills for participants</i>	Increased knowledge of gardening practices Number of volunteers trained Total volunteer time Description of volunteer activities	Administrative records Volunteer activity logs Volunteer survey	10 volunteers total, including two core volunteers Total volunteer time meets need
<i>To increase satisfaction among Latino residents with produce offered at neighborhood farmer's market</i>	Satisfaction with food selection Number of youth participating in garden Number of hours youth participated in garden Increase in youth leadership skills Increase in youth connection to culture/background Increased consumption of vegetables by youth	Youth sign-in sheets Surveys of youth (post-program)	25% of participating youth will report an increase in leadership skills 60% of participating youth will report an increase in their connection to their culture/background 40% of participating youth will report an increase in the amount of vegetables they eat
<i>To increase civic responsibility</i>	Community service Description of original barriers to obtaining food prior to project Description of barriers to obtaining food after participation in the project Quantity of produce grown and sold	Garden logs of produce grown Market logs of produce sold Garden logs of produce taken home by youth and volunteers	500 pounds of produce will be grown in the garden in 2003 300 pounds of produce will be sold at the market in 2003 150 pounds of produce will be taken home for personal use by youth and volunteer growers in 2003
<i>To increase collaboration of food-related community organizations</i>	Increased collaboration Number of organizations participating in food network Number of meetings held	Meeting participation tracking forms	8 community-based organizations will attend monthly network meetings

Worksheet 5: Evaluation Plan

Program Goal	Indicators	Data Sources	Performance Indicators

Evaluation Plan Checklist

As you devise your evaluation plan, consider and check your work against the following list of questions. Scrutiny of your planned evaluation will help you to get started on the right track toward measuring your program's effects.

Questions to Answer as You Devise an Evaluation Plan:

- Does your plan cover all of your goals?
 - Is your evaluation focused?
 - Have you selected the evaluation data that really matter?
 - Does your plan reflect your program's philosophy?
 - Are the outputs and outcomes you identify realistic in view of the scope or design of your program?
 - Is the evaluation in line with your resources?
-

Does your plan cover all of your goals? You should be able to translate each of your goals into measurable terms. If not, the goal is probably too vague and needs to be reviewed, revised and clarified. For example, a goal stating “to help our participants improve” is too vague since participants can improve in many ways unrelated to the program. A better, more specific goal might be, “to help our growers increase in marketing skills” or “for our neighborhood residents to become more involved in community service.”

Is your evaluation focused? Do not include so many outcomes that your mission becomes diffused or your data collection efforts collapse under their own weight. For example, one fictitious youth program had as its goals to: increase gardening skills; improve self-esteem; improve family relationships; increase assertiveness, strengthen leadership skills and increase school performance. These goals might be appropriate if the organization plans to work intensively with participants for many years but may not be reasonable for a summer program.

Have you selected the evaluation data that really matter? Will your data tell you what you want to know about how effective your programs are? A good way to begin to answer this question is to identify what outcome indicators have been linked to long-term improvements in the quality of life of participants that are similar to yours. For example, it is well established that good nutrition practices are critical to establishing a healthy weight. Thus, changing diets is known to bring about longer-term, important outcomes.

To identify the most relevant data, it may be helpful to conduct a literature search of studies on the topic that your program addresses (e.g., community food security, increasing business skills, increased empowerment, nutritional status, etc.). Contacting similar programs around the country to learn about the indicators they monitor can also be useful and can provide outcome information (as well as information about funding

sources or innovative services) that you can compare to yours. If you collect follow-up data on your participants after they leave the program (and few programs do), you probably have a good understanding of the program outcomes that endure over time. Even anecdotal information can help steer you toward those data that really seem to last.

Does your plan reflect your program’s philosophy? Outcomes should reflect the service-related activities that you are engaged in on a daily basis. If you collect evaluation data that your program does not address, your results on those measures are not likely to be very good. (For example, if your program is specifically aimed at providing vocational training in the food service and distribution industry, you will want to use tools to measure vocational training specific to this industry and not use generic tools that fall outside this realm of training.)

Are the outputs and outcomes you identify realistic in view of the scope or design of your program? Performance indicators and standards must always be realistic. Otherwise, your confidence in and enthusiasm about your program will suffer and the real or presumed expectations of your funders will cause you to overestimate what you are likely to achieve. This is not to say that you should aim low in order to guarantee the appearance of success; instead, be realistic about your program’s capabilities and its limitations. Is the “dosage level” your program provides adequate to result in a measurable improvement in the outcomes you have identified? That is, as you examine your logic model, are the resources and activities devoted to promoting your outcomes sufficient to achieve the outcomes you have targeted? If your program is aimed at providing healthy food to a school cafeteria’s salad bar, it might be too aggressive to hold your program accountable for decreasing the number of obese kids in the school. A more appropriate outcome might be the number or percentage of kids who now eat more fruits and vegetables at lunchtime.

Is the evaluation in line with your resources? Data can be measured in more than one way. Some require far greater sophistication than others in collecting and interpreting the information. You will need to take into account how much time and expertise your staff has to collect the data you have identified. Many programs feel quantitative measures do not capture the richness of their programs. However, rigorous qualitative evaluation is likely to be more resource intensive and perhaps too costly in terms of money and staff time. For example, there are a number of ways of assessing improved leadership skills. Some assessment methods are relatively easy to administer such as asking participants if they feel they have become stronger leaders as a result of participating in the program. Another method where a staff member observes a participant’s behavior over time might be considered a stronger evaluation approach. However, this staff member may not have the expertise to apply this method effectively without extensive training, so it would be better to use something that is more straightforward though perhaps less rich in interpretative value. The training and

experience needed to make more sophisticated measurements may come at a later time or may be requested and paid for from a funder interested in the more rigorous outcome measures.

Over Promising: A Cautionary Tale about Promising Too Much

The central mission of the Growing My Greens youth garden is to provide a safe environment for a culturally diverse group of 5th to 8th grade students, a place where they can have fun and keep out of trouble. The program is at capacity and must turn kids away every year. Participants report that they feel accepted, safe and are having a good time with other youth and staff. There are no serious reports of trouble among youth while they are on the premises of the program. Parents report that staff are enthusiastic, talented, sensitive and courteous, and staff report feeling motivated and effective.

The data that tell this story come from a systematic collection and analysis of admission and exit surveys of youth, parents and staff and records about the kind and timing of services delivered and the characteristics of students receiving services. Staff have developed a set of rules about which person collects the data, how data are formatted, when and how data are analyzed, who gets the results and what decisions will be made from the report. Some self-told youth stories are tape recorded and transcribed as part of the annual report.

Despite the fine evaluation and positive outcomes, Carol Medina, the program director is getting some pressure from funders to show that youth participating in the program are more likely to graduate from high school and are more likely to find jobs after high school. Although the central mission of the program is to provide safe and fun activities, a new goal is appended to the mission along with a small amount of new funding to help ensure that participants meet these new goals. Youth mentors are now asked to talk with the youth about career goals. Outcomes for this part of the program (high school graduation rates and job attainment) when measured, are not what the funder or program staff hoped for.

As a general rule, programs must look for funding wherever they can find it, and funders demand proof of efficacy even if their funds buy only a portion of the service intensity needed to show measurable improvement. When funding does not match the effort required for mitigating long-term, more difficult problems, the director must be wary about promising positive behavioral outcomes.

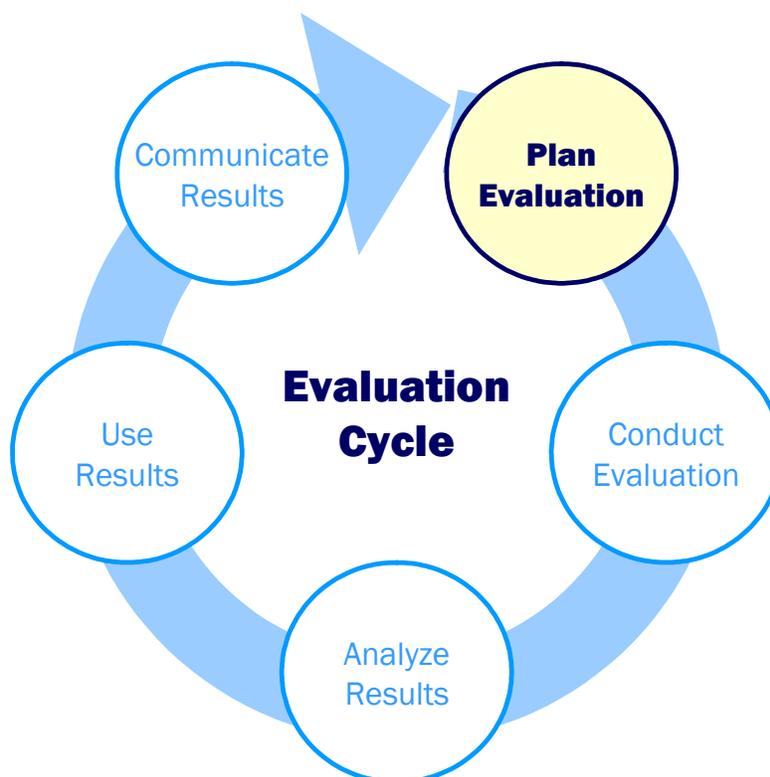
If new funds linked to new outcomes are too meager to expect important behavioral gains among participants – such as high school graduation or job attainment – Carol would do well to convince the funder that, while funding remains at this level, new outcomes should focus on participant attitudes or understanding which are more likely to change with limited resources devoted to the career counseling component of the program.

Carol Medina could also promise too much if she failed to consider what effort would really be required to determine if students did graduate or get jobs. It is one thing to identify a meaningful outcome; it may be another to find the resources to measure it reliably and regularly. It can be difficult to track participant whereabouts months or years later. It is important to consider how an outcome will be assessed before a promise is made to measure it. Further, it is important to describe to the funder the assessment process so that adequate resources are allocated to the outcome measurement itself.

Chapter 5.

Selecting Evaluation Strategies and Study Designs

- 🌸 Quantitative and Qualitative Research: A Sibling Rivalry
- 🌸 Typical Study Designs
- 🌸 Common Evaluation Methods
- 🌸 Selecting the Best Methods
- 🌸 Comparison of Evaluation Methods
- 🌸 Guidelines for Method Selection
- 🌸 Culturally Responsive Evaluation Methods



Chapter 5. Selecting Evaluation Strategies and Study Designs

As we discussed in the introductory chapter of this handbook, there are many ways to conduct evaluation. Although many consider evaluation a scientific endeavor, good evaluation practice is as much an art. It is a constant struggle to balance the search for “truth” with limited resources. In this chapter we discuss a number of methods that can be readily used to evaluate community-based service programs.

Quantitative and Qualitative Research: A Sibling Rivalry

A distinction commonly made by evaluators is the use of quantitative versus qualitative research strategies. Simply stated, quantitative evaluation uses numbers to describe and explain the topic of interest while qualitative evaluation relies more on words.⁴⁰ Qualitative evaluation is more subjective in nature and typically involves a smaller number of participants. The table on the following page provides more information on how researchers compare and contrast qualitative and quantitative research strategies.

Although many choose to classify evaluations as either qualitative or quantitative and keep explicit boundaries between the two, we believe that in good evaluation practices, the distinction becomes blurred. Both are examples of disciplined inquiry. Quantitative and qualitative strategies can be used in combination within a study and even within a single evaluation tool. Further, we argue that the strongest evaluation is based on a “mixed-method” design where quantitative data are used to provide breadth and qualitative data are used to provide depth and context.

All that said, information provided throughout most of this handbook is based on more quantitative forms of research: surveys, quantified observation and data tracking.⁴¹ We have chosen to focus on quantitative research methods for a number of reasons. First, quantitative data are felt to be stronger and more credible by a majority of funders – the data are more often associated with outcomes, accountability and objectivity. Second, purely qualitative research, when done well, is typically more expensive and requires significant staff training. Thus, it is less sustainable for a typical community-based service program.

Comparison of Quantitative and Qualitative Research ⁴²		
Characteristic	Quantitative	Qualitative
Basic Belief about the nature of reality	Relative constancy assumed. "Nature is orderly and follows specific laws. Occurrences have causes that can be discovered."	Dynamic. "Reality is what you think it is." Reality is not purely objective, and does not exist independent of the humans who interpret it.
Basic Perspectives – lenses through which the researcher interprets the world	Data is measurable: based on the natural scene worldview (empiricism, positivism)	Data is interpretive: based on the anthropological worldview (feminism, Marxism, humanism, race-based, multi-cultural)
Reasoning	Deductive (theory precedes research)	Inductive (theory emerges from research)
Goal	Results oriented: establishes relationships, demonstrates causation, makes confirmations	Process oriented: describes meaning, promotes discovery, exploratory
Role of Evaluator	Objective: evaluator is separate (outsider centered)	Subjective: evaluator is part of process (insider centered)
Basic Strategies	Designs: descriptive, correlational, quasi-experimental, experimental, causal-comparative Data collection methods: surveys, quantified observation, analysis of secondary datasets	Designs: phenomenology, grounded theory, ethnography Data collection methods: observations (case studies), in-depth interviews, focus groups, document studies, key informant interviews
Sampling	Random Larger number of cases	Purposeful, non-random Small number of cases
Analysis tools	Analysis tools: Statistical measures: means, medians, t-test, chi-square tests, ANOVA, MANOVA, non-parametric tests (Wilcoxins), correlations, Type I and Type II errors	Narrative analysis, discourse analysis, textual analysis, ethnography
Quality Assurance	Reliability: internal and external Validity: construct, content, face, predictive, discriminant, concurrent, convergent	Trustworthiness: credibility, confirmability, dependability, transferability
Depth	Surface glance	In-depth
Generalization	Strives for generalization (context free)	Strives for uniqueness (context dependent)
Reporting	Basic element of analysis is numbers	Basic element of analysis is words/ideas

Nonetheless, a strictly qualitative approach (e.g., conducting focus groups) has its use in the evaluation of a program. It is particularly useful when the research problem and the research setting are not well understood, when evaluation is in the earlier stage of

theory building rather than hypothesis testing, and when more in-depth analysis is needed. Other advantages of qualitative analysis include its ability to:

- Provide insight into the analysis by allowing participants to raise topics and issues not anticipated in the evaluation design
- Allow participants to express their feelings and opinions in their own words
- Provide anecdotal information that is powerful and persuasive
- Emphasize the importance of context, setting and participant frame of reference⁴³

In the remainder of this chapter we make little distinction between qualitative and quantitative methods. For more information on qualitative research methods please see one of the following resources:

- Marshall C and Rossman GB (1999). *Designing Qualitative Research*. Third Edition. Thousand Parks, CA: Sage Publications, Inc.
- Patton MQ. (1990). *Qualitative Evaluation and Research Methods*. Second Edition. Thousand Parks, CA: Sage Publications, Inc.

Qualitative handbooks on the Web include the following:

- User Friendly Handbook for Mixed Method Evaluations (1997). Eds. Frechting J. and Sharp L. Division of Research, Evaluation and Communication, National Science Foundation:
<http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/START.HTM#TOC>
- D. Ratcliff. *Qualitative Research Resources*. Department of Psychology, Biola University, CA: <http://don.ratcliff.net/qual/>

Typical Study Designs

When you are determining the best procedures for data collection in your program, you will want to consider adopting a strategy to produce the most meaningful data, using the least amount of resources. Base your study design on the natural intervals associated with the project activities you are offering. For example, the natural interval might be September through May if your activities take place during the nine-month school year, or April through October, if that is the growing season for your program.

The gold standard of all study designs is the **randomized-control trial** (RCT). This would involve randomly assigning participants into one of two groups: one group receives the services whose efficacy you wish to test (your project); the other does not. A comparison of the two groups' outcomes before and after program participation or just after program participation is made to determine effectiveness of services.

Very few service providers are in a situation where they feel they can or want to withhold services to a subset of their target population to demonstrate the efficacy of their program. There are, however, other types of study designs, which may not “prove”⁴⁴ that change in a population was caused by the program, but will help provide compelling evidence that the program did succeed in improving the quality of life of those participating in your activities. These evaluation designs are presented in the graphic on the following page [The (stair) Case for Program Impacts].

Following the (stair) Case graphic is a table describing each “step” in the evaluation design framework. As one moves along the staircase one finds evaluation designs that provide ever-stronger evidence that the program caused the improvement seen in participants or the target population.

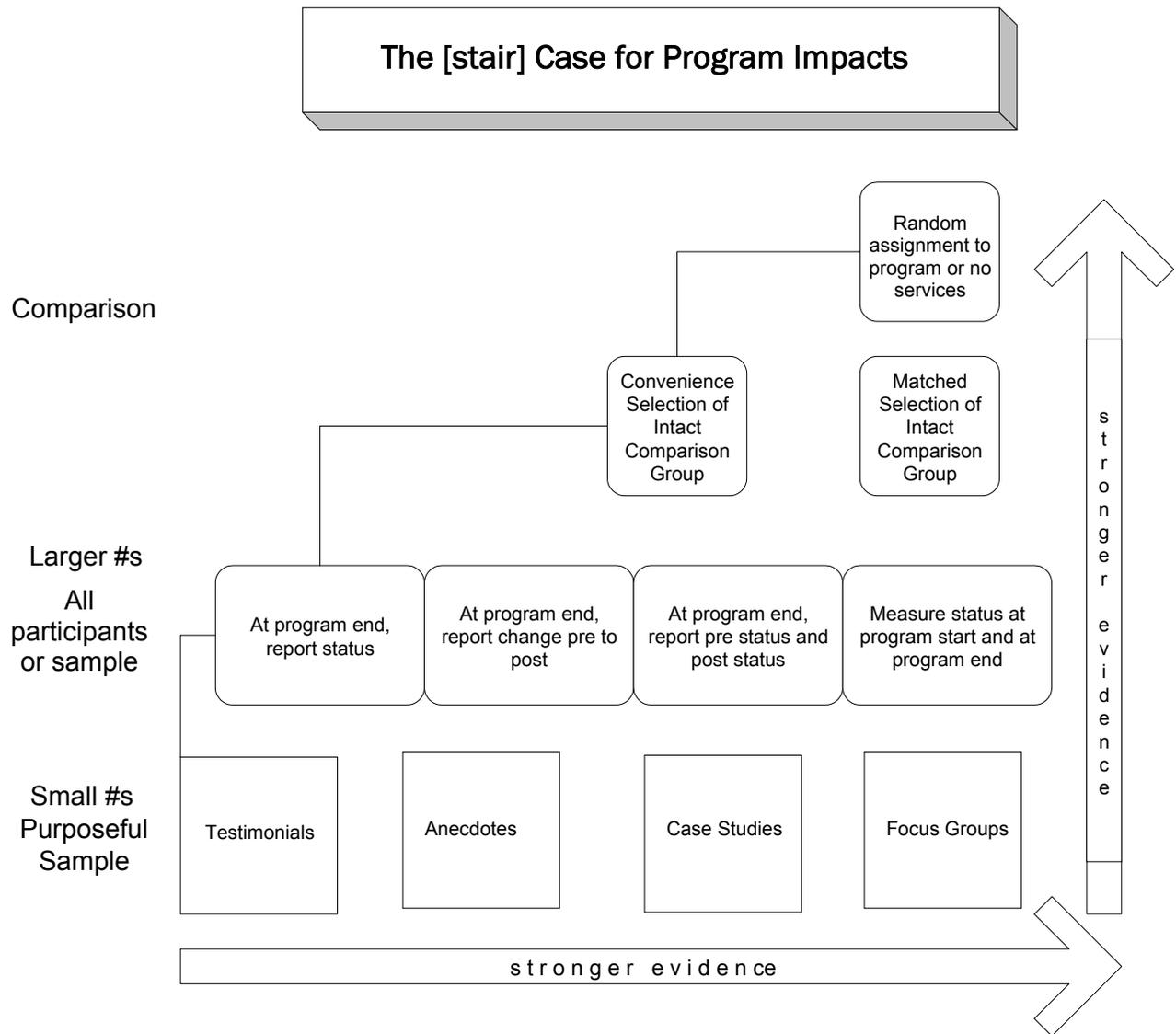
There are three major levels in the model. Tier one involves evaluation strategies such as testimonies, anecdotes and case studies. Although these evaluation modes can provide very colorful information about a program they are considered weak in their ability to demonstrate program outcomes as they are often limited to reports on a small number of participants who may not be representative of all participants. (Many program directors may choose to report on their most successful stories while omitting those of struggling participants).

Tier 2 involves more systematic evaluation strategies where a larger and more representative sample of participants is involved in the study. Outcomes are reported for all participants or a random sample of participants if a program serves large numbers. Evaluation designs tend to become more quantitative at this point because qualitative research is resource intensive with larger numbers of participants (e.g., survey data, score on skill inventories, etc.).

Tiers 3 and 4 uses the systematic methods of study designs in tier 2 (larger, more representative samples) and adds a comparison or control group. This comparison group can be created through random assignment (the strongest method), or through the selection of a “like” group or people not participating in a program (e.g., a school in the same district without a farm to cafeteria program). The addition of a comparison group makes an evaluation stronger because it provides information on how people are changing outside of a program due to national and regional trends, world events and maturation. Again, these study designs generally rely on more quantitative modes of data collection due to resource constraints.

As the evaluation design gains more rigor, the resources needed for the evaluation increase. Thus, we recommend a combination of anecdotes or case studies with outcome measures at program end or one of the designs for showing change from program start to finish (any of the options on second row of stairs). These require fewer resources than designs that rely on comparisons but provide reasonable evidence about

program impacts. When projects can afford to find comparison data in an evaluation, stronger evidence (for higher costs) can be found on the upper steps of the staircase.



© The Colorado Trust, 2002 by National Research Center, Inc.

Building the Case for Program Impacts			
Type of Design	Description	Example Measures	Strengths
Purposeful Sample			
Testimonial	Enthusiastic participant touts program	Self-conceived and self-described change.	Little research credibility but can make for compelling advertisement.
Anecdote(s)	Program staff describe success of a few participants	Observations of positive participant behaviors.	Little research credibility but can make for a compelling and readable presentation to funders; provides a “picture” to accompany the numbers.
Case Studies	Evaluator systematically reports experiences of a few participants	More or less in-depth observation of participant behavior; may include diary or events, photos, interviews, measures of participant attitude, knowledge or behavior.	Done systematically by an external evaluator can provide in-depth view of how program resources and activities appear to help.
Focus Groups	Evaluator systematically reports experiences of a smaller group of participants participating in a focused group discussion	Tell me about your decision to join {this project}? PROBE: Was it something you had heard or seen about the project? PROBE: Was anything different happening in your life?	Conducted with a small group of individuals, purposefully selected for one or more of their characteristics, in-depth questions can be pursued and group discussions can stimulate creative thinking
All Participant or Systematic Sample			
At program end, report status: Post-Only Measurement	Evaluator systematically measures all or a sample of participants	How anxious do you feel about your family’s finances? Very anxious Not at all anxious	By repeating post-only measures with the same participants over many years, a trend line of participant impacts is built and can be used for monitoring success.

Building the Case for Program Impacts			
Type of Design	Description	Example Measures	Strengths
At program end, report change pre to post: Self-report Improvement (retrospective)	Participants, staff or proxy describes change in participant pre to post that is attributable to program	To what extent did the program help you to feel more self-confident? very much.....very little How do you feel about your families' finances as a result of being in this program? much less anxious...much more anxious	Audiences want to know how participants improve so while this is the weakest way to assess improvement, it nevertheless can show what participant or observers believe the program impacts were. Well-constructed instruments keep this from being simply a testimonial.
At program end, report pre status and post status: Recalled pre-status compared to current status (post)	At program end, participant, staff or proxy rates participant status at start and then end	At end of program: Please rate the participants on each of the following characteristics considering his/her status at the start of the program and now: Sense of empowerment start:low.....high now:low.....high	Another way to measure change, this method is more credible than self-reported improvement because participant or observer must think back to how things stood just at program entry and to specify a point for that period on an outcome rating scale. Raters then indicate where they think they are now on that same scale.
Measure status at program start and at program end: Pre-Post Measures	At start of program, participant, staff or proxy rates participant status. Participant status is rated again at end of program.	At start of program: Please rate the participant on each of the following characteristics: Sense of empowerment low.....high At end of program: Please rate the participant on each of the following characteristics: Sense of empowerment low.....high	This is the best way to measure change, though not a perfect way to prove that the program caused the change. Here, pre-program status is measured at the start of the program and post status is measured at the end of the program so that memory lapses or the tendency to over or under estimate problems at program initiation will not muddy the change scores.

Building the Case for Program Impacts			
Type of Design	Description	Example Measures	Strengths
Comparisons to Other Programs			
Convenience intact comparison	The same participant measures are collected from participant in your program as well as from their peers not attending the program or attending a different type of program. The peer program is not necessarily selected for its similarity to your program, but because of its willingness to participate.	Pre-post or post only questions as shown above.	Comparisons offer the best antidote to the competing conditions that undermine claims that the program caused the participant to improve. If a participant in the program scores higher on measures of self-reliance than, say, a participant in another program, you have a better case that you were doing something that was effective.
Matched intact comparison	Agreement is made with a peer program to collect the same participant measures. Comparison programs are matched on the basis of participant and operational characteristics that are likely to affect participant outcomes, e.g., activities, participant ages, resiliency factors, demographics.	Pre-post or post only questions as shown above.	It is even more compelling if participants in your program do better on the measures of self-reliance than a matched group of participant when both sets of participants scored the same last year, they are the same age, their families have the same income, ethnicity and they live in the same neighborhood.

Building the Case for Program Impacts			
Type of Design	Description	Example Measures	Strengths
Random assignment of groups	A pool of eligible participants is divided randomly so that one group receives one set of services and the other group receives a different set of services or no services at all.	Pre-post or post only questions as shown above; completed in both groups and compared.	Even with matching, it can be claimed that there was something different (better) about your participants because they sought your program while their comparisons, looking demographically the same, might have lacked a certain motivation or home environment that helped your participants. By taking the choice of participation out of the hands of the participant, this counter claim disappears. Of course, it's best to consider random assignment when you have so many participants wanting your services that you couldn't accept them all anyhow. If you ever find yourself in such a situation or if you want to compare the effectiveness of two different kinds of services, it may be fairest to hold the equivalent of a lottery and then measure the participant who "won" the lottery and the ones who did not get into the program or the new service. The comparison of these two groups will give you the most solid evidence of your program's effects.

Common Evaluation Methods

There are as many ways to classify evaluation methods, as there are ways to collect outcome data. In this handbook we highlight some of the methods most useful in evaluating community-based service programs without significant costs. Evaluation methods can be broken into three categories: interviews/surveys, observation and document studies. The methods are described below.

Interviews, Focus Groups and Surveys

Interviews, focus groups and surveys involve sets of questions asked of program participants or others whose perspectives could inform on the success of the program (e.g., staff, volunteers, etc.). The use of interviews and self-report data as a research strategy begins with the assumption that the participants' perspectives are meaningful, knowable, and able to be made explicit.⁴⁵ Surveys can range from highly structured (complete determination of questions and response categories) to the less structured in-depth interview (little predetermination of the topics or response options). Highly structured surveys are generally conducted using a self-administered questionnaire (SAQ), the mail or sometimes the telephone.

Surveys using less structure are generally conducted in-person or on the telephone (more often referred to as interviews). An interview, rather than a paper and pencil survey, is selected when interpersonal contact is important and when opportunities for follow-up of interesting comments are desired.⁴⁶ Interviews and surveys can be given at the individual or group level. A focus group is an example of a group-level interview. A skilled interviewer⁴⁷ or moderator may read a survey aloud while several respondents use a paper and pencil survey to answer the questions.

The Important Role of the Interviewer

The interviewer plays an important role in the evaluation data collection process. Consider the following points when selecting an interviewer:

- Interviewers should be trained individuals who are good listeners, sensitive and empathic.
- They should be able to establish a non-threatening environment in which participants feel comfortable to be candid.
- Interviewers should be unbiased listeners.
- It is important to note that personal characteristics of the interviewer (e.g., age, sex, race/ethnicity and appearance) may influence how acceptable they are to the individuals being interviewed.

- Interviewers must speak and understand the language of the participants involved in the research (see Chapter 7 for more information on cultural responsive data collection).

More recently, technology has become important in the administration of interviews and surveys. Audio computer-assisted self-interviewing (ACASI) is a methodology that allows the respondent to listen to the interview with a headset/walkman and answer the questions on a computer or simply read the questions off the computer screen. This interviewing mode works particularly well with sensitive topics as it provides a greater sense of anonymity. Further, this method is attractive to youth because it allows for varying response times and uses modes they are familiar with in a social rather than academic context.⁴⁸

Using Program Participants to Collect Evaluation Data

The trend in empowerment evaluation and Asset-Based Community Development⁴⁹ calls for interviewers to be pulled from within communities and programs rather than imposed from outside. Through this process, participants skills are strengthened, their dispositions changed and their civic capacities increased. This form of data collection process also works to build and strengthen relationships between participants and empowers participants as program resources not just recipients. Although this trend in using internal interviewers may be important programmatically, it should be noted that it is considered weaker methodologically as it is often more subjective and less consistently collected.

Focus Groups: A Structured Group Interview

A focus group study is a “carefully planned series of discussions designed to obtain perceptions on a defined area of interest in a permissive, nonthreatening environment.”⁵⁰ Focus groups are typically a gathering of six to 12 people who share some characteristics relevant to the evaluation. An interviewer or “moderator” leads the group in one to two hour discussions. The hallmark of focus groups is the explicit use of group interaction to generate data and insights that are more likely to emerge when respondents communicate. The technique allows observation of group dynamics, discussion and firsthand insights into the respondents’ behaviors, attitudes and language.^{51,52}

The discussion guide used with the groups tends to start with general questions designed to enhance group interaction and foster conversation; and then leads into more personal, potentially controversial or “hot” topics that drive the main discussion. “Probes” or probing questions are also used to elicit responses to the main questions.

A limited number of questions are used so that discussants have the time to share their thoughts and ideas without feeling rushed through the dialogue. Discussion questions are limited in their complexity and length to help with ease in understanding.

Primarily open-ended questions are asked rather than “yes” or “no” questions to learn underlying thoughts and feelings; anticipated or past actions or influences in their lives.

Observation

Observational strategies are methods by which an individual or individuals gather firsthand data on programs, processes, or behaviors. Observation protocols (i.e., carefully developed sets of steps, reporting or rating guidelines and instruments) are often used to assure that all observers are gathering the same types of data and applying similar criteria to collection. The protocol can take a variety of forms, ranging from a narrative description of events seen, to a checklist or a rating scale of specific behaviors and activities.⁵³ Observers can also use field notes as well as technological tools such as tape recorders, video cameras or laptop computers to help record events for later analysis. (See Appendix III and IV for a copy of an evaluation rubric template and two illustrations.)

Document Studies

For some topics, existing documents, records or datasets can be used to determine the attainment of an outcome. Documents are sometimes categorized into two separate categories: public records and personal documents.⁵⁴

Public records are materials created and kept for the purpose of “attesting to an event or providing an accounting”⁵⁵ and are often tracked outside of the program. Examples of public records that are useful in community food projects are Census data, Federal food assistance program participation information (number, locations, participation rates), The Centers for Disease Control’s Behavioral and Risk Factor Surveillance System and Youth Risk Behavior System, the HHS and USDA’s Food Security Supplement to the Current Population Survey.

Personal documents are first-person accounts of events and experiences. These “documents of life” include diaries, portfolios, photographs, artwork, schedules, scrapbooks, poetry, letters to the newspaper, etc.⁵⁶ Analysis of these types of data is often more qualitative in nature or involves some form of evaluation “rubric.”⁵⁷

Selecting the Best Methods

Different types of evaluation methods will be appropriate for different types of outcomes. As you decide what methods will be appropriate, keep in mind not only the outcomes you desire to measure, but also the credibility of the collected data to your intended audience. A testimonial may be a powerful and effective demonstration of the influence of your program, but some funders may doubt the “generalizability” of such stories; that is, whether or not this impact was felt equally by all participants. A testimonial in combination with survey data may help to complete the picture, and provide a firmer foundation of evidence. Likewise, survey data alone may be

insufficient to show the range of impact that your program has on the lives of its participants.

In the preceding text, we introduced nine data collection methods in three categories: surveys, observations and document studies. These nine methods are:

- Hand-distributed self-administered questionnaires
- Mailed self-administered questionnaires
- Phone interviews
- In-person interviews
- Group interviews (highly structured)
- Focus groups (less structured group interview)
- Audio computer-assisted self-interviewing (ACASI)
- Observations
- Document studies

(These methods are compared on page 82 and include some guidelines for selecting the appropriate method for your evaluation needs.)

As mentioned above, you may choose to join any of these nine evaluation methods with other informal techniques such as using testimonials. Although this handbook concentrates on the nine methods listed above, here are some alternatives to consider.

- **Testimonials:** individual statements by people indicating personal responses and reactions.
- **Anecdotes:** individual statements by staff of participants indicating positive personal behaviors
- **Photographs, slides and videos:** use of photography to capture visual images.
- **Diaries, journals:** recording of events over time revealing the personal perspective of the writer/recorder.
- **Logs:** recording of chronological entries, which are usually brief and factual.
- **Action cards:** use of index cards on which participants record what they did – the “action” – and when they reach their goal; primarily used in self-assessment.
- **Simulations:** use of models or mock-ups to solicit perceptions and reactions.
- **Problem stories:** narrative accounts of past, present or future situations as a means of identifying perceptions using fictional characters to externalize the problem situation.
- **Creative expression:** use of art forms to represent people’s ideas and feelings as through stories, drama, dance, music, and art.

- **Unobtrusive measures:** the gathering of information without the knowledge of the people in the setting such as the wear and tear on a “planted” mat in front of a display.

Comparison of Evaluation Methods

For captive audiences, such as the participants of a community food project activity, researchers have commonly relied on hand-distributed self-administered questionnaires, personal interviews, structured group interviews and focus groups. More recently, technology has played a major role in survey administration whereby participants are administered surveys on the computer or with the use of an audio recorder or portable media player.⁵⁸

Participants can also be surveyed using the more traditional methods of mail and phone, although these options are often reserved for non-captive populations (these modes, however, are most likely the best options for surveys of parents). In the next table, “low” indicates that a data collection method is not good at achieving a characteristic, whereas “high” indicates an optimal method.

Evaluation Method Comparison									
	Hand Distributed SAQs	Mail SAQs	Personal Interview	Phone Interview	Group Interview (Structured)	Focus Groups	ACASI	Observation	Document Studies
<i>Speed of Administration</i>	Fast	Slow	Slow	Moderate	Fast	Fast	Moderate	Slow	Varies
<i>Providing a High Response Rate</i>	High	Low	High	Low	High	Moderate	Moderate	High	High
<i>Obtaining Candid Responses</i>	High	High	Low	Low	Moderate	Low	High	High	NA
<i>Eliminating Interviewer Bias</i>	High	High	Low	Low	Moderate	Low	High	Low	High
<i>Getting at In-Depth Topics</i>	Low	Low	High	Moderate	Moderate	High	Moderate	Low	Low
<i>Permitting the Use of Visual Aids</i>	High	High	High	Low	High	High	Moderate	N/A	NA
<i>Enforcing Question Order</i>	Low	Low	High	High	High	High	Moderate	N/A	NA
<i>Reducing Cultural Barriers</i>	Low	Low	High	Moderate	Low	Moderate	Low	High	NA
<i>Accessing Captive Respondents</i>	High	Low	High	Low	High	High	Moderate	High	NA
<i>Assessing Non-captive Respondents</i>	Low	High	Low	High	Low	Low	Low	Low	High
<i>Expense of Hard Costs</i>	Low	Moderate	High	High	Low	Moderate	High	Low	Varies
<i>Staff Time Needed to Collect Data</i>	Low	Low	High	High	Low	Low	Moderate	High	Varies
<i>Staff Training/ Evaluation Skills</i>	Low	Low	High	High	Moderate	High	Less	High	Moderate
<i>Eliciting Interest from Youth</i>	Low	Moderate	Low	Moderate	Low	Moderate	High	N/A	NA
<i>Allowing for Mixed Literacy Levels & Mixed Cultures⁵⁹</i>	Low	Low	High	High	Moderate	Moderate	High	High	High
<i>Burden on Participants</i>	Moderate	Moderate	High	Moderate	Moderate	High	Moderate	High	Low ⁶⁰ / High ⁶¹

Guidelines for Method Selection

You may administer instruments by handing them out and collecting them, or completing personal interviews when participants are easily accessible to your program office. If you do not have much interpersonal contact with the people you would like to survey a phone or mail survey method might be a better alternative. Guidelines to determine the best method are presented in the following table.

Guidelines for Selecting a Data Collection Method		
Survey Mode	Recommended Circumstances	Not great when ...
Hand-distributed Self-Administered Questionnaire	<ul style="list-style-type: none"> • Captive audience • Data need to be collected in short time frame • Minimal staff time is available for data collection • Outcomes can be measured in concise manner (survey must be fairly short) and at right literacy level • Survey questions are self-explanatory and do not need interviewer/interviewee interaction • Questions may be asked on sensitive topics • Question order is not as important 	<ul style="list-style-type: none"> • Instruments closely resemble a test • Respondents are already over-surveyed or over-“tested” • Respondents need varying lengths of time to complete the survey
Mailed Self-Administered Questionnaire	<ul style="list-style-type: none"> • Respondents are not captive • The respondent has little personal contact with staff • Minimal staff time is available for data collection • Respondent has a high concern for anonymity • Respondents prefer to answer the questions in a setting comfortable and familiar to them 	<ul style="list-style-type: none"> • Questions are longer and require more in-depth responses
Personal Interview	<ul style="list-style-type: none"> • Questions are more complex and require more in-depth responses • Questions may best be answered with interviewer-interviewee interaction • Respondents vary significantly in terms of literacy levels • A person perceived as neutral has time to schedule and administer the interviews 	<ul style="list-style-type: none"> • Questions are asked on sensitive topics • Minimal staff time is available for data collection • Staff delivering service also ask questions of participants
Phone Survey	<ul style="list-style-type: none"> • Respondent is not captive • The respondent has little personal contact with staff • Outcomes can be measured in a concise manner (survey must be fairly short) 	<ul style="list-style-type: none"> • Questions are asked on sensitive topics • Minimal staff time is available for data collection • A person perceived as neutral has time to administer the interviews

Guidelines for Selecting a Data Collection Method		
Survey Mode	Recommended Circumstances	Not great when ...
Group Interview using Written Surveys (Structured)	<ul style="list-style-type: none"> • Captive audience • Data need to be collected in short time frame • Minimal staff time is available for data collection • Outcomes can be measured in a concise manner (survey must be fairly short) • Questions are asked on sensitive topics 	<ul style="list-style-type: none"> • Respondents need varying lengths of time to complete the survey • Respondents are already over-surveyed or over-“tested” • Instruments closely resemble a test • There is no central location where respondents can convene
Focus Groups	<ul style="list-style-type: none"> • Situations where group interaction is important • In-depth information is needed • There is limited staff time to collect information • Staff is well trained to facilitate the focus group 	<ul style="list-style-type: none"> • Questions are asked on sensitive topics • Participants are at varying levels of power • A broad sample of opinions is needed • Participants can influence other areas of each other’s lives
Audio Computer Assisted Self Interviewing (ACASI)	<ul style="list-style-type: none"> • Captive audience • Minimal staff time is available for data collection • Outcomes can be measured in a concise manner (survey must be fairly short) • Questions are asked on sensitive topics • Respondents need varying lengths of time to complete the survey 	<ul style="list-style-type: none"> • Program does not have technology necessary to administer the survey • Data need to be collected in a shorter time frame
Observations	<ul style="list-style-type: none"> • Captive audience • Participants are unaware, unwilling or unable to discuss a particular topic • Understanding the context of events is as important as the event itself • Outcomes can be easily observed • Trained observers are available to record events 	<ul style="list-style-type: none"> • Outcomes are difficult to observe or observation would be inappropriate /intrusive • Program has little access to trained observers • You can rely on participant self-report • Staff make judgments about their own participants
Document Studies	<ul style="list-style-type: none"> • The information exists to measure the outcome • The information is reliable and accurate 	<ul style="list-style-type: none"> • Data or documents do not exist

Culturally Responsive Evaluation Methods

The evaluation methods selected for your study should be culturally responsive to the individuals being evaluated. Some methods may appear foreign, obtrusive or inappropriate for some cultures. Thus, engaging cultural stakeholders in the process of

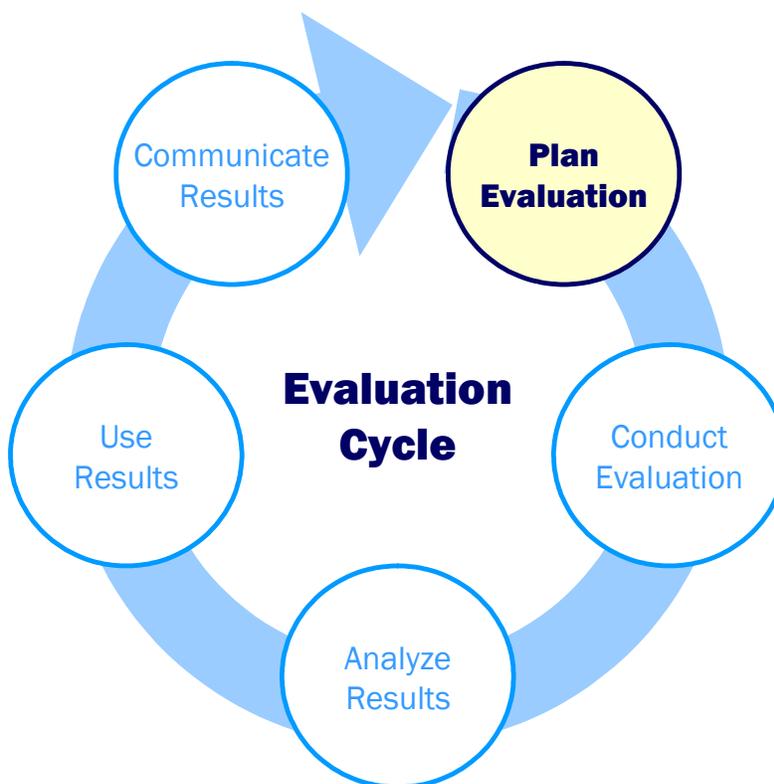
method selection is recommended. You may need to think outside the box of standard evaluation methods or alter the way in which a method is administered to be respectful of those being evaluated. For example, a survey may be administered door to door rather than mailed to rural families. Or, focus groups may be transformed into “learning circles” co-facilitated by the evaluator and a trusted member of the community.

When developing and implementing evaluation methods for diverse groups, reflect on and incorporate the ethnic and cultural characteristics, experiences, norms and values of the populations you serve. Consider different communication patterns and styles of interaction across racial/ethnic and other culturally diverse groups. How can you honor these in your evaluation? One expert notes, “Some groups use indirect means of communication more frequently while others emphasize cooperation over competition.”⁶² An evaluator seasoned in multicultural evaluation explains, “Evaluation should be built upon the values of the community.” As an example, tribal programs are often centered on strengthening their communities. Evaluation can become a means of telling a story, “one more of many stories that chronicle this people in their place.”⁶³ Evaluators become vehicles for telling these stories in a methodologically sound way.

Chapter 6.

Designing and Customizing Evaluation Tools

- 🌸 Searching for Tools
- 🌸 Evaluating Existing Tools
- 🌸 Developing Your Own Tools
- 🌸 Designing Culturally Responsive Evaluation Tools
- 🌸 Developing Effective Tools
- 🌸 Pilot Testing Your Evaluation Tools



Chapter 6. Designing and Customizing Evaluation Tools

After you have determined the types of data you would like to collect and the evaluation strategies you will use, it is time to design your evaluation tools. The tools selected to measure your outcomes are critical to the success of your evaluation program. The CFSC Evaluation Program has worked with NRC to develop the companion to this handbook, *Community Food Project Evaluation Toolkit*. This toolkit includes over 40 template surveys and evaluation procedures for various elements of community food security initiatives, including general evaluation tools, project level tools (for farmers' markets, community gardens, community supported agriculture, and farm to school projects) as well as system-level tools for evaluating food coalitions, networks and councils. Several CFP grantees were involved in editing and pilot testing these tools to ensure their relevancy. You can order a copy of the Toolkit on the CFSC website at www.foodsecurity.org.

Searching for Tools

Before creating your own tools from scratch, you may want to identify assessment materials developed by other organizations with similar data needs. The organizations may be providers of like activities in the state, region or elsewhere in the country. You can even share tools with other community food projects. There are many reasons to look first for tools developed by others:

- It saves time – others already have struggled through the difficulty of creating the tool.
- The tool has been pilot-tested – others have used it and hopefully have ironed out any major problems with it.
- The tool may be validated –quantitative tools can be tested through statistical methods to demonstrate that they are measuring what they purport to measure and that they are consistent or reliable.
- The data from other programs can be used as control or reference data for comparison.
- Other types of important data can be identified. When reviewing another program's materials, you may discover other useful pieces of information.

Evaluation materials that have already been developed can be identified by:

- Conducting a literature search at the library or on the Internet under the topic (e.g., household food security, nutrition status, leadership skills, empowerment, etc.)
- Contacting similar programs in the region or the nation
- Calling experts in the field and interviewing them about any information sources of which they are aware

- Contacting publishers for catalogs of tests that can be purchased
- Searching the web

Many of the tools you will find may be more quantitative in nature, particularly if you are looking to measure social and psychological characteristics of participants (e.g., conflict resolution skills, empowerment, etc.). If the tool is published, you can usually obtain information about its validity and reliability, as well as recommended procedures for administration, scoring, interpretation of results and norms from the company that markets it or from its author. Public domain instruments are not copyrighted or sold and can be copied and used free of charge. In the best-case scenario, you may find multiple tools measuring the same concept. You can then review all those instruments and determine which approaches the topic in the way that fits best with your program goals. The effort you put into selecting measurement instruments will pay off in the long run in terms of the quality and utility of information you collect and monitor. A list of guidelines for good instrument construction is provided in the next section, which can also help you determine the best instrument.

If you like portions of an instrument, but not the entire tool, use those portions and combine them with questions from other instruments and your own questions. Single questions, modified to fit your target population are unlikely to violate copyright laws.

Copyright Laws ©

The lawful use of instruments varies greatly. Some can be copied from library books while others must be bought directly from the author. If an instrument is clearly copyrighted or produced by a publishing company, DO NOT use it without purchasing it or contacting the author for permission. It can be difficult to tell whether or not it is legal to copy and use an instrument or its single questions. Contacting the author directly is a good practice, if possible. The author may be very happy to have the instrument used, but would like to insure its proper use by asking you to acquire some accompanying literature or he or she may want you to share your results.

Web-based searches for tools may result in a larger number of useful tools for community-based organizations. Although these tools are less likely to be validated or used in multiple populations, they might work well for your CFP. Focus group discussion guides and interview scripts are often included as appendices of full evaluation reports displayed on the Web by government and other non-profit organizations. When selecting an evaluation tool, pay careful attention to the credentials and experience of its authors.

Evaluating Existing Tools

Quantitative Tools

Tests can be performed on evaluation tools to help determine if they are reliably measuring what is intended. It is beyond the scope of this handbook to present the methodologies and statistics used in these tests,⁶⁴ but it is important that you understand the concepts so that you can evaluate the properties of a tool before using it. In quantitative research, there are two main characteristics of instruments that are commonly tested and reported: validity and reliability. These are called the psychometric properties of the instrument.

Put simply, validity means that an instrument measures what it is supposed to measure. For example, a bathroom scale that repeatedly tells you your weight, rather than whether you gained weight over Thanksgiving, isn't a valid bathroom scale (although it may be considered reliable if it consistently informs you of your correct weight).

Reliability, on the other hand, means that a measure gives you consistent answers – the bathroom scale should not tell you that on Monday you weighed 100 pounds, on Tuesday you weighed 200 pounds and on Wednesday you weighed three pounds. Reliable measures give similar answers when measuring similar circumstances. The following table applies these concepts to community food project evaluation.

Validity and Reliability			
Instrument Characteristic	Interpretation	Example for Fictitious Instrument Measuring Leadership Skills	Subcategories of Characteristic
Validity	Describes how well the instrument measures what it was intended to measure	20 participants take an instrument that is designed to measure improved leadership skills. A program staff member spends two hours with each participant and determines which have a high # of skills or a low # of skills. The staff member's findings are compared to the instrument. If the test has a high correspondence with the staff member's determination, it is considered "valid."	Face validity, construct validity, discriminant validity, concurrent validity, predictive validity, convergent validity
Reliability	Describes how consistent the instrument is with repeated measurements over time or items	20 participants are asked to take the "leadership quiz," twice, one week apart, with no CFP activities provided between the retesting. If the retest corresponds highly with the first test, it is considered "reliable."	test-retest reliability, inter-rater reliability, intra-rater reliability, split-half reliability, internal consistency, specificity

Qualitative Tools

The quality of qualitative tools is more difficult to judge because it is more difficult to apply statistics (qualitative tools do not typically produce numerical information). The questions or protocols themselves are not as often critiqued as is the analysis and interpretations of the evaluators. The standards that exist for qualitative research are related to the “trustworthiness” of the evaluation. A qualitative study is considered trustworthy if it gives the reader confidence in its findings. Cuba and Lincoln (1989) set forth four criteria that comprise trustworthiness: credibility, dependability, transferability and confirmability. These criteria appear in the table below.⁶⁵

Standards Used to Judge the Quality of Qualitative Research

Credibility: is achieved when those who read the research study perceive the situation described by the research study as related to his or her own experience (similar to internal validity in quantitative research).

Dependability: is achieved when the researcher provides a sufficiently clear account of the research process to allow others to follow his or her thinking and conclusions (similar to reliability in quantitative research).

Transferability: is achieved when a study’s findings “fit” contexts beyond the immediate study situation (similar to external validity in quantitative research).

Confirmability: is achieved if the results are confirmed or corroborated by others (similar to objectivity in quantitative research).

Developing Your Own Tools

Beyond these psychometric tests and criteria, it is important to remember that every evaluation is unique. Just because a tool is considered to be the “gold standard” in the field or has been used in another large community food project, it may not work for your project. It is important to consider the age, cultures and contexts where the tools have been used in the past. Also, the use of a pilot-test (discussed later in the text) will help ensure that evaluation materials are a good fit.

Many non-profit organizations can remember working in an environment where data were collected and never used because the information was not what the organization or staff really needed for program planning or the data did not provide the picture of what was occurring.

Designing Culturally Responsive Evaluation Tools

In addition to the nuts-and-bolts of scientific instrument design, program staff need to ensure that the assessment tools they choose or create are appropriate for the intended population. Evaluation tools are generally developed within a particular context, for a specific population. Many instruments have not been used or tested in other settings. A tool that has been developed and is culturally appropriate for one group (e.g., for use

with White, middle-class college freshmen) may not be appropriate for use with people of different cultures, different socio-economic status or within different urban or rural settings. One author refers to this as the “Three Musketeers” problem. In the development of measurement tools, “all for one and one for all” may not necessarily apply. “If different people have different cultural and racial experiences and present their abilities differently, it is unlikely that a single measure could be developed that would work equally well for all.”⁶⁶ The American Counseling Association advises: be “cautious in using assessment techniques, making evaluations and interpreting the performance of populations not represented in the norm group on which an instrument was standardized.”⁶⁷ They also state that one should “recognize the effects of age, color, culture, disability, ethnic group, gender, race, religion, sexual orientation and socioeconomic status on test administration and interpretation and place test results in proper perspective with other relevant factors.”⁶⁸ Further, programs with participants from several different racial and cultural groups may need to consider the implications of using only one measurement tool. Cross-cultural assessment may require more intensive review and adaptation of the instruments you choose.

Programs should consider the way participants and some communities may feel about or relate to the process of evaluation: test taking, having sometimes personal questions asked of them, etc. These feelings may be intricately related to their culture: influenced by cultural values, beliefs and practices. Sociopolitical factors of poverty, racism, immigration and culture all contribute to differences in how one may feel about evaluation.⁶⁹

A couple methods you might use to ensure that you are using culturally appropriate measurement tools follow:

Make Decisions Collectively: Evaluations are more successful when they “incorporate a broad spectrum of perspectives and experiences.”⁷⁰ Hence, programs may consider selecting diverse individuals and stakeholders to participate on an evaluation advisory team to ensure that, in the earliest conceptual stages of instrument development, cultural factors are considered. Your team may include program and evaluation staff, as well as individuals from stakeholders groups such as participants, volunteers, board members and other members of your community. Frequent exchange of information between those developing your data collection instruments and those with a particular interest and expertise in multicultural issues may be helpful to ensure that instrumentation is culturally and socio demographically appropriate.⁷¹ An advisory team may help answer the following two questions:

- Does the instrument design and implementation strategy reflect “sound practice in responding to diverse populations in multiple and meaningful ways?”⁷²

- Are you able to determine the extent to which the differences among the various participants (e.g., cultural and linguistic diversity) are “acknowledged and effectively addressed to ensure higher levels of achievement of all?”⁷³
- Do the instrument or related evaluation materials contain language considered prejudice or cultural assumptive?⁷⁴

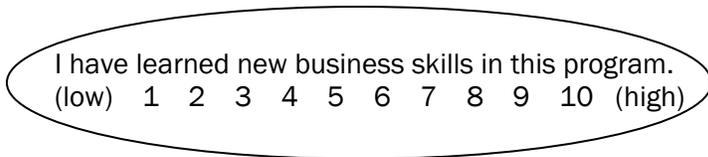
Remaining unbiased does not require an evaluator to make culture-free considerations. However, keep in mind that membership in a ‘particular group’ does not automatically render an evaluator, or an advisory team member, free of bias or make one an expert in the educational issues, needs, etc. of that group. In fact it may limit one’s perspective. Therefore, by utilizing a diverse advisory group to critique the evaluation design, give advice and review findings, a program may best achieve a culturally competent evaluation.⁷⁵

Pilot-test Your Tools: As with any evaluation tool, always pilot-test the tool on a few participants from your program. Choose those who reflect the diversity of culture, age and literacy variation in your program (pilot testing is covered in greater detail later in the text on page 101). A dry run with these participants will help you determine how to better adapt the tools.

Tool Adaptation: Items must be deemed bias free and conceptually and linguistically equivalent for their intended audiences. The following table describes these methods that can be used to adapt assessment tools to better fit the population of interest.

Tool Adaptation ⁷⁶	
Translation	
Forward	The original tool in the source language is translated into your audience’s primary language and then bilingual researchers or assistants are asked to compare the original version with the adapted version.
Backward	The tool is translated into another language and then re-translated back to the source language. This practice may be repeated several times, then comparing the final back-translated version to the original version.
Equivalence	
Functional	Do not assume that behaviors play the same role or function across cultures. Verify the role or function behavior plays in different cultures.
Conceptual	Some behaviors and concepts may have different meanings across cultures. Consider the similarity in meaning attached to behavior or concepts referred to in the tool.
Metric	This refers to the psychometric properties used in the tool. Make sure the scales measure the same constructs across different cultures.
Linguistic	The actual translation process.
Tool Bias	
	Contents of the tool may be more familiar to one group than to another or have differential predictive validity across groups. Make certain that the tools you choose do not systematically discriminate against your culturally diverse audiences.

In addition to the above suggestions, Marin and Marin⁷⁷ recommend that careful consideration be given to the use of Likert-type response scales when developing tools for culturally diverse audiences. Participants of some cultures may not be comfortable making subtle distinctions used in Likert scales. See example of a Likert scale below.



Several appendices provide additional information about creating or modifying tools to be age appropriate. See Appendix V which offers hints for developing instruments for youth and children, Appendix VI with tips for conducting focus groups with teens and Appendix VII which presents information for developing instruments for older adults.

Developing Effective Tools

Because your program fulfills a unique role in the community, you will want to demonstrate how your program contributes to the community and is uniquely effective. These outcomes will be most specific to your mission and may provide useful information to aid internal program improvement. Further, there may be few tools developed by others that will work well in your program. A framework for developing and identifying data collection tools is provided in the following section.

Tool Design

A number of issues need to be considered when drafting an evaluation tool. Some simple guidelines are listed below:

- **Keep it short and to the point.** The simpler the activity, the more likely it will be understood and completed. To enhance simplicity, make the questions specific, short, logical, non-threatening and at literacy levels appropriate for the population you serve.
- **Guarantee anonymity or confidentiality, if possible.** A program may choose to administer an anonymous or confidential tool. Anonymity means that no identifying information will be collected from respondents, not even their phone numbers. Confidentiality indicates that programs are enforcing clear rules that prohibit access to any information that would identify a particular respondent.
- **Begin the evaluation activity with more general, less threatening questions.** Place the most sensitive or personal questions closer to the end. This includes questions on socio-demographic characteristics like age, household income, etc.
- **If using a self-administered questionnaire – make it friendly and attractive.** Surveys that are well laid out and logical will ensure higher response rates.

Use caution when adding unnecessary mood boosters like cartoon graphics that might bias results; but here are some suggestions of how to make the instrument more appealing:

- Using appropriate size fonts
- Using more creative, but easy to read fonts
- Making sure the order of your questions is logical and easy to follow
- Keeping the wording simple, using common language appropriate to your audience
- Leaving enough space to answer questions appropriately
- Avoiding instruments that look like a “test”
- Printing surveys on pleasingly colored paper
- Using graphics
- Allowing enough “white space” to avoid seeming overcrowded with items

Writing Your Own Questions

Wording the questions is often the most difficult part of constructing an evaluation tool. There are a number of guidelines that, if followed, will help produce more “scientifically valid” questions – questions that are unbiased and clear. The table on the following page addresses the concepts of clarity, fairness and neutrality when writing quantitative questions for any audience.

Guidelines for Developing Questions ⁷⁸			
What To Avoid	Explanation	How NOT To Say It	How To Say It
<i>Vague wording</i>	Avoid words or expressions in questions and response items that can be interpreted in more than one way.	Example: How do feel about the volunteers at program x?	Alternative: Volunteers at program x care about me.
<i>Double-barreled questions</i>	Avoid questions where respondents are asked to provide an answer to more than one topic simultaneously.	Example: Staff at program x provide my child a lot of support and a sense of safety.	Alternative: Staff at program x provide a lot of support for my child. And I feel my child is safe while participating in activities in program x.
<i>False assumptions</i>	Do not assume characteristics about the respondent that may not be true.	Example: How many times did you eat fast food in the past week? ___ 1 to 5 times ___ 6 to 15 times ___ 16 or more times	An alternative: Have you eaten fast food in the past week? ___ yes ___ no If yes, how many times did you eat fast food in the past week? ___ 1 to 5 times ___ 6 to 15 times ___ 16 or more times OR How many times did you eat fast food in the past week? ___ 0 times (never) ___ 1 to 5 times ___ 6 to 15 times ___ 16 or more times

Guidelines for Developing Questions ⁷⁸			
What To Avoid	Explanation	How NOT To Say It	How To Say It
<i>Over-lapping categories</i>	Question responses should not overlap.	Example: How many servings of fruits and vegetables did you eat yesterday? ___0 servings ___1 to 3 servings ___3 to 5 servings ___5 or more servings	Alternative: How many servings of fruits and vegetables did you eat yesterday? ___0 servings ___1 to 3 servings ___4 to 5 servings ___6 or more servings
<i>Option asymmetry</i>	Question response categories, in most cases, should have an equal number of positive and negative options.	Example: I am good at making decisions ___Strongly agree ___Agree ___Disagree	Alternative: I am good at making decisions ___Strongly agree ___Agree ___Disagree ___Strongly Disagree
<i>Option incompleteness</i>	When providing response categories, be sure to include all probable options. Also, you may want to use an “other” category to make sure you have covered all potential responses.	Example: How did you hear about this program? ___child’s school ___Dept. of Health & Human Services office ___community bulletin board ___flyer	Alternative: How did you hear about this program? ___child’s school ___ Dept. of Health & Human Services office ___community bulletin board ___flyer ___friends/family ___other _____
<i>Biased question wording</i>	The wording of questions can influence participant ratings when opinions are not firmly held or the information requested is of a sensitive nature.	Example: A primary goal of this program is to promote healthy eating. Do you think your family is eating more healthy foods as a result of this program? ___yes ___no	Alternative: To what extent has this program helped you make healthier food choices for your family? ___Did not help ___Helped a little ___Helped a lot

Guidelines for Developing Questions⁷⁸

What To Avoid	Explanation	How NOT To Say It	How To Say It
<i>Response order effects</i>	The order of response categories can influence ratings when opinions are not firmly held or the information requested is of a sensitive nature	Example: To what extent has this program helped you make healthier food choices for your family? ___ Helped a lot ___ Helped a little ___ Did not help	Alternative: To what extent has this program helped you make healthier food choices for your family? ___ Did not help ___ Helped a little ___ Helped a lot

When writing instrument questions, programs may be challenged to find the best way of wording their response options. For example, a “yes or no” question may be worded and ordered differently: Yes/No, Yes/No/Maybe, No/Maybe/Yes/Strong Yes, YES/yes/not sure/no/NO, Yes/Sort of/Not really, or Yes/Kind of/Not really. Response option scales may be sought for agreement, quality, importance, frequency, likelihood, amount, change or knowledge attainment. Appendix VIII provides some different options for each of these response scale categories.

The final consideration in selecting response options is the use of open-ended versus multiple-choice questions. An open-ended question is one for which the responses to the question are left blank, often with some lines for the respondent to complete. Multiple-choice questions provide a possible list of responses for each question (see the box below for an example).

Example of an Open-Ended vs. Multiple-Choice Question

Open-Ended Question:

In what ways have your family’s food choices changed since attending this program?

Multiple-Choice Questions:

How have your family’s food choices changed since attending the program? Do you now eat

much less healthy	somewhat less healthy	same amount	somewhat more healthy	much more healthy
----------------------	--------------------------	-------------	--------------------------	----------------------

OR

How have your family’s food choices changed since attending the program? (Check all that apply)

- We eat more fruits
 - We eat more vegetables
 - We eat more locally grown food
 - We eat more organic food
 - We eat fast-food less often
-

Although open-ended questions are useful because they do not place words in the mouths of respondents, they may elicit unintended responses. For example, one person may respond to the open-ended question above with “We eat more fruit and vegetables,” while another responded with, “We stopped eating Burger King” or, from another respondent, “We eat less sugar.” This variation in responses can be problematic because it may represent different interpretations of the question rather than different attitudes and it is difficult to provide meaningful summaries to your audiences.

Furthermore, younger respondents and participants with lower literacy levels may have difficulty completing open-ended questions. It is easier to respond to a statement than create one. In addition, recognizing that something is true is cognitively easier than generating all possible true responses. For these reasons, it is best to use open-ended questions sparingly.

When drafting your response choices for multiple-choice questions, you should consider carefully the literacy level of your respondents. As part of your pilot test, you may want to ask the same question using a couple of different scales to help you decide which types of scales will work best for your program.

Evaluation of Sensitive Topics

Some community food projects work with participants on issues that may be considered sensitive. There may be concern about a participant's willingness to reveal sensitive or socially undesirable information. Collecting data on topics like household finances or unhealthy food choices may require extra thought about how best to assure the respondents that the information they provide is confidential. As was discussed earlier in this section, written or self-administered questionnaires provide greater anonymity than telephone or personal interviewing (personal interviewing generally guarantees the least anonymity). Some recommendations to help respondents feel more comfortable disclosing sensitive information follow:

- Guarantee anonymity whenever possible (do not ask respondents to put their name or a unique identifier on their surveys) and assure participants that others in the program will never see their answers.
- When introducing the evaluation, emphasize importance of truthful responses and that there are "no right or wrong answers." Try to convey that participants will not be judged.
- The sensitivity of the topic may vary by population subgroup. Pilot-testing the tool in your participant population will help you determine the sensitivity of each topic.

Tips for Increasing Accuracy of Sensitive Information	
Self-Administered Questionnaires	<p>Consider using answer sheets separate from the instrument.</p> <p>Allow respondents to place the survey in a locked box or send it through the mail to reduce personal interaction with administrator.</p>
Personal Interviews/Phone Interviews	<p>Use an administrator who is unfamiliar to respondent. Familiarity with the interviewer decreases the accuracy of self-reported answers. Also, interviewers with socio-demographic characteristics like those of the respondent can help to increase accurate reporting.</p> <p>Train interviewers in the importance of presenting a neutral attitude toward responses, including avoiding subtle and non-verbal cues of acceptance.</p> <p>Create an environment where respondents feel secure that others will not overhear their responses.</p> <p>Address confidentiality concerns, if necessary, by explaining to respondents how their responses will be handled, who will see them and how the results will be reported.</p>

Pilot Testing Your Evaluation Tools

Once you have crafted your tools and designed your data collection protocols, the next step is the pilot test. Even if your instrument has been used and tested in other settings, you should test it to make sure it will work for your population.

To conduct a pilot test, select five to ten of your program participants who vary in terms of age, literacy, gender, ethnicity and any other characteristics that may influence the way someone may respond to or perceive your evaluation tools.

Pilot test the instrument with each respondent individually. Ask him or her not only to complete the survey or interview but to also “think aloud” while completing it. The respondent might tell you, “I’m not sure what this question is asking but I think it’s asking this...” or “I don’t know this word...” or “None of the choices on this scale apply to me but if I have to choose I would pick this one.” The information you receive from this pilot test respondent should help you decide whether the instrument is appropriate for a particular age group or culture, answers the questions intended and is easy for respondents to understand. In addition, time how long it takes each respondent to complete the survey. After completing the instrument, ask the respondent a number of questions:

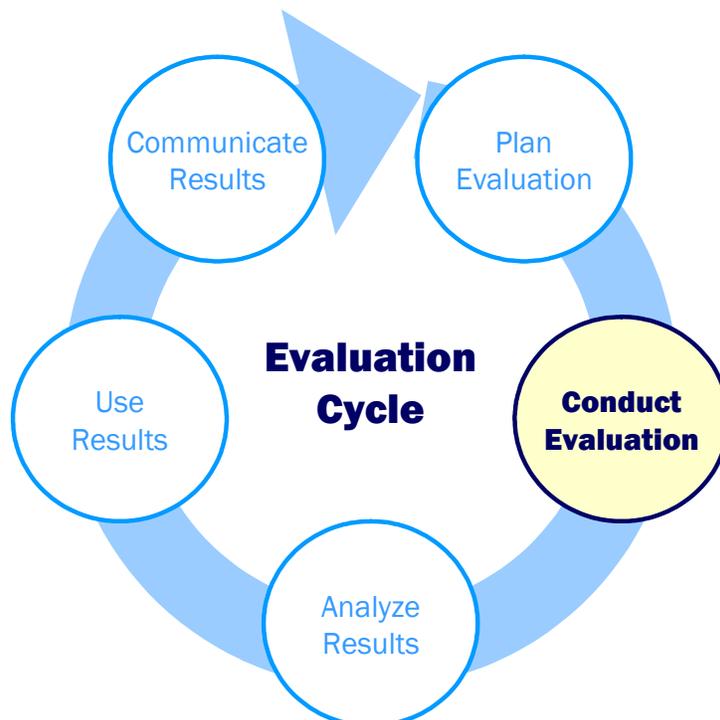
- Overall, how did you feel about the questions?
- Were there any parts that were hard to answer or confusing?
- Were there any parts of the survey that made you feel uncomfortable or any questions that you felt like you did not want to answer?

- Did you or do you think other people would feel angry being asked any of the questions in the survey?
- Do you think other people would be tempted to answer less than truthfully on any of these questions? If so, which ones?
- How do you think other participants in the program might feel about participating in this activity?

You may also want to ask additional questions specific to topics, questions or response scales that you feel more uncertain about. After discussing the tool, look over the completed surveys to make sure respondents answered the questions correctly. You may find questions that need additional work.

Chapter 7. Collecting Data

- 🌸 Determining How Many to Include in Your Evaluation
- 🌸 The Use of Incentives in Evaluation
- 🌸 Developing Data Collection Protocol
- 🌸 Guidelines for Evaluation Administrators
- 🌸 Understanding and Protecting the Rights of Evaluation Participants
- 🌸 Collecting Data in a Culturally Responsive Way
- 🌸 The Key to Ethical Evaluation



Chapter 7. Collecting Data

The data collection process is often the most time consuming component of evaluation. It is important to assign staff members with the specific responsibilities to collect data and allow them enough time to ensure that it is done accurately.

Determining How Many To Include in Your Evaluation

Often the question emerges, "How many people should we ask?" The answer is, "It depends." It depends on the number of people served in a year and the cost in program resources to get those people to answer the questionnaire.

Guidelines for How Many People to Include in Your Evaluation

1. If you can, get them all.
 2. No matter how many people you serve, if resources permit, it is easier to survey them all to avoid designing special data collection methods that sample only a portion of your participants. Plan to ask every individual (or one person from every family) to complete the instruments (before and after service delivery if need be). Plus, with data collected from many participants, you will have greater strength to examine the impact of activities on subgroups (e.g., immigrants versus longer-term residents, non-English speakers versus English speakers, etc.).
 3. If you serve fewer than 100 individuals in a year, try to survey them all. With only a small sample size of participants, it is best to retrieve information from all of them.
 4. If you serve over 100 individuals in a year and cannot afford to have them all surveyed (due to staff time expense or the expense of data collection and analysis when you are successful), sample as many as you can but no fewer than 100.
 5. You can get a reliable read on participant outcomes even with 100 participants out of 3,000. Opinion polls of the American public are based on far smaller percentages (only 1,000 residents out of 80 million or a little over one thousandth of one percent).
-

Develop a Plan for Selecting Your Sample

For larger programs choosing to evaluate a limited number of participants, the next step is selecting your sample. If you expect 500 participants in a year, perhaps choose every fifth. If you expect 200 in a year, choose every other participant. Your goal would be to have 100 participants included. This is the best way to select your sample but it takes some careful attention and consistent management. It may be easier to set aside time in a single month to capture outcomes from every participant rather than limiting your sample size. For example, choose the second week of September to reach every participant served.

However, be cautious about how you choose those who participate in the evaluation. It may not be appropriate to include all participants who enter your program in the second half of the year or in December because they are in some ways fundamentally different than others (all elderly, new immigrants, etc.). Do not just include those

participants engaged in one or two selected activities unless you are confident that the characteristics of these participants are representative of all the participants attending your program.

Responding to Concerns About Time Burden on Your Participants

- As you begin to design your database and refine the information you already ask, you may find that a lot of information about the individuals you serve is collected simply to meet the requests of various funders. You may have concerns about asking for more. Consider the following:
 - With careful planning and good database design, you may actually end up asking for less information by reducing redundancy and unnecessary questions. Reviewing your data collection forms offers an opportunity to improve the overall flow of the questions, making the forms easier to complete.
 - In general, people enjoy the opportunity to give their opinions, especially if their opinion will make a difference. Declare your desire to enhance your activities based on the information you receive from those you survey.
-

Response Rates

It is rare to get 100% of any sample or population asked to participate in an evaluation to actually do so (unless you have a captive audience; and even then there may be problems). A response rate is the proportion of respondents completing the evaluation divided by the number of people who were invited to participate in the evaluation. This response rate gives the audience a sense of how representative your data are of the population you serve. For example, if 100 food recipients are asked to complete a survey at two separate sites, and one site gets ten completed surveys returned (a 10% response rate) and the other gets 90 (a 90% response rate) returned, the second site’s data are likely to provide a better representation of the population served at that site. When response rates are low, one has to wonder if those participating in the survey are a good measure of others. They may be a program’s biggest supporters or biggest critics, not representing the average participant.

Good Response Rates for Surveys	
Mail Survey:	50%+
Phone Survey:	50%+
Handout Survey:	80%+
In-Person Interview:	90%+

You may want to include your response rate in a sentence or two when documenting the data collection methods used in your report. When calculating a response rate, be sure to include those who could not be reached (for phone or mail surveys) as well as individuals who refused the evaluation.

Factors Which Increase Response Rates for Self-Administered Questionnaires

- Keep the survey instrument short, simple and attractive
 - Guarantee anonymity or confidentiality
 - Explain purpose of survey either verbally or with a cover letter, including an appeal for help to improve services or activities
 - Make survey completion as convenient as possible – e.g., the last 15 minutes of an already scheduled meeting
 - Ask participants to complete the survey while on premises
 - Charge program staff with responsibility for making sure surveys are administered and collected appropriately
-

The Use of Incentives in Evaluation

A community food project may choose to use incentives, such as a monetary donation, as a way of enticing individuals to participate in their evaluation. Incentives are a good way to increase participation and demonstrate that a program values participants' time. One of the downsides of incentives is the cost, although some argue that incentives may introduce more bias because they compel participants to answer more favorably because of payment received or may select for specific sub-populations based on the type of incentive. For example, if a coupon worth \$1 is given for completion of a survey, participants with lower incomes may be more likely to participate than higher income participants.

Many CFPs may not need incentives because participants are often motivated and place a high value on the program. Further, some programs employ their participants or provide services in a way where the participants serve as captive audiences for the evaluation (e.g., they are all on site). In these cases it may be unnecessary to provide incentives for participation in evaluation activities.

However, some CFPs may need incentives to help increase participation in their efforts. Incentives can be used to increase participation if low response rates are expected or become an issue. Incentives can be used in a variety of ways:

- Advance payment for future work: incentives are given to all participants asked to take part in the evaluation regardless of their response. These incentives are often less significant and play more on the "guilt factor" rather than reward people with reasonable compensation for the work performed. (This approach is often used by charitable organization asking for donations. For instance personalized address labels, pens, a penny, etc. are mailed to motivate people to give.)
- Payment for work completed: participants who complete the survey are given an incentive. These incentives are often more significant like t-shirts, hats, snacks, bumper stickers, coupons for use at farmers' market, etc. A good option

for programs with limited budgets is the use of a lottery or drawing. Every participant completing a survey is entered in the lottery to win a prize.

Note that if incentives are paid for work completed, there must be a system in place to protect the anonymity or confidentiality of the respondents (e.g., people cannot be asked to sign the survey so they get their prize). A common way to get around this dilemma is to ask individuals turn in a second form indicating the survey is complete at the same time the actual completed survey is returned. The surveys and completion forms are kept in two different boxes.

Instead of offering money as an incentive, consider alternatives such as a coupon to the local farmers' market, gift certificate to a local eatery, a jar of locally made salsa, etc.

Developing Data Collection Protocol

Ensuring accurate data requires that procedures be put into place to encourage consistent data collection. Consistent data collection is important to guarantee accurate, trustworthy results. For each data collection tool you are using, the following questions must be answered. (A worksheet follows for you to complete for each of your evaluation tools.)

Who is eligible? All participants? Only those who successfully completed the program (What is considered a successful completion?), those who have been involved at least a certain amount of time? Only those who have attended at least a certain percent of the program's activities during a particular season?

Will all eligible participants be included? If no, you will need to develop a sampling plan.

Who will complete the evaluation tool? The participant, a proxy (e.g., English-speaking child, a parent or guardian), a volunteer, a program staff member? If others rate or observe the participants, are all raters using a consistent set of guidelines? Do they have a common understanding of the words used (e.g., the definition of "yield" or "good marketing skills")? Have they been trained on the guidelines?

When are the data collected? On what date(s) and time(s)? At program start, program completion, mid-season, follow-up after program completion? When the program begins for the day or near the end? More than one date may need to be scheduled if it will improve the overall response. Provide adequate amount of time for completion so that participants do not feel rushed or like they are missing out on other program activities.

Will the evaluation tool be administered in a language other than English? If so, arrangements will need to be made for the survey to be translated and individual(s)

administering the survey will need to speak this other language as well to recite the explanatory text and be available to answer questions.

Who will administer the evaluation tool? Administrative staff, program leaders, the program administrator, selected participants? Name names.

Who will gather evaluation tool administration supplies? This may be the same or a person other than the survey administrator. Name names. Supplies may include pens/pencils, copies of surveys (in English and other languages if appropriate), survey completion box.

What introductory text will be recited when distributing the evaluation tool? Will all staff administering the tool be using similar instructions? A common script should be used as a guide to introducing yourself and the evaluation to the participants. This script may include information about why the evaluation is being conducted, how it will be used, whether anonymity or confidentiality will be provided and general instructions for participation. Sample introduction scripts are provided later in this chapter.

Where will completed tools be returned? A box with a drop slot may be used or perhaps survey completers may be instructed to return completed surveys upside down on a desk or chair at the back of the room. If the intention is to collect responses anonymously from participants, the way in which completed tools are returned may impact this decision.

Protocol Modifications for Participants with Limited English or Low Literacy Skills⁷⁹

- Items (and responses) can be read aloud to participant.
 - Item wording can be simplified, so that questions are more appropriate to the reading level of the participant.
 - If given in interview form, the survey can be presented in a more lighthearted manner, so that it does not “feel like a test.”
 - Provide an adequate amount of time for participants to complete the survey.
-

Sample Data Collection Protocol: Program Satisfaction Survey	
Question	Response
1) Who is eligible to complete the survey/?	All participants who have attended the program at least two months.
2) Will all eligible participants be included?	Yes.
3) Who will complete survey?	Participants.
4) On what date(s) and time(s) will the data be collected?	September and April (one week will be randomly selected and the surveys will be administered every day of the survey week), near the end of the program (around 3:40pm).
5) Will the survey be administered in a language other than English? Explain.	Yes. Surveys will be translated into Spanish by our volunteer, Frank, and available for participants to complete.
6) Who will administer the survey (in English)? In other language?	Cynthia will hand out the surveys 20 minutes before the program ends.
7) Who will gather survey administration supplies?	Zoë will gather supplies (survey copies, pencils and survey box) and provide to Cynthia for administration.
8) What explanatory text will be received when distributing the survey?	<p>“This survey is being used to get your opinions on the program. It is not a test and there are no right or wrong answers. Please take your time and be sure to answer each question based on what you really think.”</p> <p>“Please do not put your name on this survey. Your answers are completely private.”</p>
9) Where will completed surveys be returned?	A box with a drop slot will be placed at the door. Once participants complete the survey, they will fold it in half and drop it in the box.

Worksheet 6: Questions to Help Design a Data Collection Protocol

Program: _____	
Question	Response
1) Who is eligible to complete the survey/evaluation tool?	
2) Will all eligible participants be included?	
3) Who will complete the survey/evaluation tool?	
4) On what date(s) and time(s) will the data be collected?	
5) Will the survey/ evaluation tool be administered in a language other than English?	
6) Who will administer the survey/evaluation tool?	
7) Who will gather survey/ evaluation tool administration supplies?	
8) What explanatory text will be recited when distributing the survey/evaluation tool?	
9) Where will completed surveys/tools be returned?	

Guidelines for Evaluation Administrators⁸⁰

The individual selected to administer the survey, facilitate the focus group or otherwise manage the evaluation process is responsible for effectively collecting data on the project. This includes making sure that participants understand the purpose of the evaluation, are comfortable participating in the evaluation process and the evaluation takes place according to established protocol.

The evaluation administrator may be a program staff member or volunteer who is not directly responsible for the project work with participants, as participants may not feel as comfortable indicating their feelings about the activities, staff or their actions. If there is no other person besides program staff who can reasonably administer the evaluation, then it is imperative that the administrator encourage participants to give accurate and straightforward answers to the evaluation and to stress the privacy of the evaluation process.

Following are considerations for the evaluation administrator to take into account.

Understand the Evaluation Tools

Prior to conducting the evaluation, it will be important for evaluation administrators to be familiar with the evaluation tool(s). Administrators may want to complete the survey or answer the questions themselves, reading the directions out loud. To anticipate participant questions regarding the tool, administrators may role-play questions and answers and then discuss and agree upon consistent responses to common questions in advance.

Standardize the Administration Process

To ensure that information collected from all participants is comparable, each administration should be standardized. In other words the evaluation tools should be administered to each participant in the same way as much as possible. For self-administered surveys, someone should be available to answer any questions or clear up any confusion the participant might have about the survey. For interviews, the interviewer/administrator should read the questions as written; in the order they are presented in the instrument, and should allow the participant to select their answers without influence by the interviewer. Also, in both situations, but particularly in the interviewer-administered situation, there should be no negative reactions to any attitudes or behaviors that the respondent reveals.

Maintain Neutrality toward Participants

To ensure that participants do not feel compelled to answer evaluation questions in a way to please the interviewer or administrator, neutrality must be maintained. The administrator should keep a neutral tone and expression while still being friendly and professional. The administrator should not be judgmental regarding their knowledge,

attitudes, or behavior; any judgments the administrator does have should not be able to be detected by the participants.

Use Common Administration Protocol and Procedures

The data collection protocol worksheets described earlier in this chapter will help guide an evaluation administrator through each administration step before the evaluation takes place. As the protocol describes, common introductory text may be recited before surveys are distributed or evaluation questions are asked. This common text serves as a guide for survey administrators, but the language should be unique and reflect the administrators own style of communication to help build rapport with participants right from the start. Following is a sample script for use with a survey administration.

Sample Survey Administration Script

Hi I'm (insert your name here). I am here to ask you to participate in our Family Farm program evaluation. The survey I am about to pass out is being used to get your ideas about our farm-to-school program. It is not a test and there are no right or wrong answers--we want to know what your experiences and opinions are. Take your time and be sure to answer each question based on what you really think. If you cannot read or don't understand a question, please raise your hand or come find me and I can help you.

The survey should take you only 8 to 10 minutes (May differ depending on survey length). Make sure you do NOT put your name on the survey (if survey to be anonymous) so that we can keep your answers private. Okay, let's begin.

Once the evaluation is completed by participants, kindly thank them for their time and effort. At this time, incentives should be provided if appropriate.

Assistance during a Survey Administration

There may be times during a survey administration when a participant needs extra assistance with understanding a question, have difficulty reading a question or complete the survey more slowly or quickly than others. Following are a few suggestions for managing these concerns.

- Discretely check on a participant who appears to be going too slowly or too quickly. If additional assistance is needed, you may want to offer to read them the questions in another room or space so as not to disturb other survey participants.
- If a participant does not understand a question or a response, read the question or response aloud to them; emphasize key words that may help convey meaning and ask if there are any words that the participant does not understand. Try to paraphrase the confusing word or phrase, giving word or phrase options until

the participant understands the question. However, keep as close to the original question as possible without leading the participant to any particular response.

- If a participant is concerned about who will see answers remind them that only members of the program team and the evaluators will see their individual surveys. No one will ever see his or her name attached to the survey. And no information will be presented on the individual respondents; only group averages will be used in reports.

Understanding and Protecting the Rights of Evaluation Participants

Individuals choosing to participate in an evaluation must be guaranteed certain rights. They have the right to privacy of their answers and of their participation in the study. They have the right to refuse participation and they have the right to understand the risks and benefits of participating. These protections apply to evaluation participants of all ages, especially more vulnerable populations like children, the elderly and those socio or economically disadvantaged. These protections are outlined in the following sections.

The Guarantee of Anonymity or Confidentiality

Evaluators and program staff alike must strictly honor a participant's right to privacy. If you promise anonymity, no identifying information may be collected from respondents, not even their phone numbers. If you promise confidentiality, programs must enforce clear rules that prohibit access to any information that would identify a particular participant, unless you have specifically received their consent to do so. If there is no need to link respondents' names with the responses, then do not ask for their name (e.g., for participant satisfaction) or use separate identifying codes from actual names and destroy the code-name link (by purging the name) as soon as the necessary data are linked. If the data need to be linked to other data collection forms (e.g., pre to post outcome changes), use an identification code or number rather than putting the name on the top of the form. If you use identification codes, you no longer can promise anonymity, only confidentiality. Respondents should always be told that their answers are anonymous or confidential and will be reported only as a group. These basic tenets apply whether data are kept in paper files or on computer.

For younger participants, the terms "anonymity" and "confidentiality" may have little meaning, so an introduction like the following may be used.

Please help us improve our program by answering the following questions. We are interested in your honest opinions, whether they are positive (good) or negative (bad). DO NOT put your name on this form, so that NO ONE will know your answers (not your parents, the teachers in this program, or anyone else).

Respect for the participant is the cornerstone of ethical evaluation and research. Part of respect is maintaining the confidentiality of participants. The evaluation administrator may not discuss any information disclosed by participants with anyone unrelated to the evaluation. Even if a respondent asks the administrator to tell someone an answer they gave in a survey, the administrator is not permitted to do so.

Participants should be reminded of their confidentiality rights whenever the administrator feels it is appropriate, and may refer participants to their consent form (if used), which guarantees these rights. The only exception to confidentiality: If the respondent shares information revealing that he or she might harm themselves or others, then the evaluation administrator must follow procedures that are legally mandated in the state to report such information.

Voluntary Participation

Project participants may not be forced to participate in the evaluation. Instead, an administrator might stress the importance to their project to learn how best to deliver services or give other appropriate explanation. Participants should understand that they are free to not answer individual questions and their participation or withdrawal from the evaluation would have no impact on the services they receive.

Obtaining Consent to Evaluate

Consent forms are designed to protect evaluation participants by informing them of their rights and any risks and benefits as a participant in the evaluation. The statement of consent describes the benefits and possible dangers to participation in the evaluation. It allows participants to understand what the evaluation is about and what they are being asked to do and obtains their written permission to engage in the evaluation process. Projects may also choose to include a waiver of liability for their project or organization.

For projects that work with children or youth, it is recommended that for youth to participate in a research or evaluation study, their parents or guardians must sign an informed consent statement permitting their child to take part in data collection. It may make most sense to get parent or guardian signature at the same time that parents sign their approval for program participation of their child.

The US Department of Health and Human Services, Office of Human Research Protections (OHRP), offers guidelines for what to include within informed consent documents. Following is a link to OHRP's Tips on Informed Consent:

<http://www.hhs.gov/ohrp/humansubjects/guidance/ictips.htm>.

Institutional Reviews

For a few projects, like those connected with a university or government entity, staff may be required to submit the proposed evaluation tool, study design, methodology and informed consent documents to an Institutional Review Board (IRB) regardless of whether their program serves youth or only adults. Requirements may differ per organization so projects are encouraged to check with appropriate agencies as needed.

Collecting Data in a Culturally Responsive Way

There are many factors that may influence a program's evaluation administration success with culturally diverse communities. Highlighted below are some matters to ponder.

- Be intentional when selecting facilitators or interviewers for qualitative evaluations, matching their culture to participant culture when possible.
- Consider inclusion of translators or co-facilitators/interviewers as necessary. A cultural advisor may host or co-facilitate to help participants feel at ease.
- Consult with cultural advisors about culturally appropriate refreshments and customs associated with food, if food served during administration.
- Ask recruited evaluation participants about any special needs they might have and provide appropriate accommodations.
- Learn of the cultural community's experience, locally or historically, with past research or evaluation and consider these events when collecting data.
- Be respectful of cultural norms when selecting evaluation administrators: Be aware of body language, communication styles and gender roles.
- Provide transportation and/or childcare if necessary and consult with cultural advisors on how this should most appropriately be provided.
- Going door to door to recruit and interview individuals may be culturally appropriate and pragmatically necessary. One evaluation team shares the following story:

Homes in this rural community are located in isolated areas, often with no telephone. Mail is retrieved only periodically from the Post Office box located many miles from recipients' residences. Some letters sit unopened for a long time before a family member or friend reads them. Therefore, Native American members of the community assessment committee were hired to recruit and interview participants (in their homes).⁸¹

The Key to Ethical Evaluation

In addition to issues of consent, anonymity and confidentiality, people conducting evaluation research should be informed about other ethical guidelines that apply. As with many disciplines, there are codes of ethics for evaluation. Most of these codes

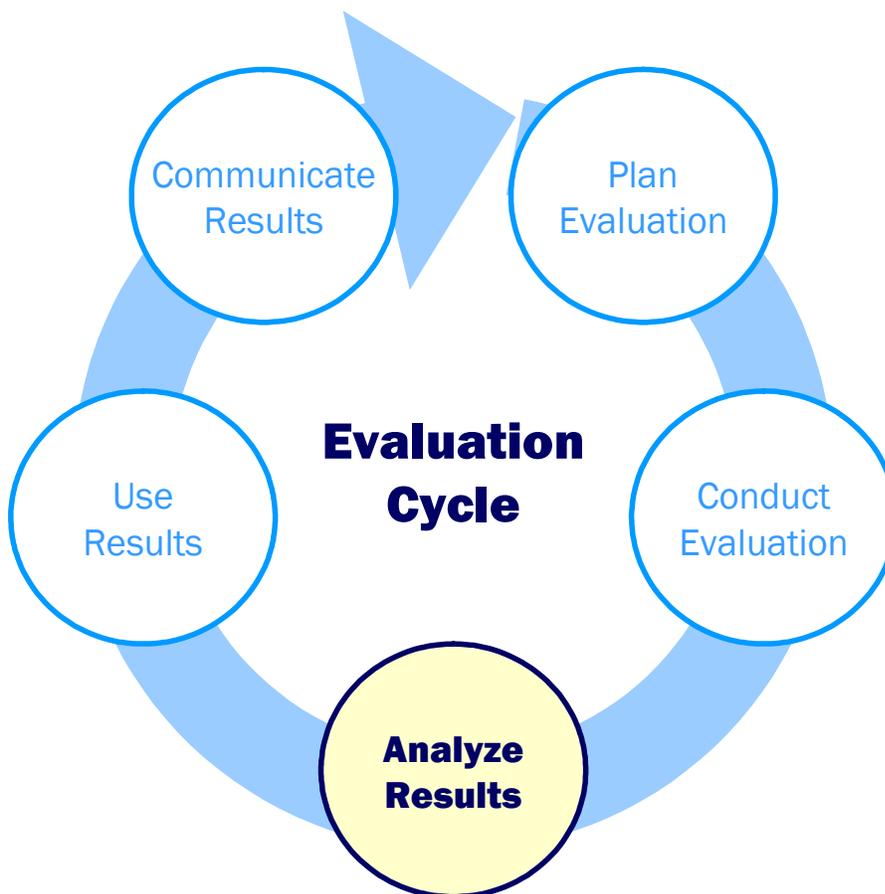
speak to the evaluators need to be aware of their own belief systems, values, needs and limitations and the effect of these on their work. In addition, the codes guide evaluators to respect the security, dignity and self-worth of the respondents, program participants, and other stakeholders with whom they interact. Thus, it is necessary for evaluators and all those involved in evaluation to understand the ethical issues in the evaluation process. The following information introduces ethical issues regarding respect for others.⁸² For more information on ethical principles please refer to the ethical guidelines set forth by federal regulations about protection of human subjects, or the ethical principles of such associations as American Evaluation Association, American Anthropological Association, American Educational Research Association or American Psychological Association.

- Where applicable, evaluators must abide by current professional ethics and standards regarding risks, harms and burdens that might be engendered to those participating in the evaluation; regarding informed consent for participation in evaluation; and regarding informing participants about the scope and limits of confidentiality.
 - Evaluators have the responsibility to respect differences among participants, such as differences in their culture, religion, gender, disability, age, sexual orientation and ethnicity. Evaluators must be mindful of potential implications of these differences when planning, conducting, analyzing, and reporting their evaluations. It is essential that evaluators try to eliminate the effect of these potential biases on their work and that they do not knowingly participate in or condone activities of others based upon such prejudices.
 - Evaluators must accord appropriate respect to the fundamental rights, dignity and worth of all people. They not only respect the rights of individuals to privacy and confidentiality, but also to self-determination and autonomy, and are aware that special safeguards may be necessary to protect the rights and welfare of persons or communities whose vulnerabilities impair autonomous decision-making.
 - Children are a group with special needs for extra protection and support. While research on children and their living conditions is valuable and important, their needs and interests often have to be catered for in other ways than when research concerns adult participants. Children are developing, and their needs and abilities change from phase to phase. Children are consequently exposed to other risks than adults.
 - Informed consent may be bigger issues with children because they often are more willing to obey authority, regardless of their own wishes, and do not always have a full understanding of the consequences of giving researchers information. For instance, children may view the publication of anonymous data as a breach of confidence. The consent of their parents or guardians is often sufficient to safeguard children's interests, but there may also be conflicts of interest between children and their parents or guardians.

Chapter 8.

Analyzing Your Data

- 🌸 Creating an Evaluation Notebook
- 🌸 Developing an Analysis Plan
- 🌸 Quantitative Data Analysis
- 🌸 Qualitative Data Analysis
- 🌸 Analysis of Secondary Data



Chapter 8. Analyzing Your Data

Data analysis may very well be the most intimidating part of program evaluation. A large part of data analysis is statistical in nature. While many of us may have academic experience in statistical analysis, translating this into a useful analysis for direct services may be daunting. Analysis of program data, however, has become essential to program management, improvement and continued funding. In this chapter we hope to simplify the world of data analysis by describing each step in its practical relevance and application in order for programs to find ways of adapting such strategies to their setting.

It may be helpful for your program to identify the individual on staff with the most knowledge and interest in analyzing your program data. This person should be comfortable using a computer, and not be afraid of numbers. Often an administrative assistant or someone who manages the bookkeeping or finances will make a good choice (make sure they agree). Review this chapter with whoever is chosen to assist with data analysis. If your program has limited staffing resources or analytic abilities, you may consider as an alternative, seeking an outside evaluator or graduate student to help with these next steps of your evaluation.

Creating an Evaluation Notebook

Creating an evaluation notebook can be an excellent tool for tracking all of your evaluation information. Once the analysis starts, there is plenty of important information to keep track of and having it collected in one place will make it easier. The notebook should include items you may already have created, such as copies of the final evaluation tools (hard copy and electronic on disk) and data collection protocols. You will also want to add any items you may create as a result of going through this chapter, such as your analysis plan, codebooks, coding sheets and data printouts. You may also find it helpful to have pages of reflections on your evaluation process – what went well, what did not and improvements you might want to make for the next go-round. This notebook will ensure that all current information is in the same place as well as provide a quick look-up when questions arise during the next evaluation or as new staff are assigned evaluation tasks.

Developing an Analysis Plan

A simple analysis plan can be created by elaborating on your evaluation plan worksheet (see Chapter 4, Worksheet #5). The analysis plan will list all of the specifically analysis to be performed on the evaluation data. An example of such an analysis plan is presented along with a worksheet on the following pages.

The plan is quantitative in nature; that is, it focuses on counts and proportions. More complex analysis plans will need to be created for more complex evaluation designs or those using qualitative data collection methods. (For more on evaluation designs, see Chapter 5. Selecting Evaluation Strategies and Study Designs). Even with more complex analyses, however, it is important to “start with the end in mind,” and to let your evaluation questions guide the analysis. The worksheet provided often can be used for simple or even complex evaluations.

Sample Evaluation Plan: Neighborhood Garden Project				
Program Goal	Indicators (Outputs and Outcomes)	Data Sources	Performance Standard	Data Analysis
To increase gardening skills for participants	<p><i>Increased knowledge of gardening practices</i></p> <ul style="list-style-type: none"> • Number of volunteers trained • Total volunteer time • Description of volunteer activities 	<ul style="list-style-type: none"> • Administrative records • Volunteer activity logs • Volunteer survey 	<ul style="list-style-type: none"> • 10 volunteers total, including two core volunteers • Total volunteer time meets need 	<ul style="list-style-type: none"> • Counts of volunteers trained • Counts of volunteers hours worked • Counts of volunteer hours by activity type
To increase satisfaction among Latino residents with produce offered at neighborhood farmer's market	<p><i>Satisfaction with food selection</i></p> <ul style="list-style-type: none"> • Number of youth participating in garden • Number of hours youth participated in garden • Increase in youth leadership skills • Increase in youth connection to culture/background • Increased consumption of vegetables by youth 	<ul style="list-style-type: none"> • Youth sign-in sheets • Surveys of youth (post-program) 	<ul style="list-style-type: none"> • 25% of participating youth will report an increase in leadership skills • 60% of participating youth will report an increase in their connection to their culture/background • 40% of participating youth will report an increase in the amount of vegetables they eat 	<ul style="list-style-type: none"> • Counts of youth participating in garden • Counts of youth hours • Counts and percentages of youth answering "strongly agree" or "agree" to statements: • Since I came to the garden, <ul style="list-style-type: none"> • <i>I am more of a leader</i> • <i>I feel more connected to my culture</i> • <i>I eat more vegetables</i>

Sample Evaluation Plan: Neighborhood Garden Project

Program Goal	Indicators (Outputs and Outcomes)	Data Sources	Performance Standard	Data Analysis
To increase civic responsibility	<p><i>Community service</i></p> <ul style="list-style-type: none"> • Description of original barriers to obtaining food prior to project • Description of barriers to obtaining food after participation in the project • Quantity of produce grown and sold 	<ul style="list-style-type: none"> • Garden logs of produce grown • Market logs of produce sold • Garden logs of produce taken home by youth and volunteers 	<ul style="list-style-type: none"> • 500 pounds of produce will be grown in the garden in 2003 • 300 pounds of produce will be sold at the market in 2003 • 150 pounds of produce will be taken home for personal use by youth and volunteer growers in 2003 	<ul style="list-style-type: none"> • Counts of pounds of food produced • Counts of pounds of food sold • Counts of pounds of food taken home for personal use
To increase collaboration of food- related community organizations	<p><i>Increased collaboration</i></p> <ul style="list-style-type: none"> • Number of organizations participating in food network • Number of meetings held 	<ul style="list-style-type: none"> • Meeting participation tracking forms 	<ul style="list-style-type: none"> • 8 community-based organizations will attend monthly network meetings 	<ul style="list-style-type: none"> • Counts of organizations • Counts of meetings • Average attendance per meeting

Worksheet 7: Analysis Plan

Evaluation Questions	Indicators	Data Sources	Performance Indicators	Data Analysis

Quantitative Data Analysis

This section describes techniques for performing quantitative data analysis, methods that produce numerical summaries of your findings. Instructions for performing simple qualitative analysis techniques begin on page 136. Those who already know how to use a statistical program such as SPSS or SAS, and are familiar with the process of creating electronic datasets from surveys or other sources may wish to skip this section. Much of this section focuses on using information from a survey. If you have quantitative data from another source, you may continue to follow most of these same steps. For example, forms used to count and classify customers at a Farmer's Market may also be entered into an electronic dataset and analyzed in a similar fashion as that of recording food distribution.

Preparing Your Data for Analysis

Say you have just completed your first survey administration. Now you have a stack of completed surveys in front of you and you are not quite sure how to go about producing some useful information from your pile. You probably are considering a "hand tally" of the surveys at this point, but know in the back of your mind that there must be a much more efficient way of analyzing data from these surveys that may involve using your computer.

You are right. The question is how to go from the large stack of data to a concise computer print out. Basically, you will be taking your stack of surveys and creating a "numeric" electronic dataset that can be analyzed. You may be wondering what will make the dataset "numeric." Almost all analysis programs run more efficiently when they tally numbers rather than words. Once you have some experience with this, you will also find that you can complete the data entry much more quickly using numbers rather than letters or words. Consider the numbers just codes for the words. For example, you may use the number 1 as a code for the answer "yes," 2 for "kind of" and 3 for "not really." This will make more sense as we go further in this chapter.

Before creating your electronic numeric dataset, you will need to prepare the surveys for data entry using the following steps.

Coding and Identification Numbers

If your surveys were administered anonymously, asking respondents not to include their names, then each survey must be assigned an identification number before entering it electronically. This number will allow you to go back to an actual survey at any time for clarification if needed. The unique number is placed in the same spot on each survey (e.g., the upper right corner of the page) and can range from 1 to the number of surveys administered. The survey excerpt on the following page gives an example of how to ID surveys. Specifically note the area highlighted.

Survey Excerpt – Example of ID’ing Survey Forms



Valley Garden Survey

ID=01

This survey is being used to get your opinions on Valley Garden so we can improve it for you and others. We are interested in your honest answers. Please do not put your name on this survey. Your answers are completely private

1. **How long have you been working at Valley Garden?**

_____ weeks OR _____ months OR _____ years

Cleaning Your Data

Next you will want to “clean” the survey data by going through each survey and each survey response to make sure respondents have followed instructions and that the surveys were completed in an acceptable manner (without mistakes). Often survey respondents make common mistakes when completing a survey. The table on the following page lists such errors and provides methods for “cleaning” the results.

Common Survey Response Errors		
Problem	Example	Solution
More than one response: the respondent selects more than one response when instructed to select only one	<p>1. How would you rate the quality of the food sold at this market?</p> <p><input checked="" type="checkbox"/> Excellent <input checked="" type="checkbox"/> Good <input type="checkbox"/> Fair <input type="checkbox"/> Poor</p>	Flip a coin between the two responses and select only one answer (“white out” the answer not selected by the toss). This coin toss will randomly select one response preventing bias.
Wrong questions completed: the respondent completes questions he was not intended to	<p>A new participant who has not attended any courses yet completes question 3</p> <p>2. In addition to Cooking Like Crazy, how many other courses offered by Ida Farms have you attended? None (0) ___1 ___2 ___3+</p> <p>If you have not attended any other courses, please go to question 4.</p> <p>3. Compared to the other courses offered by Ida Farms that you have attended, how would you rate Cooking Like Crazy?</p> <p><input type="checkbox"/> Much more useful <input checked="" type="checkbox"/> Somewhat more useful <input type="checkbox"/> About equally useful <input type="checkbox"/> Less useful</p>	Blank out the inappropriate response (e.g., white out the answer to question 3).
Blank questions: respondent leaves answers blank	<p>10. What is your age (select one)?</p> <p><input type="checkbox"/> 18-24 <input type="checkbox"/> 25-44 <input type="checkbox"/> 45-64 <input type="checkbox"/> 65 or older</p>	In most cases, you will not be able to fill in these answers without possibly biasing the results, so leave them blank.

Open-Ended Questions

Your survey may include one or more “open-ended” questions (e.g., “In what ways have your family’s food choices changed since attending this program?”). For these questions, you may wish to “code” the responses before you perform the data entry. Please refer to the subsection “Coding Your Data” in the next section on “Qualitative Data Analysis.”

Creating an Electronic Dataset

Once your evaluation data are coded and cleaned, it is ready to be entered or “keypunched,” into the computer, thus preparing it for analysis. This creates your electronic dataset. There are a variety of computer software programs on the market

that can assist you with data entry and the subsequent analysis. Three major types of software programs are generally used for these procedures: databases, spreadsheets and statistical packages. All three types of programs allow data entry, storage, retrieval and analysis capability. They vary, however, in terms of the major function they are intended to perform. Also, you may choose to use one program for data entry and storage, but a different program for data analysis. The chart on the following page describes computer software program options and the major functions and strengths of each.

Types of Data Storage and Analysis Programs			
Program Type	Commonly Used Software Products	Major Function of Program	Strengths
Spreadsheets	Microsoft Excel, Lotus 1 2 3, Quattro-Pro	Simple mathematical calculations performed on data	Produces tables and graphs of data once analyzed Can test for statistically significant differences in data
Database Programs	Access, Paradox, Filemaker Pro, DBASE, Rbase	Storage and retrieval of data	Easiest data entry and retrieval Easiest to use with large text fields like names and addresses Ability to generate mailing labels and custom reports with little effort
Statistical Packages (professional, expensive)	SPSS, SAS	Statistical analysis: simple and complex analyses	Can test for statistically significant differences in data Quicker and more complex analysis of data Produces tables and graphs of data once analyzed
Statistical Packages from the Internet	EpiInfo (http://www.cdc.gov/epiinfo/) SimStat for Windows (http://www.simstat.com/simstw.htm) Arcus QuickStat (http://www.camcode.com/arcus.htm)	Statistical analysis: simple and complex analyses	Can test for statistically significant differences in data Quicker and more complex analysis of data Free (EpiInfo) or less expensive (SimStat, Arcus) than SPSS or SAS (about \$150)

If you are currently using one of the programs listed above for other activities and are comfortable with it, this might be your best choice for data storage and analysis. For specific instructions on using Microsoft Excel to conduct data analysis, see Appendix IX: Descriptive Analyses Using Excel or NRC's handbook, *Excel for Data Analysis*, which can be found at the following Web link: <http://www.n-r-c.com/excelhandbook.pdf>.

The layout for creating your electronic dataset will vary somewhat by the type of program you choose. As a general rule, when creating any type of data file, ensure that each column represents participant responses to the same question. An example of a data entry file is provided in the table below:

Example of a Data Entry File					
Name	Participant ID	Age	Gender (female=1, male=2)	Participant Satisfaction (1-10)	Quality of Life Score (1-10)
Sam Rhee	001	14	2	9	7
Serena Smith	002	16	1	8	8
Tom Martinez	003	15	2	9	9

Important characteristics of this file are:

- Each participant has been assigned an individual row or “record”
- The response for each participant to the same question is recorded in the same column (e.g., everyone's age is in the third column from the right)
- Numbers substitute for text and a key is used to explain the coding (e.g., female = 1, male = 2)

Creating a Codebook

It will be helpful to design a “codebook” or layout of how your data will be entered into a file created by the software program. An example of turning a survey into a codebook is included on page 130. A “codebook” provides a reference guide for understanding your data printout. You will generally be using shortened words and numeric codes to represent the questions and responses on your data collection tool so the codebook will help you decipher these codes. In addition, the codebook will help others better understand the analysis work that was done when it comes time to re-administer the study. Some basic terminology and concepts used in data entry and analysis programs are presented in the table on the following page.

Terms and Concepts Used in Data Storage and Analysis Programs		
Term	Explanation	Example
Record	One row of data. You will generally set up your program so that each participant has an individual row or “record.”	Synonyms ID Age Gender One record → 01 16 1 One row → 02 18 2 One person → 03 15 2
Variable or field	The descriptor for one piece of information. A variable or field will represent each question on your survey or data form. Variables are generally represented as	One variable or field ↓ ↓ ↓ ID Age Gender 01 16 1 02 18 2

Terms and Concepts Used in Data Storage and Analysis Programs																		
Term	Explanation	Example																
	column heads in statistical software programs - they are the basic unit of database or analysis software. Each row is composed of a series of fields.	Variables included in this dataset are: Participant ID, age and gender.																
Value	The response a specific variable can take is the value. This is generally the specific reply to a question (often coded to be a number).	Values for the variable sex are "1" (female) and "2" (male). Values for age are the age in years of the respondents.																
Field width	The number of typed characters that correspond with the largest possible value a variable requires.	If your participant identification numbers range from 1 to 320, you will need a field that is the width of "3 characters" (or 3 numeric digits) for the variable "ID", e.g., <table border="1" data-bbox="959 737 1362 873"> <tr> <td>Field width:</td> <td>#1</td> <td>#2</td> <td>#3</td> </tr> <tr> <td>Number:</td> <td>0</td> <td>0</td> <td>1</td> </tr> <tr> <td>...</td> <td colspan="3">to</td> </tr> <tr> <td>...</td> <td>3</td> <td>2</td> <td>0</td> </tr> </table>	Field width:	#1	#2	#3	Number:	0	0	1	...	to			...	3	2	0
Field width:	#1	#2	#3															
Number:	0	0	1															
...	to																	
...	3	2	0															
Alphanumeric	Indicates a data field where letters are used instead of or along with numbers.	If the values for sex are "female" and "male" instead of "1" and "2", the field is specified as "alphanumeric."																

Refer to the data entry chapter of your software manual for more information on how to enter data into your program so that the program "knows" what information exists in each field.

Example of a survey that now serves as a “Codebook”

“Handwritten” notations have been added to the survey below to “code” the survey. See the structure of this dataset on the following page.



Column A: ID = 001

Tell Us What You Think about Ida Farm!

1. Please rate the food that you have received from Ida Farm by circling the number that comes closest to your opinion:

	poor	fair	good	very good	excellent
<p><i>Column B - Q1 i.</i></p> <p>The quality of the vegetables delivered</p>	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
<p><i>Column C - Q1 ii.</i></p> <p>The variety of vegetables delivered</p>	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
<p><i>Column D - Q1 iii.</i></p> <p>The timeliness of the deliveries</p>	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5

2. To what extent has this food delivery service met your needs? *Column E - Q2*

None of my were needs met..... 1

Only a few of my needs were met 2

Most of my needs were met 3

The service met all of the expected needs 4

3. Has the delivered food helped you to eat more fresh vegetables? *Column F - Q3*

No, I am eating fewer vegetables..... 1

No, I am eating about the same amount 2

Yes, I am eating a few more 3

Yes, I am eating a great deal more..... 4

4. How has the quality of your life changed as a result of the services you received by this program? Is it... *Column G - Q4*

Much worse..... 1

Somewhat worse 2

Somewhat better 3

Much better 4

...than when you began the program?

5. What foods do you like best? *Column H - Q5*

1 Tomatoes

2 Cabbage

3 Snow peas

4 Squash

5 Cucumbers

6 Carrots

7 Other _____

6. Are there any improvements you would like to see made to the Ida Farm program? (Check as many as apply.) *Column I - Q6*

1 More flexibility in scheduling deliveries

2 Delivery of fruits

3 Add easy preparation recipes

4 Other _____

Example: Data Analysis Structure for [as described by example codebook on previous page] Sample Participant Satisfaction Survey

Variables in dataset	Variable Label	Values
ID	Participant identification number	ID number used on survey
Q1 i	Quality of vegetables	1=poor, 2=fair, 3=good, 4=very good, 5=excellent
Q1 ii	Variety of vegetables	1=poor, 2=fair, 3=good, 4=very good, 5=excellent
Q1 iii	Timeliness	1=poor, 2=fair, 3=good, 4=very good, 5=excellent
Q2	To what extent has service met your needs	1=none of my needs met, 2=only a few needs met, 3=most needs met, 4=all of my needs met
Q3	Has it helped you to eat more vegetables	1=no, I am eating fewer vegetables 2=no, I am eating about the same amount 3=yes, I am eating a few more 4=yes, I am eating a great deal more
Q4	How has your quality of life changed	1=much worse, 2=somewhat worse, 3=somewhat better, 4=much better
Q5	Food liked best	1=tomatoes, 2=cabbage, 3=snow peas, 4=squash, 5=cucumbers, 6=carrots, 7=other
Q6	Improvements	1=more flexibility in scheduling deliveries 2=delivery of fruits, 3=add easy preparation recipes, 4=other

Example of Survey Data Typed into Dataset

A (ID)	B (Q1i)	C (Q1ii)	D (Q1iii)	E (Q2)	F (Q3)	G (Q4)	H (Q5)	I (Q6)
001	4	5	3	3	2	3	1	3
002	5	4	5	4	4	4	3	1
003	2	3	1	2	2	2	2	2
...
320	4	2	2	5	3	1	3	4

Checking Your Work

Many errors in statistical analysis are often the result of mistaken data entry. Methods to reduce these common errors are:

- Make sure your surveys are “cleaned” and have IDs coded before data entry
- Double-check data entry (e.g., one staff member reads data while another checks data on the computer screen)
- Do random spot-checks of data
- When reading data printouts, always check the number of cases to make sure all of the data are being read and all the values fall into the appropriate ranges
- Take breaks from data entry and analysis – it’s good for your eyes, head and hands.

Statistical Analysis of Data

After all of your data are entered into the selected software program, analysis can begin. The complexity of this task will vary based on needs. In most cases, you will be producing basic, summary information about the characteristics and outcomes of your project’s participants, customers, merchants, etc. Statistics are an important tool to help describe your data and confirm whether the differences found from year to year, one type of participant to another, one type of program to another, or from program start to finish are significant or merely due to chance. You commonly will use descriptive statistics such as means and frequencies (including percentages) to present your data. In addition, if you are interested, you can use more advanced or “inferential” statistics to test hypotheses and look for statistically significant trends.

Descriptive Statistics

Descriptive statistics are numerical descriptions used to summarize a larger mass of data. Funders will not be interested in looking over all 300 of your participant satisfaction instruments; they will instead be interested in a general picture, for example the % of participants who were very satisfied, or your average participant satisfaction rating. Typical descriptive statistics and the interpretation of each example are included in the table on the following page.

Inferential Statistics

Many times the purpose of calculating numbers goes beyond the description of the dataset. You may want to test a hypothesis that one group of participants is gaining more from your program than another group (e.g., males versus females), that one educational mode is working better than another (e.g., individual versus group settings), or that your program results have improved over time. Statistical tests will help you determine if these apparent differences have occurred simply by chance alone. For more information on these types of tests and statistical significance, refer to Appendix X. For more detailed information on either of these types of information, please refer to one of the following texts:

- Phillips, John L., 1996. How to Think About Statistics. New York, NY: W.H. Freeman and Company.
- McClave, James T. and Deitrich, Frank H. 1985. Statistics. San Francisco: Dellen Publishing Company.
- Huck, Schuyler W. 2000. Reading Research and Statistics. Addison Wesley Longman.

Definitions of Statistical Terms

Term	Definition	Description	Example
Frequency	The percent of scores falling into each response category.	This is the most basic statistic; it provides a proportional breakdown of responses indicated for each question.	Participant Sex/Gender Female 75% Male 25% Total 100%
Mean	The sum of all scores divided by the number of scores summed.	The mean is often referred to as “average,” and in a balanced dataset is the best estimate of central tendency. (Note: a mean can be influenced by a value that is quite different from most other values. In the example shown in the next column, if the lowest score of 50 had been only 20, the mean score would be 74.)	Mean Participant Satisfaction Score (scale = 0-100): Five participants with the following scores- 80, 90, 95, 50, 85 Mean score = $80 + 90 + 95 + 50 + 85 = 400 / 5 = 80$
Median or 50th percentile	The score that is halfway between the lowest and highest value when all the scores are listed in ascending order.	Another measure of central tendency, the median describes which measurement falls in the middle of the dataset. It can be a better measure of central tendency than the mean if some atypically high or low score influences the data. (Note: if the score of 50 had been only 20, the median would still be 85, showing that the median is less vulnerable to extreme values than the mean.)	Median Participant Satisfaction Score (scale= 1-100): Five participants with the following scores- 80, 90, 95, 50, 85 Sorted in ascending order: 50, 80, 85, 90, 95 The median (middle score) = 85 Or 50, 80, 85, 90, 95
Percentile	The percent of scores that fall below a given score.	Percentiles provide more descriptive data than just the mean or median alone can provide, by demonstrating a score's relative standing to the rest of the dataset.	If the 10th percentile of students' test scores is 60, 10 percent of the students scored 60 or less while 90% scored higher than 60
Standard deviation	The square root of the sum of all squared deviations around the mean divided by the number of deviations summed minus one.	The standard deviation describes how dissimilar or similar the scores are or how closely they cluster around the mean. It's a measure of the spread of scores.	Standard Deviation Participant Satisfaction Score: Five participants with the following scores- 80, 90, 95, 50, 85; Mean = 80 $(80-80)^2 + (90-80)^2 + (95-80)^2 + (50-80)^2 + (85-80)^2 = 1250$ $1250 / 4$ (number of scores -1) = 312.5 $\sqrt{312.5} = 17.68$

Interpretation of Statistics

Statistic	Example (from previous page)	Interpretation
Frequency	Participant Gender Female 75% Male 25% Total 100%	75% of our participants were female; 25% were male Three-fourths of our participants were female, while the other fourth were male. Three times more females than males participated in the program.
Mean	Mean Score on Gardening Skills Inventory (scale = 1-100): Five participants with the following scores-100, 90, 95, 70, 95 Mean score = $100 + 90 + 95 + 70 + 95 = 450$ $450/5 \text{ scores} = 90$	The average Gardening Skills Inventory score was 90 (range 0-100). Gardening knowledge was high, with an average score of 90 out of 100.
Median or 50th percentile	Median Score on a Gardening Skills Inventory (scale 0-100): Five participants with the following scores-100, 90, 95, 70, 95 Sorted in ascending order: 70, 90, 95, 95, 100 The median (middle score) = 95	After the program, half of all participants scored 95 or above on the gardening skills inventory.
Percentile	At program start, the 10th percentile of students' test scores is 60.	Ten percent of the students scored 60 or less, while 90% scored higher than 60. Note that the median (the 50th percentile) in this example would be higher than 60.
Standard deviation	Standard Deviation Participant Satisfaction Score: Five participants with the following scores-80, 90, 95, 50, 85; Mean = 80 $(80-80)^2 + (90-80)^2 + (95-80)^2 + (50-80)^2 + (85-80)^2 = 1250$ $1250/4 \text{ (number of scores -1)} = 312.5$ $\sqrt{312.5} = 17.68$	Compared to scores among participants citywide, our participants' satisfaction scores were much more diverse (17.68 program participants vs. 12.5 citywide).

Qualitative Data Analysis

Notes or records from observations, interviews and focus groups represent examples of qualitative data. There are a variety of methods that can be used to analyze qualitative data, many of which fall outside the scope of this text because they are quite elaborate and require significant staff time and resource. In this text we cover three primary methods to synthesize qualitative data: classifying data, coding data and using “composite” descriptions of data. In all cases the goal is to reduce the volume of data without losing key information.

Classifying Your Data

This task involves placing selected responses into a series of categories. The analyst reviews the written documents or transcripts of spoken data (e.g., audiotape of focus group session) and categorizes the pieces of text to represent either important concepts, common patterns among participants or distinct responses by different population subgroups.⁸⁸ The analyst creates names or labels for the categories that express the general theme that each response in that category has in common. If more than one analyst categorizes such themes, a key should be developed that clarifies the categorization being used. For example, for the open-ended question, “What parts of this program were the most meaningful?” the responses might be categorized in the following way:

Example of Classifying as Set of Qualitative Data

(Bold headers are the labels you create for the categories into which you place the verbatim responses listed.)

Time Spent with Program Staff

- I liked hanging out with Antonio. He always listened to me.
- The mentors. They are really cool.
- Claudia taught me a lot.
- Ty made me feel important.

Peer Relationships

- The other kids!
- I made many friends this summer.
- Learning that there are other people like me
- Hanging out with other kids in the gardens
- Meeting new people
- Spending time with kids who don't care what kind of shoes you have
- Johnny and Roberto

Income/Employment Opportunities

- Getting paid for working in the garden
- The money
- \$\$\$\$\$
- Spending time doing something different than hanging out this summer
- It's a better job than McDonalds or Taco Hell.
- The job is fun. I liked coming every day.

New Knowledge and Skills

- Learning how to garden
- Learning new things about plants
- Feeling confident that I can do something well
- Understanding how to lose weight by eating more fruits and vegetables

Other

- The food
- Getting outside of the city
- Feeling better about my future
- Getting away from my family problems

Categories can be defined either before the data are collected (“predetermined categories”) or after reading through the raw data (“emergent categories”). Less frequent responses are placed in the “other” category. One may choose predetermined categories when prior research provides such groupings or when such categories can be preset using the evaluator’s practical experience with the group. Emergent categories are used when there are no clear expectations for responses and/or when the respondent population’s diversity compels one to develop the categories in a more formative manner.

Coding your Data

A step beyond classifying your data by topic is to assign numeric codes for each of the responses in the same category. So in the example above, each response falling into the Program Staff category would be assigned a “1,” all responses falling into the Peer Relationships category would be assigned a “2” and so on. You may want to code responses for two reasons: 1) most software programs prefer numbers to text – it speeds processing time and allows a greater number of computations, and 2) coding qualitative information will allow the data to be used in quantitative ways.

Example of Codes for an Open-ended Question

Question 5: “What parts of the program were the most meaningful?”

- 1=Time spent with program staff
 - 2=Peer relationships and support, friends, buddies, etc.
 - 3=Financial aid/employment opportunities
 - 4=New knowledge and skills
 - 5=other
-

The responses can then be analyzed and presented in the following manner: (“N” equals the number of participants that responded.) In the example below, each respondent gave one response (the “# of Participants per Response” column totals 25). However, a participant may provide multiple responses that can be coded into several response categories. In this case, the “# of Participants per Response” would equal greater than 25 and the “Percents of Participants” would total more than 100%.

Example of the Presentation of Results of a Coded Open-ended Question: Components of the Program Most Meaningful to Participants

Response	Code	# of Participants per Response	Percent of Participants (N=25)
Peer relationships and support	2	7	28%
Financial aid	3	6	24%
Time spent with staff	1	4	16%
New knowledge and skills	4	4	16%
Other	5	4	16%
Total		25	100%

Consistency in Coding

Consistency in coding is very important. In the coding example above, individuals could be unsure whether to code a response like “my counselor became a very good friend” in code one (time spent with staff) or two (peer relationships). If more than one staff member or volunteer will be applying codes to your questionnaires, be sure to train them all on the codes and their meaning. Establishing a consensus about what answers will be included in each code is fundamental to making coding credible. Once this consensus is reached, record it in an evaluation notebook and place examples of each coded item into the appropriate spot in the notebook. This way, the instructions for coding, with real life examples, can be used consistently over the years.

Using Composite Descriptions to Synthesize Your Data

A “composite” response is a statement that generally describes the responses in each of the categories you have developed. This composite can sometimes take the form of a quote from a specific respondent that embodies the responses in the dataset. For example, a composite statement for the category of Income/Employment Opportunities might be:

When asked the most meaningful part of the program, a number of the youth reported the opportunity to earn money and to work at a more interesting job.

You can use a series of composite statements to create a summary paragraph if more explanation is necessary.

Enhancing Data with Direct Qualitative Statements

Many funding reports may be strengthened by the use of testimonials from participants. Synthesized data can be complimented by adding meaningful comments from participants. Direct quotes can also be added to a report to provide an example of a coded category or to highlight information collected.

Analysis of Secondary Data

In some cases, there may be data collected by others that may be useful in your analysis either as evidence of impact or to provide context for your program. For example, if you have worked with a community market to purchase food from local farmers, you might get records of the market’s produce sales for a specific time period. Other examples of secondary data might come from the US Census, Food Security Supplement to the Current Population Survey, U.S. Department of Commerce, or the National Food and Nutrition Survey (NFNS). (See the USDA Community Food Security Assessment Toolkit for more information on other helpful datasets.)

When using data collected by others, it is important to consider the primary use of the data and make sure it fits with your purposes. Resist the temptation to use data just

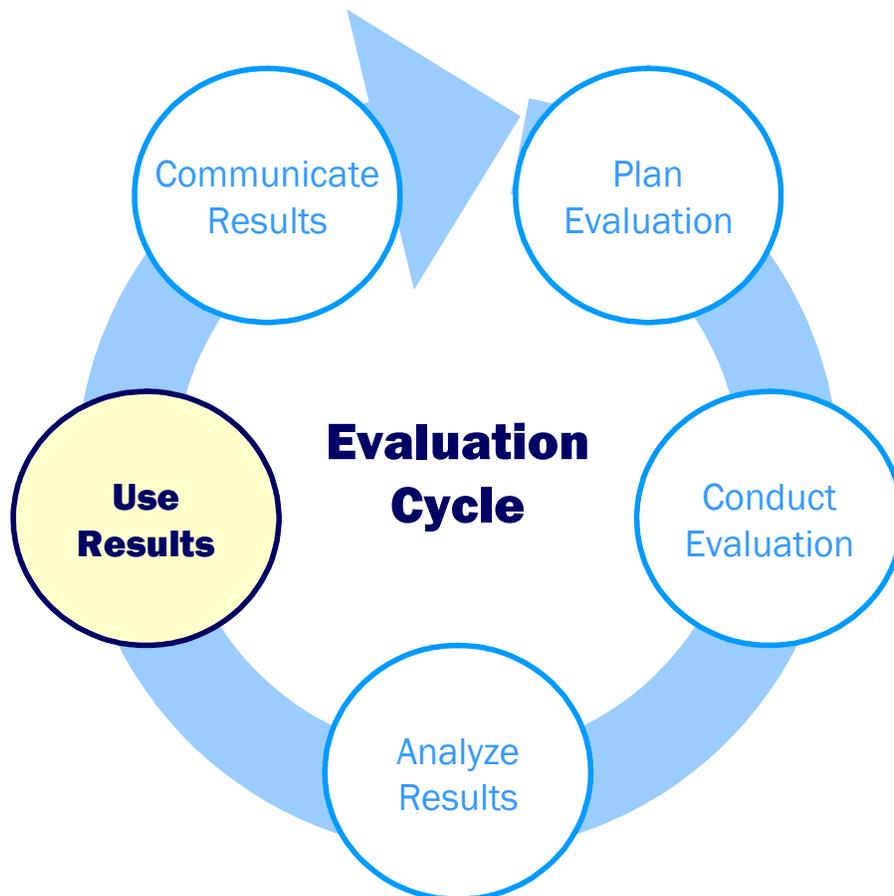
because it is there. If the data are not meaningful or accurate you will be doing yourself a disservice to rely upon it for your evaluation. For example, don't use records of low birth weight to demonstrate poor nutrition when the National Food and Nutrition Survey data provides more targeted data.

A second caution is to make sure that the population base of the secondary data match with your population base. For example, if you have a program goal to reduce obesity in a specific community, relying on obesity statistics for an entire county or region of a state may underestimate the impact of your local program because many who have not participated are used in the counts of secondary data for the county and region.

The modifications to secondary data will vary based on the source. Sometimes the data already will come in a useable form (e.g., Census Data). Other times you may need to manipulate the data to either create summary statistics or to pull out population subgroups more useful in your evaluation (e.g., Number of children enrolled in Free or Reduced lunch programs in your school rather than district-wide estimates). Make sure to always find out if you can obtain the secondary dataset electronically, because it is often through these manipulations that more useful information is found.

Chapter 9. Understanding and Using Your Results

- 🌸 Understanding Your Data
- 🌸 Using Your Results for Program Improvement



Chapter 9. Understanding and Using Your Results⁸⁹⁹⁰⁹¹⁹²⁹³

Understanding Your Data

Once the data are collected and analyzed, staff need to understand the data and what story the results tell. Then staff can develop a plan for the use and dissemination of the results.

Making Sense of the Results

The first step to understanding your analyzed data involves looking at your results and making sure the numbers are consistent with your knowledge of the program. There are a few key questions you should ask yourself.

- Do the numbers make sense to me?

Do the data show that 100% of participants are male when in fact you know that a number of your participants are female? Is the average number of market visitors 200, but you know that the daily totals are generally over 300?

Always test your results against your own experience and when the data are way off from what you understand or believe, start checking around. The sample of information on which the data are based may be skewed. A wrong number may have been entered into your database. The report may be mislabeled so that you are reading about the average number of new market visitors, rather than the total number of market visitors.

- What do the data tell you about the participant population and how has that changed since last year (or month or quarter)?

Is the intern program attracting younger teens? Does the market serve a large number of WIC recipients? How many co-op members are over the age of 55?

Answers to these questions may shed light on changes in the outcomes you observe.

- How have your activities changed since last year (or month or quarter)?

Are you now providing additional nutrition counseling at the schools? Have you added a job-training component to the youth garden project? Have you teamed with a senior center to provide food in congregate meal settings?

Answers to these questions, too, may shed light on changes in the outcomes you observe.

The Use of Comparison Data

Once you feel that you can trust the data and have a good handle on how the participant population and activities have changed, you want to determine if participant outcomes (e.g., increased knowledge of organic gardening practices) are improving. Improving participant outcomes, by themselves, will not provide indisputable scientific proof that your program deserves the credit for change. However, improving participant outcomes can offer compelling evidence of program effectiveness.

If a project is unable to show evidence of improving its participant outcomes, project staff may need to examine whether any external factors are putting downward pressure on participants' abilities to improve (e.g., greater economic hardship, less community support or less service provided because of budget cutbacks). Many circumstances may contribute to making participant outcomes poorer over time. Perceived stability of participant outcomes from year to year and even small declines in participant outcomes from year to year can be mistaken for program failure.

The best way to determine if a program is working is by comparison; by comparing program outcomes over time. A program can compare their results to results of other like programs (organizations in the same geographic vicinity or those serving similar populations in other parts of the county, state or nation). Compare program outcomes to outcomes from norms developed by the authors of the instruments you use to measure outcomes.

It may be quite difficult for some unique or unusual programs to find similar programs with which to compare results. For these programs, it may be especially important to compare their own outcomes over time.

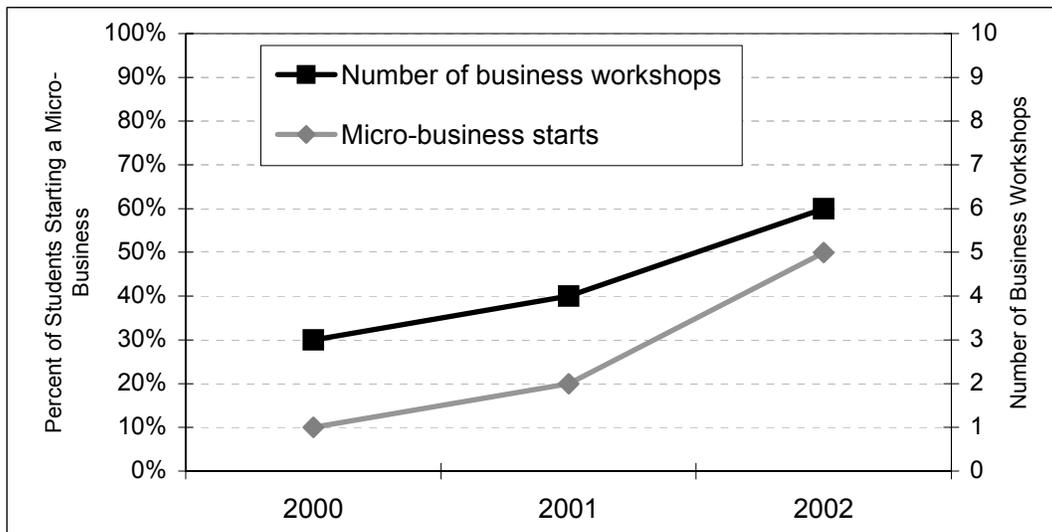
Comparing Your Own Outcomes over Time

A single snapshot of participant status will be useful as a descriptive tool, like the following example.

“Our participants averaged 16 years old and participated in the programs two days a week for four hours per day on average. Over 70 percent lived at home. Over 60 percent of those in the program for at least six months successfully started their own agriculture micro-business. Ninety-two percent stayed in the program for over six months (leaving a drop-out rate of only eight percent).”

Data about program activities and outcomes over time will make the case for program success. For example, data from the last three years of the program show that the proportion of participants starting micro-businesses increased while at the same time new training was added to teach small business management and entrepreneurial skills. The graph on the following page demonstrates this success.

The data show that as more business training was phased in, youth were more likely to successfully start a new micro-business. The link of service enhancement to better outcomes - makes the strongest case that the program is responsible for participant improvement.



Evaluation data becomes more powerful as it is compared to itself over time. Sequential data may reflect program improvement efforts and allows stakeholders to see progress being made year after year.

Comparison to Other Programs (Networking and Benchmarking)

At a presentation to the school board, members were told that students rated the quality of the salad bar on average as 75 on a 100-point scale, where one was “Very Bad” and 100 was “Very Good.” One board member asked of the rating, “But is that good? How do kids usually rate food?”

One of the best ways to make sense of data that emerge from your evaluation process is to have data from other CFPs to which you can compare your own. How do we know what is tall or what is small unless we compare? Without other data about typical participant ratings, it is impossible to say whether “Good” on the scale should be considered good enough

The truth is that most participants will give most services pretty high ratings. But how participants rate other community projects should provide a dose of reality about what is typical of participants in general and what is unique about your own participants. The same is true for rates of increased food security for a neighborhood with a new garden, the rate of job acquisition after job training and the increase in healthful eating or improvement in the general quality of participant life after participation in food project activities.

Evaluation networks can be developed among similar projects, for example, those that serve participants in the same geographic area (U.S., state, county, city, neighborhood); those that serve the same type of participant (youth, recent immigrants, homeless); those that offer the same type of service (Community Supported Agriculture Projects farm to cafeteria programs, farmers' markets). Evaluation networks can be as broad as all CFPs offered to all participants in the U.S. or as narrow as food service job preparation programs offered to 16-year-old youth who dropped out of school.

Many CFPs belong to organizations such as CFSC or have contacts with staff in programs in other parts of their community or in other states. The participants of CFPs share common characteristics despite their more obvious differences. The contacts, local and elsewhere, should be seen as valuable, yet untapped, resources. A network of projects - even if it is only two such projects - aimed at sharing information about what works in community food projects as well as funding, can expand the power of any evaluation done at one location alone.

As networks grow, so grow the opportunities to identify the best practices among those organizations with the best outcomes. These best practices can become service benchmarks against which local projects compare themselves. The outcomes of a network of projects can provide empirical data to help set performance standards. The numbers - like the percent of students changing their eating habits or the average sales at a farmers' market - are the quantitative goals staff aim to achieve.

Furthermore, programs that participate in an evaluation network will benefit from learning about the impact of good outcomes on fundraising. Also a network with comparable evaluation and service data can build a database powerful enough to fuel meaningful research on what works best in a variety of locations and with a variety of participants and types of service delivery.

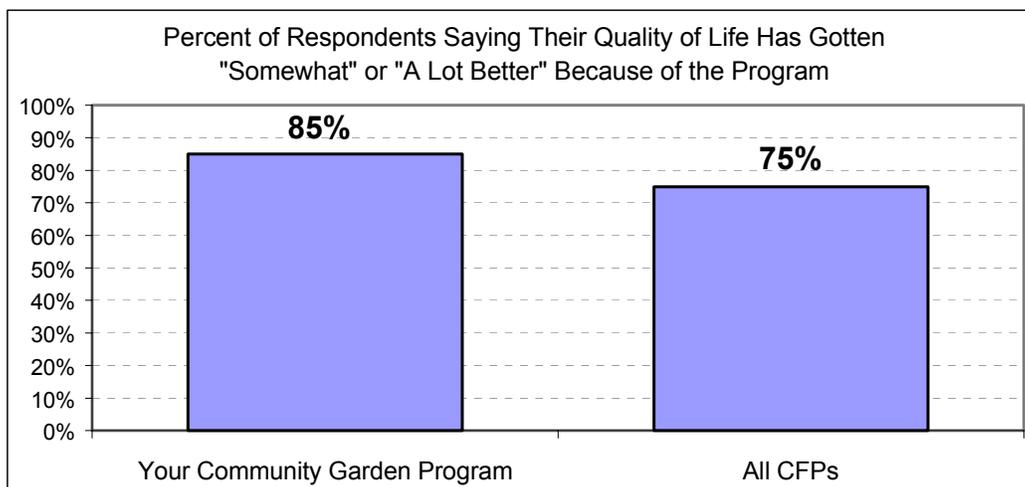
Comparison Data

It is often difficult to find a program similar enough to your program that offers outcome data that can be credibly compared to your outcome data. For example, is it acceptable to compare a youth garden project within a school to one that is run by a church? Can we compare outcomes of a new farmer training program in Washington, D.C. that emphasizes traditional food production and community integration of participants to one in Seattle that emphasizes the use of organic farming techniques? Or can we compare results of the youth gardens to the new farmer training programs?

The old saying, “you can’t compare apples to oranges,” is often invoked to defame comparisons of seemingly different programs. But of course, you can and you do. Apples and oranges are about the same size, but oranges provide more vitamin C per serving than do apples. Both apples and oranges have more in common, and more to offer the vitamin C seeker, than potatoes. And for the camper, the apple clearly is superior because it requires little juice clean up after consumption. The further back you are willing to step from the objects to be compared, the easier it is to notice the similarities and to overlook the smaller things that would diminish your comfort in comparing. Compare what any parent of twins says about their kids to what the world at large says. Parents are so close to the children that they see all the distinguishing features; the rest of the world sees two identical people.

An Example of the Use of Comparison Data

Imagine that a community garden and a number of other community food projects have asked their participants two questions: “To what extent has your quality of life improved (if at all) since you became a participant in this program?” and “To what extent do you credit (if at all) this program for the change in your overall quality of life since you became a participant?”



Results from this fictitious example show that 85% of participants felt the community garden program had made their quality of life better, while the average across all projects (say there were 10 of them) was 75%. Certainly some of this difference may be due to participant expectations and the kind of service being provided, but when

participants' perceptions are so positive (for whatever reason) a program can feel confident that participants believe the help they are receiving is effective. These kinds of comparisons help build a case that a program's interventions are working.

If results were not so beneficial - showing, for example, 55% for your program compared to 65% for the average from the evaluation network - that would send a signal to program staff to investigate further and to discover how activities might be changed so that participants may have a more positive perception of the program.

Take the Good with the Bad

Don't ignore outcomes that fall short. Although it is certainly more fun and rewarding to look at the successes (and they should be celebrated), the negative results are at least equally important. It would be one's natural tendency to focus on the positive results and ignore the negative. However the areas where you failed to achieve your goals are opportunities for education - a challenge to your staff and organization to serve your participants better. Chapter 9 of this handbook presents ways of applying your data analysis learnings to program improvement.

Using Your Results for Program Improvement

Once you have determined if results point to areas needing improvement or if most outcomes are so positive you need only steer a straight course, you still should wait before telling the world. It often makes sense to get results to staff before they go to other stakeholders because staff often have more direct contact with participants and it is those staff whose activities have the greatest day to day impact on participants. With some community food projects, volunteers and board members may share an active role along with staff regarding the day-to-day activities and interactions with participants. In these settings, it may be appropriate to share results with this combined group in order to learn from results for program improvement. We

Cautionary Note About Comparison Data (To Others or Yourself Over Time)

Be careful not to over-interpret small differences. There is a certain amount of variation that occurs in data simply by chance. The information you collect will, in most cases, confirm your own intuition and knowledge gained from day to day program work.

Just as it's prudent not to change your eating habits radically based on one publicized finding, don't jump to redistribute resources or change your program's direction based on small differences in new data, especially if you are suspicious about the results.

Look for more information about the issue by doing further investigation of your data. You may also want to check the research literature or another similar program to see if your findings are supported or not.

Over the long haul, consider augmenting your questions to participants to include more data collection on the topic in question.

We recommend that any report about program evaluation include a section that explains how staff plan to address the findings. What will you do to keep the program healthy? What will you do to improve the program where improvement is needed? What will be

enhanced? What will be replaced? How will this affect program results? Why will this help participants?

This is the toughest part of evaluation. There are no simple formulas for making the connections between what you find and what you do. This is where staff and board experience and training will make the greatest difference. See the results review and dissemination cycle on the following page. Even the most experienced evaluations are better at telling you what is happening than why it is happening. Although a strong evaluation will provide some clues about the reasons for better or worse results, it is from staff that the best answers are likely to come. Staff are closest to the program and will have useful ideas to explain your evaluation findings. Just as a blood test is limited to telling you the levels of nutrients and minerals in your system, the evaluation results must await your doctor's (read staff) expert opinion about how to diagnose or make recommendations to improve your health.

A few evaluation principles can help program managers assure that the entire evaluation enterprise does not end with only, "That was interesting." Important activities to help integrate the results into your program relate to communicating and reviewing results with staff and developing an action plan.

Debriefing Evaluation Findings with Staff

To ensure evaluation results are used within your organization, communicate the findings with staff and at the same time, involve staff in the process of making sense of the data, brainstorming potential actions to take based on the findings. Staff are more likely to use results if they understand and have ownership of the evaluation process. However, before beginning this process, it is important to set the stage for this process to occur: to ensure that these results are disseminated and discussed in an environment of learning, not judging.

Promoting a Learning Environment

The W.K. Kellogg Foundation cites the following barriers to the use of evaluation results by staff:

- Fear of being judged
- Concern about time and effort involved
- Resistance to change
- Dysfunctional communication and information sharing systems
- Un-empowered staff⁹⁴

In Chapter 4 of this handbook, we provided information on the importance of establishing a "culture of inquiry" or environment where evaluation data are valued rather than feared. As evaluation data are released to staff, the principles of a learning environment are most important, as staff may feel nervous about less than positive results. It is very important that staff do not feel singled out and that all feel

empowered to work together as a team to acknowledge the good work accomplished as well as solve future challenges together. Staff may need to be reminded that they participated in the conception and implementation of the evaluation. (See Chapter 4 for more information on setting the environment for evaluation.)

For larger programs or ones where there are serious challenges to the learning environment for one reason or another, one might consider sending results of the evaluation system to front line staff only, without managers being able to see those results, if that reinforces a “no threat” approach.

Beginning the discussion

A debriefing of evaluation results with program staff might best begin with a group process whereby staff are given a copy of the evaluation results,⁹⁵ (including any comparison information that has been collected) and asked a series of questions:

- What parts of these findings did you expect?
- Was there information that was surprising? What was it and why?
- Do you still have unanswered questions, or areas you thought you would know more about after the evaluation?
- Is there additional information we need to collect to better understand these data or take action?
- Are there changes we should make to the evaluation system as a result of these data?
- What do these results mean for us/ our program?
- What are we doing well? Where are our participants or community seeing the most positive impact?
- Are there changes we should make to the program as a result of these data?
- Do we need to take action on these data? Why? What actions?

Depending on the complexity of the data results, you may consider giving staff the information ahead of time for review prior to the group discussion. In addition, consider that it might take multiple meetings to review the data as a group, discuss the above listed questions and create a plan for action.

If your program is too large to involve all staff in these types of discussion, select a number of staff (from various levels of the organization) to work on a team through this process. This team can then present its recommendations to all staff for review.

You may also want to break the discussion into several sessions. Staff may need time to digest the results and think through their meaning. Perhaps at the first meeting, only the first two points suggested above are discussed. Participants are then instructed to give more thought to the next questions before a next meeting. The next two main

points could be covered at one or two sessions. Additional discussion may be needed to create an action plan.

Creating an Action Plan

As every community food project is unique—having its own goals and objectives, methods of achieving those objectives, performance standards, community and cultural setting and, of course, evaluation results—the action plan that is developed must also be unique. The evaluation action plan and the process used to develop and implement it can be as brief or extensive as necessary. However, it is critical that an action plan be considered, so that the evaluation results are integrated into program operations and used to help improve service delivery and community impact. Without an action plan, programs may sit with their data and yet have little direction for making program improvements. An action plan provides this direction and helps staff focus in on the road to improvement.

One way to develop an evaluation action plan is to integrate it with the logic model and participant/client/customer evaluation measures. On page 42 of Chapter 3, examples were shown of outcomes with their related indicator or measure and performance standard. A table of this type could be extended to contain three more elements to become an action plan (see the example on the next page). These elements are:

The criteria needed to judge change

- This may be the same as the performance standard, or you may choose to set it a bit lower than the performance standard, with the understanding that no action is needed until the performance is somewhat below the standard.

Procedures for implementing change

- The actions that will be taken to make a change, when it is deemed necessary.

Timeline

- How often a review of the performance standards will be undertaken. This will mostly be driven by the frequency with which you collect the evaluation data to be reviewed.

This type of approach can be used whether your evaluation shows positive or negative results. The discussion with staff should spotlight program successes, the factors that contribute to those successes and how to build on those strengths. Highlighting the positive with staff will help create buy in to the review process and likely make staff more eager to participating in the action plan for improvement. Even if the negative results are grave, staff need the opportunity to celebrate and feel good about their efforts. Then, as a group they can move forward.

You should also keep the “big picture” in mind. It may be easy to concentrate on the few narrow outcome indicators, but some time should be spent on consideration of the fundamental mission and impact of your program. What is the overall story your results tell? Do these results warrant any bigger changes in the purpose or methods of your program? Is a shift in priorities or resources warranted? Or do these results confirm the suitability of your activities as is, to meet the community’s needs?

Sample Plan to Use Evaluation Results⁹⁶

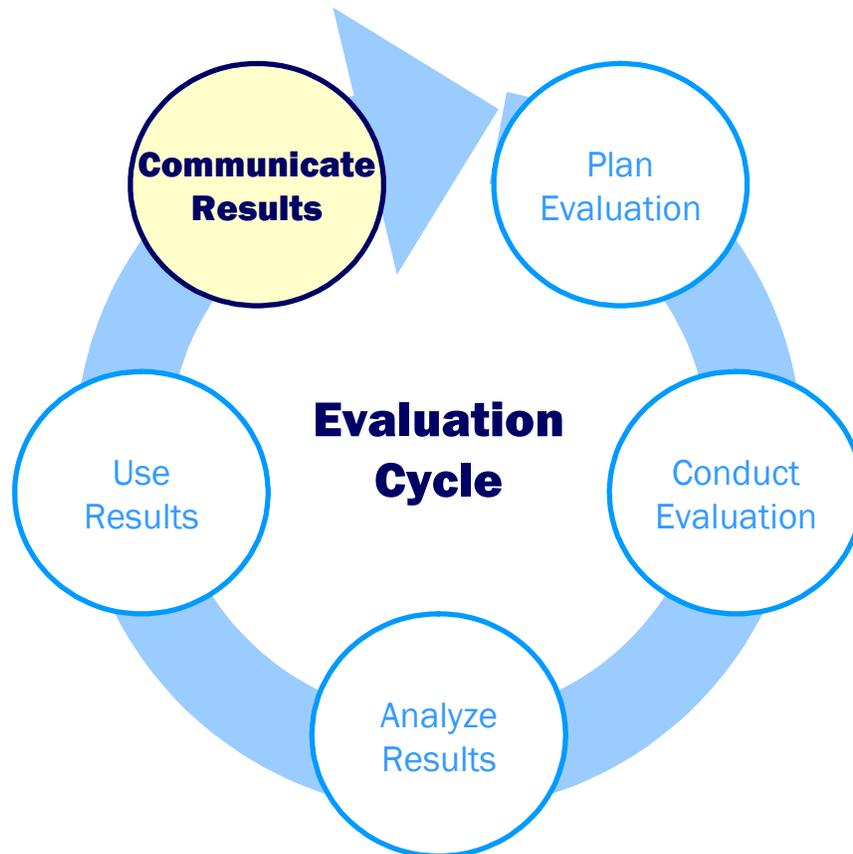
Project Activity or Program Goal	Outcome	Measure or Indicator	Performance Standard	Evaluation Results	Action Plan	Timeline/Staff
To increase gardening skills for participants	Increased knowledge of gardening practices	Knowledge score on garden skill inventory	On average, a 50% increase in knowledge in a pre/post test	Average Change in knowledge = 52%	No changes necessary	Not applicable
Latino residents will report greater satisfaction with produce offered at neighborhood farmer's market	Satisfaction with food selection	Average rating on market intercept satisfaction rating scale	At least 80% of Latino shoppers will report being "very satisfied" with the variety of food choices provided at market	63% of shoppers were very satisfied	Hold focus group with sub-set of shoppers to determine methods to increase satisfaction	Next month -Jose will recruit -Lydia will host
To increase civic responsibility	Community service	Number of hours spent in community service	90% of youth will participate in at least 20 hours of community service per semester	60% of youth contributed 20 hours	Review program curriculum to determine how to increase emphasis on community service	Complete before progress report due to funder -Margo and John
To increase collaboration of food-related community organizations	Increased collaboration	Number of organizations participating in food network	8 community-based organizations will attend monthly network meetings	10 organizations joined network, but 4 are infrequent participants	Call 4 organizations and inquire more about barriers/incentives to participation	This month -Malaika

Worksheet 8: Plan to Use Evaluation Results

Plan to Use Evaluation Results						
Project Activity or Program Goal	Outcome	Measure or Indicator	Performance Standard	Evaluation Results	Action Plan	Timeline/Staff

Chapter 10. Communicating Results

- 🌸 Determining Your Audiences
- 🌸 The Basic Evaluation Report
- 🌸 Effective Methods for Data Presentation and Dissemination
- 🌸 Attributing Causation
- 🌸 Using Your Results for Fundraising and Grant Writing



Chapter 10. Communicating Results

In addition to using evaluation results to enhance your program and manage staff and resources, your findings will also be of interest to other stakeholders, potential funders and the community at large. Community organizations, individuals and funding agencies will benefit by understanding the demographics of individuals that your program serves. They will also be interested in learning how your program is affecting these people's lives.

On par with communicating results to stakeholders, funders and the community, is the significance of communicating results to the individuals who themselves participated in the evaluation. Throughout history, groups of people have been studied and studied and studied again, with no knowledge of "what happened" and no perceived change as a result of their participation. Of particular emphasis on this point are people who are vulnerable, of lower economic status and racial/ethnic minorities. Too often these populations have been "left out" of the communication stream, even though they were the ones with the most significant investment. Evaluators owe their participants this follow-through and, by doing so, provide future evaluators with the same opportunity.

Whether you report results verbally or in writing to your board, the USDA, the public, and/or other funders, be specific about the rigor of your data collection procedures. The evaluation system itself, like a good financial accounting system often instills confidence in the reader or listener. The fact that your program has taken the time and energy to collect evaluation data and that your measurement system includes staff review of results will help convince others that you accept responsibility for the funds given, you are continually striving to improve your performance and you are serious about increasing the well-being of your participants.

Determining Your Audiences

Before beginning data dissemination, it is important to consider the various audiences you will have for your data and communication modes and styles that will be most appropriate and meaningful for each. Common audiences for community food project data are:

- Board Members
- Participants (including program participants, family members and volunteers)
- Policymakers
- Community groups and leaders
- Funders (current and potential)
- Community members

When thinking about each audience, you will want to consider the reader's relationship to and knowledge of your program as well as the reader's primary area of concern. The closer the relationship of the reader to your program, the more likely he or she will be interested in more details of the evaluation findings. Those audiences with less stake or relationship will look for more synthesized, "big picture" data presentation. The amount of data provided and the way in which it is presented will depend on the particular audience being addressed. It will be helpful to consult with stakeholders and CFP colleagues as you gauge how to provide the right message to different audiences.

The Basic Evaluation Report

Because your audiences' interests may vary, it is important to consider a number of different methods to present and disseminate information. The foundation for all of these communications is the Basic Evaluation Report; a detailed and comprehensive review of the evaluation process and findings. We recommend you create this report before engaging in other evaluation dissemination activities because it will give you a chance to document and thoroughly review the data in a way that will ensure that you bring forth the most important data for each additional communication and audience. The structure of a Basic Evaluation Report is detailed below. Appendix XI provides a sample Executive Summary from a fictitious report for a farmers' market evaluation. This fictitious example is based on the results from one data source: a survey of farmers operating stands at the market. Evaluation reports often present data from multiple evaluation sources. Appendix XII provides a sample annotated evaluation tool that would be included with the Executive Summary in an evaluation report. This annotated evaluation tool is based on the same fictitious farmers' market evaluation.

Basic Evaluation Report Template

Project X Evaluation Report

Executive Summary: This is the first portion of the report, and may be the only portion of the report some audiences ever read. Thus, the information presented should be that of most importance. In addition, information should be concise (many use bullets) and non-technical. Many argue that this summary is best written after the rest of the report is complete.

Key information to include in the Executive Summary (usually 2-4 pages in length):

- Purpose of evaluation (one or two sentences)
- Methods (one or two sentences)
- Major findings (1/2 to 2 pages)
- Conclusions and recommendations (1/2 to 2 pages)

Introduction and Background: This section of the report provides information on the program and the purpose of the evaluation. It also provides a brief set of information about the evaluation methods. A more thorough methodology section should be included as an appendix.

Key information to include in the Introduction and Background (usually ½ to 1 page in length):

- Program history and operations
- National or other research that underpins the agency philosophy and delivery of service (as possible)
- Program goals and objectives
- Purpose of evaluation
- Methods used for evaluation
 - Evaluation method (survey, interviews, focus group, document review)
 - Number of participants in evaluation
 - Response rate (if applicable)

Evaluation Results: The results section is the heart of the evaluation report. It presents all of the evaluation data included in the document. Data are presented in tables and graphs along with interpretative text.

Key information to include in the Results section:

- Participant information (comparing over time as possible)
- Number of participants
- Participant demographic characteristics
- Participant risk and protective factors

B. Activity information (comparing over time as possible)

- Number of units of service
- Types of activities provided
- Program attendance

C. Outcomes (comparing over time as possible)

- Descriptive statistics for all outcome data (frequencies, means, etc.)
- Comparison of outcomes to performance indicators
- Comparison of outcomes to other programs or norms as possible

Conclusions and Recommendations: The conclusions and recommendations section is the portion of the report where important findings are summarized and plans for action steps are outlined. It is also a time to explain why differences between expected and actual results may have occurred. It may also be a good point to mention any significant external factors that negatively contributed to your results. If you fell short – do you plan to modify activities, if so, how – if not, why not? If you met or exceeded expectations, what decisions will you make? Will you maintain or expand activities? Make sure recommendations made are feasible for your program.

Appendices: The appendix is the place to include important information that may be too technical or detailed for the body of the report.

Key information to include as Appendices:

Methodology: Although a few summary sentences have been made about the methodology in other sections of the report, a more detailed methodology is best presented as an appendix. Categories of information often covered in the detailed methodology include tool development, selection of participants, administration methods, data analysis procedures, etc. The methodology section should include any pertinent data to establish the trustworthiness of results.

Evaluation Tool(s): A copy of the tools used in the evaluation should be provided so the reader can see how questions were asked, data were recorded and the context of the data collection. It is also helpful to enter summary data into the tool as a concise way of communicating overall findings. This is referred to as an “annotated instrument” 1 (a sample of a “annotated evaluation tool” is included on page 130).

Effective Methods for Data Presentation and Dissemination

The Basic Evaluation Report, while being the most comprehensive report of your data, may not be best suited for all stakeholder groups and audiences. While many staff, board members and volunteers may have the interest to read through an entire evaluation report, many other audiences will not. Some of the common methods used to communicate evaluation data in a more summary form are:

- Press releases
- Newsletters (to community and/or participants)
- Annual reporting to funders
- New grant proposals
- Presentations at meeting to boards, community groups, policymakers
- Web pages
- Video and audio clips

We provide examples of the first four of these in Appendix XIII. The documents are all based on the fictional farmers’ market described in the sample Executive Summary.

For all of these communication modes, there are several concepts to keep in mind: simplicity, variety and candor.

Concepts in Communicating Evaluation Results		
Simplicity	Variety	Candor
Audiences look for brevity and clarity. Although there are many analyses to present, stick to the main questions of interest to your audience.	Data are often best presented in a visual manner such as tables and graphs. Use a word processing program that has the ability to utilize these types of graphics.	Don't be afraid to report outcomes that didn't meet your goal. Give readers the context and any thoughts your staff have about why the goal was not met.
Write short paragraphs. Text including evaluation data can become overwhelming if too much information is provided.	Use anecdotes and other stories to help illustrate more quantitative evaluation data and provide depth to your findings.	Reporting goals that were not met can also be a prime opportunity to show how you've used the information to make the adjustments and you may be able to argue for funds to support a new aspect of the program.
Avoid jargon. Do not assume that your audience members are familiar with typical concepts and terminology used in your area of specialty.	Make the communication visually interesting by adding pictures and graphics. Keep readers focused by highlighting important points with text boxes and/or different fonts but be careful not to use too many fonts that clutter the page.	Be honest – your report will be considerably more credible if you note both the strengths and weaknesses of your program ⁹⁷

Attributing Causation

Understanding What Your Evaluation Results Can and Cannot Tell You and Therefore What You Can and Cannot Say About Them

There are many methods professional evaluators and researchers use to help them identify if the program caused the change in participants rather than something else, but many of these methods are outside the scope of many non-profit organizations (see Chapter 5 for more information on randomized control trials and other evaluation designs).

Don't feel inhibited to offer your best speculation about why you think your program participants have or have not improved after receiving your program activities. If you have asked some of your participants and they have given credible answers, use a quote that makes sense from those interviews to help explain changes. Remember, we are all familiar with this kind of guesswork. It's what we do when we have gotten a cold – or gotten over a cold. We speculate how we get colds and from whom. (I got it from my daughter; I haven't been getting enough sleep; I went outside without my coat on; etc.)

For many programs, it is enough to demonstrate changes in knowledge, skills or behaviors within your participants. Further, most programs do not operate in isolation so it is unrealistic to believe that your program's effects can be attributed solely to the activities you provide.⁹⁸

Some methods that might work towards establishing compelling evidence that your program is responsible for participant changes follow:

- Compare your data with like programs or norms from national or regional data.
- Compare your data over time, showing linkages to service enhancements. If the proximity of participant changes is closely tied to programmatic changes, a stronger case is made.
- Present outcomes for participants getting different amounts of service. The case for program impacts is strongest when outcomes are better for participants who have been in the program longer or who have received more intense service even over a shorter period of time.
- Refer to national evaluation data that have proven the impact of the kind of service you offer.

Using Your Results for Fundraising and Grant Writing

In addition to using your results to enhance your program and manage staff and resources, your results are also of interest to funders. In fact, for many agencies it is a

funder’s emphasis on accountability that prompted the program to gather data and analyze results in the first place.

One of the most important uses of evaluation data is to provide a persuasive presentation to groups and organizations that might provide financial support for projects and programs. Because different audiences have different information needs, the reporting of evaluation data should be tailored to the interests of the targeted stakeholders.

Proposals for funding should present data concisely and tie the results to the vision, mission or goals of the grantor. If the goals of the grantor are unknown, the results of the evaluation should emphasize the value of the project and suggest a compelling case for the need for continued support.

Although most managers are skilled in writing grant proposals, solicitations for funds and reports to funding agencies, the use of statistics generated by the organization in a report may be new terrain. To demonstrate to your audience that you understand the data presented, it is important to refer to your data correctly. Some examples of the do’s and don’ts of data presentation are described in the table on the following page.

Hints for Better Data Interpretation		
Example	Better Example	Concept
Of the farmers returning surveys, 20 reported that they had increased their income from farming due to selling at the market.	Of the farmers returning surveys, 66% (20 out of 30) reported that they had increased their income from farming due to selling at the market.	Proportions are more informative than numbers. If providing a number, make sure both the numerator and denominator are included.
Of the volunteers attending the workshop, 12 of 14 reported that they found the training useful.	14 volunteers attended the training. Of these volunteers, 12 or 86% reported that they found the workshop useful.	Do the math for readers.
Most students in the program showed increases in vegetable and fruit consumption.	Almost 70% of students in the program reported eating at least four fruits and vegetables per day, an increase from two per day prior to the program start.	Be specific. Show your data.
Stated Outcome and Performance Standard: 65% of participants will graduate from high school. Reported Outcome and Performance Standard: 80% of higher income participants graduated from high school, while only 50% of lower income participants graduated.	Stated Outcome and Performance Standard: 65% of participants will graduate from high school. Reported Outcome and Performance Standard: Overall, 70% of program participants graduated from high school. Graduation rates varied by income status, however. Almost 80% of higher income participants graduated, while only 50% of lower income participants completed high school.	Promised and delivered outcomes should be reported in a comparable format.

Hints for Better Data Interpretation		
Example	Better Example	Concept
Participant scores jumped from 80 to 82.	The slight increase in participant agricultural knowledge scores from 80 to 82 was too small to conclude there was an improvement in program outcome.	Test the differences' using inferential statistics so that too much is not made of the small differences year to year in small programs.
The agricultural skills scores were great.	The scores on the agricultural skills test were all above those experienced in other farming programs. This exceeded our goals for the program.	Keep the section of the report where data are reported as neutral as possible. Do not add subjective commentary. (This should be done in the introduction or summary of the data--not while reporting the specific numbers.)

In Closing

Community food projects across the nation are including evaluation as an essential component of their programmatic initiatives. They are learning from their clients, consumers, participants, volunteers and other important stakeholders. CFPs are making programmatic adjustments and improvements to provide better services and have a greater impact on their communities. They are educating themselves and then educating their funders – past, present and future – advocating, communicating and cultivating to continue the good work in their communities.

The process of integrating evaluation into community food project work is unique for each organization. We hope the information provided in this handbook strengthens your capacity to conduct evaluations in a meaningful way.

Best of luck!

To Evaluate This Handbook:

To help the CFSC Evaluation Program continue to evaluate its efforts and improve services for community food projects, please complete the evaluation form on this *Community Food Project Evaluation Handbook*. To access and complete the evaluation form, log on to the CFSC website at www.foodsecurity.org (Programs/Evaluation Program/Evaluation Materials section).

To Learn More about CFSC Evaluation Resources:

Valuable resources are offered in the CFSC website's Evaluation Program section. To access these resources, log on to <http://www.foodsecurity.org/evaluation.html>. There you can learn more about upcoming evaluation workshops and trainings and other worthwhile evaluation resources and opportunities.

Appendix I: Program Evaluation Standards

These standards were taken from the following resource:

The Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs*, 2nd Ed. Thousand Oaks, CA: Sage Publications, Inc.

Program Evaluation Standards

Utility Standards

Utility standards are intended to ensure that an evaluation will serve the information needs of intended users.

- **Stakeholder identification:** Persons involved in or affected by the evaluation should be identified so that their needs can be addressed.
- **Evaluator credibility:** The persons conducting the evaluation should be both trustworthy and competent to perform the evaluation, so that the evaluation findings achieve maximum credibility and acceptance.
- **Information scope selection:** Information collected should be broadly selected to address pertinent questions about the program and be responsive to the needs and interests of participants and other specified stakeholders.
- **Values identification:** The perspectives, procedures and rationale used to interpret the findings should be carefully described so that the bases for value judgments are clear.
- **Report clarity:** Evaluation reports should clearly describe the program being evaluated, including its context and the purposes, procedures and findings of the evaluation, so that essential information is provided and easily understood.
- **Report timeliness and dissemination:** Significant interim findings and evaluation reports should be disseminated to intended users so that the information can be used in a timely fashion.
- **Evaluation impact:** Evaluations should be planned, conducted and reported in ways that encourage follow-through by stakeholders, so that the likelihood that the evaluation will be used is increased.

Feasibility Standards

Feasibility standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

- **Practical procedures:** The evaluation procedures should be practical to keep disruption to a minimum while needed information is obtained.
 - **Political viability:** The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained and possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted.
 - **Cost-effectiveness:** The evaluation should be efficient and produce information of sufficient value that the resources expended can be justified.
-

Program Evaluation Standards

Propriety (Ethical) Standards

Propriety standards are intended to ensure that an evaluation will be conducted legally, ethically and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

- **Service orientation:** Evaluations should be designed to assist organizations to address and effectively serve the needs of the full range of targeted participants.
 - **Formal agreements:** Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated to adhere to all conditions of the agreement or formally to renegotiate it.
 - **Rights of human subjects:** Evaluations should be designed and conducted to respect and protect the rights and welfare of human subjects.
 - **Human interactions:** Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation, so that participants are not threatened or harmed.
 - **Complete and fair assessment:** The evaluation should be complete and fair in its examination and recording of strengths and weaknesses of the program being evaluated, so that strengths can be built upon and problem areas addressed.
 - **Disclosure of findings:** The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation and any others with expressed legal rights to receive the results.
 - **Conflict of interest:** Conflict of interest should be dealt with openly and honestly so that it does not compromise the evaluation process and results.
 - **Fiscal responsibility:** The evaluator's allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible, so that expenditures are accounted for and appropriate.
-

Program Evaluation Standards

Accuracy Standards

Accuracy standards are intended to ensure that an evaluation will review and convey technically adequate information about the features that determine worth or merit of the program being evaluated.

- **Program documentation:** The program being evaluated should be described and documented clearly and accurately, so that the program is clearly identified.
 - **Context analysis:** The context in which the program exists should be examined in enough detail that its likely influences on the program can be identified.
 - **Described purposes and procedures:** The purposes and procedures of the evaluation should be monitored and described in enough detail that they can be identified and assessed.
 - **Defensible information sources:** The sources of information used in a program evaluation should be described in enough detail that the adequacy of the information can be assessed.
 - **Valid information:** The information gathering procedures should be chosen or developed and then implemented so that they will assure that the interpretation arrived at is valid for the intended use.
 - **Reliable information:** The information gathering procedures should be chosen or developed and then implemented so that they will assure that the information obtained is sufficiently reliable for the intended use.
 - **Systematic information:** The information collected, processed and reported in an evaluation should be systematically reviewed and any errors found should be corrected.
 - **Analysis of quantitative information:** Quantitative information in an evaluation should be appropriately and systematically analyzed so that evaluation questions are effectively answered.
 - **Analysis of qualitative information:** Qualitative information in an evaluation should be appropriately and systematically analyzed so that evaluation questions are effectively answered.
 - **Justified conclusions:** The conclusions reached in an evaluation should be explicitly justified so that stakeholders can assess them.
 - **Impartial reporting:** Reporting procedures should guard against distortion caused by personal feelings and biases of any party to the evaluation, so that reports fairly reflect the evaluation's findings.
 - **Meta-evaluation:** The evaluation itself should be formatively and summatively evaluated against these and other pertinent standards, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses.
-

Appendix II: Electronic Evaluation Resources⁹⁹

American Evaluation Association: Personnel Evaluation - www.eval.org/EvaluationDocuments/perseval.html
American Evaluation Association: Program Evaluation - www.eval.org/EvaluationDocuments/progeval.html
Bureau of Justice Assistance Evaluation Website - www.bja.evaluationwebsite.org/
Center for Disease Control Evaluation Working Group - www.cdc.gov/eval/resources.htm
Community Toolbox: Community Building Tools - <http://ctb.lsi.ukans.edu/>
DOE: Education Program Evaluation - www.ed.gov/offices/OUS/PES/index.html
Educational Resources Information Center/ Assessment & Evaluation Clearinghouse - <http://ericae.net>
Evaluators' Instrument Exchange - <http://141.218.173.232:120/xchange/default.htm>
Evaluating Comprehensive Community Change - www.aecf.org/publications/evaluation/index.htm
Evaluation Exchange, Harvard Family Research Center - <http://gseweb.harvard.edu/~hfrp/eval.html>
Evaluation for Learning: Greater Kalamazoo Evaluation Project Newsletters - www.wmich.edu/evalctr/eval_nsltr/evalnsltr.htm
Free Management Library: Evaluations and Research Methods - <http://www.mapnp.org/library/evaluatn/evaluatn.htm>
GAO: US General Accounting Office / Designing Evaluations - www.gao.gov/policy/guidance.htm
Getting Smart, Getting Real: Using Research and Evaluation Information - www.aecf.org/publications/getsmart/aecget.htm
Grantmakers for Effective Organizations Resources - http://www.geofunders.org/main/resources/selected_publi_paper.htm
Human Services Research Institute: The Evaluation Center - <http://tecathsri.org/>
Innovation Network Helping Nonprofits Succeed - www.innonet.org
Institute for Research in Social Science Public Opinion Questionnaire Database - <http://www.ciesin.org/datasets/irss/irss.html>
James Irvine Foundation - www.irvine.org/
Management Assistance Program for Nonprofits - www.mapfor nonprofits.org
National Evaluation Data Services - <http://neds.calib.com/>
NSF: Online Evaluation Resource Library - www.nsf.gov/
NSF's User-Friendly Handbook for Mixed-Method Evaluations - www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/start.htm
Online Evaluation Resource Library - www.oerl.sri.com
Online Evaluation Workbook: Essentials of Survey Research - www.tfn.net/%7Epolland/qbook.html
Planning & Conducting Performance-based Evaluations (Wholey & McLaughlin) - www.ed.uiuc.edu/sped/tri/evalwkshp.htm
Precede/Proceed Model for Development. & Evaluation of Health Ed. Programs - www.med.usf.edu/%7Ekmbrown/PRECEDE_PROCEED_Overview.htm
Program Evaluation - <http://www.extension.psu.edu/evaluation/category.html>
Process and Outcome Trochim's Research Methods Knowledge Base - <http://trochim.human.cornell.edu/kb/index.htm>
Program Logic Model Excerpt - www.uottawa.ca/academic/med/epid/excerpt.htm
Program Manager's Guide to Evaluation - www.acf.dhhs.gov/programs/core/
Qualitative Research Resources - <http://don.ratcliff.net/qual>
Sociometrics - www.socio.com/
The Measurement Group Evaluation & Dissemination Center - www.themeasurementgroup.com/evalbtn.htm
UNICEF Research and Evaluation - www.unicef.org/reseval/
United Way's Outcome Measurement Resource Network - <http://national.unitedway.org/outcomes>

Appendix III: Rubric Template¹⁰⁰

(Describe here the task or performance that this rubric is designed to evaluate.)

	Beginning 1	Developing 2	Accomplished 3	Exemplary 4	Score
Stated Objective or Performance	Description of identifiable performance characteristics reflecting a beginning level of performance.	Description of identifiable performance characteristics reflecting development and movement toward mastery of performance.	Description of identifiable performance characteristics reflecting mastery of performance.	Description of identifiable performance characteristics reflecting the highest level of performance.	
Stated Objective or Performance	Description of identifiable performance characteristics reflecting a beginning level of performance.	Description of identifiable performance characteristics reflecting development and movement toward mastery of performance.	Description of identifiable performance characteristics reflecting mastery of performance.	Description of identifiable performance characteristics reflecting the highest level of performance.	
Stated Objective or Performance	Description of identifiable performance characteristics reflecting a beginning level of performance.	Description of identifiable performance characteristics reflecting development and movement toward mastery of performance.	Description of identifiable performance characteristics reflecting mastery of performance.	Description of identifiable performance characteristics reflecting the highest level of performance.	
Stated Objective or Performance	Description of identifiable performance characteristics reflecting a beginning level of performance.	Description of identifiable performance characteristics reflecting development and movement toward mastery of performance.	Description of identifiable performance characteristics reflecting mastery of performance.	Description of identifiable performance characteristics reflecting the highest level of performance.	

Appendix IV: Sample Rubrics¹⁰¹

Behavior Skill	(Beginning)	(Developing)	(Accomplished)	(Exemplary)
	Rarely	Sometimes	Most of the Time	Always
On Time and Prepared				
1. Arrives to work on time				
2. Brings necessary equipment				
3. Completes project activities				
Respects Peers				
1. Respects others property				
2. Listens to peers				
3. Responds appropriate to peers				
4. Respects others opinions				
5. Refrains from abusive language				
Respects Project Leader/Staff				
1. Follows directions				
2. Listens to Project Leader/Staff				
3. Accepts responsibility for actions				
Demonstrates Appropriate Character Traits				
1. Demonstrates positive character traits (kindness, trustworthy, honesty)				
2. Demonstrates productive character traits (i.e., patience, thorough, hardworking)				
3. Demonstrates a level of concern for others				
Demonstrates a Level of Concern for Learning				
1. Remains on task				
2. Allows others to remain on task				

Collaboration Rubric¹⁰²

	Beginning 1	Developing 2	Accomplished 3	Exemplary 4	Score
Contribute					
Research & Gather Information	Does not collect any information that relates to the topic.	Collects very little information--some relates to the topic.	Collects some basic information--most relates to the topic.	Collects a great deal of information--all relates to the topic.	
Share Information	Does not relay any information to teammates.	Relays very little information--some relates to the topic.	Relays some basic information--most relates to the topic.	Relays a great deal of information--all relates to the topic.	
Take Responsibility					
Fulfill Team Role's Duties	Does not perform any duties of assigned team role.	Performs very little duties.	Performs nearly all duties.	Performs all duties of assigned team role.	
Share Equally	Always relies on others to do the work.	Rarely does the assigned work--often needs reminding.	Usually does the assigned work--rarely needs reminding.	Always does the assigned work without having to be reminded.	
Value Others' Viewpoints					
Listen to Other Teammates	Is always talking--never allows anyone else to speak.	Usually doing most of the talking--rarely allows others to speak.	Listens, but sometimes talks too much.	Listens and speaks a fair amount.	
Make Fair Decisions	Usually wants to have things their way.	Often sides with friends instead of considering all views.	Usually considers all views.	Always helps team to reach a fair decision.	
				Total	

Appendix V: Designing Age-Appropriate Tools for Youth and Children

Many youth-serving programs have focused their evaluation strategies on adults—interviewing staff, parents and teachers about the programs, yet published studies have shown that data provided by adolescents (10 to 18 years) are of significantly better quality than that provided by younger children and are often as trustworthy as their parents' responses. However, collecting evaluation information from youth presents unique challenges due to the varying cognitive and social developmental stages as they mature.

The table below reviews the developmental stages of adolescents as described by Rubenstein.¹⁰³ (Note that for youth ages 10 to 14, cognitive status is significantly different than for their older counterparts.) It is important to take the youth's developmental stages into consideration when considering survey administration modes and instrument verbiage.

Overview of Adolescent Development			
Phase	Ages	General Issues	Cognitive Issues
Early Adolescence	10-14	Focus on physical development Feel they are the center of attention Feel invulnerable to usual problems in the world	Think in concrete terms Recall memory is developed to the level of a young adult Difficulty thinking about the future Think about events in terms of their own experience Difficulty thinking about hypothetical situations
Middle Adolescence	15-17	Compelled to make independent decisions Advice and feedback from peers becomes very important Often reject parental values More concerned about attractiveness and more comfortable with the opposite sex	Some abstract thought Can plan ahead Begin to perceive relationship between present actions and future consequences
Late Adolescence	18-21	Able to function independently Often more willing to listen to parental advice Continue to value peer relationships	Abstract thought is fully developed Able to think about both short-term and long-term goals

If you are designing your own instruments to use with youth, consider the following guidelines

Designing Instruments for Adolescents			
Recommendation	Reason	How NOT to Say It...	How To Say It...
Write questions at the literacy level of the participants of your program.	Establishes rapport.	How much vitality have you felt during the past month?	How much energy have you had during the last month?
Avoid using outdated phrases or slang.	Reduces credibility of questions and may be the focus of ridicule rather than serious thought.	Do you think this program is the bomb?"	Do you enjoy this program?
Write questions consistent with the participant's stock of knowledge.	Avoids frustration and embarrassment for the adolescent.	What is your parent's occupation?	Where does your mother work? What does she do where she works?
Write questions or statements in first person.	Takes developmental stages into consideration. Prior to late adolescence, young people think concretely and may interpret questions literally. Adolescents may answer questions asked in the third person on behalf of the person asking (an adult) rather than him or herself.	People of my age care more about being happy than healthy.	I care more about being happy than healthy.
Consider simple open-ended questions versus complex multiple-choice questions. (Note: Simple multiple-choice questions are best. Avoid using too many open-ended questions, as they require more staff time to code and analyze them.)	Reduces the amount of time needed to review unfamiliar options. Open-ended questions allow adolescents to answer in their language.	What activities do you take part in after school? <input type="checkbox"/> sports <input type="checkbox"/> do my homework <input type="checkbox"/> boy/girl scouts <input type="checkbox"/> help parents on their farm <input type="checkbox"/> hang out <input type="checkbox"/> go to the mall <input type="checkbox"/> volunteer	What activities do you take part in after school? _____ _____ _____ _____

Designing Instruments for Adolescents			
Recommendation	Reason	How NOT to Say It...	How To Say It...
Ask questions rather than respond to statements.	Adolescents have difficulty understanding how to respond to Likert scales (agree/disagree)	My family is important to me. <input type="checkbox"/> strongly agree <input type="checkbox"/> agree <input type="checkbox"/> neither agree nor disagree <input type="checkbox"/> disagree <input type="checkbox"/> strongly disagree	Is your family important to you? <input type="checkbox"/> NO! <input type="checkbox"/> no <input type="checkbox"/> yes <input type="checkbox"/> YES!
Maintain similar response categories across questions.	Facilitates questioning process for adolescents.	How many times, if ever, have you been physically injured in a fight at school? <input type="checkbox"/> Never <input type="checkbox"/> Once <input type="checkbox"/> Twice <input type="checkbox"/> 3 times <input type="checkbox"/> 4 times or more How many times, if ever, have you been physically injured in an accident at school? <input type="checkbox"/> Never <input type="checkbox"/> 1-2 times <input type="checkbox"/> 3-4 times <input type="checkbox"/> 5 or more times	How many times, if ever, have you been physically injured in a fight at school? <input type="checkbox"/> Never <input type="checkbox"/> 1-2 times <input type="checkbox"/> 3-4 times <input type="checkbox"/> 5 or more times How many times, if ever, have you been physically injured in an accident at school? <input type="checkbox"/> Never <input type="checkbox"/> 1-2 times <input type="checkbox"/> 3-4 times <input type="checkbox"/> 5 or more times
Use positively stated questions and do not mix positive and negative questions.	Improves question interpretation. Research shows that adolescents interpret the meaning of positively and negatively worded questions differently.	I do not like riding a bus to school. I like my teacher.	I like riding a bus to school. I like my teacher.

Designing Instruments for Adolescents			
Recommendation	Reason	How NOT to Say It...	How To Say It...
Allow participants to answer "I don't know."	Removes pressure from adolescents to report incorrect answers if they do not have another option.	What is the highest grade in school that your mother has completed? <input type="checkbox"/> she did not attend school <input type="checkbox"/> between grades 1 and 6 <input type="checkbox"/> between grades 7 and 12 <input type="checkbox"/> graduated from high school <input type="checkbox"/> some college <input type="checkbox"/> graduated from college	What is the highest grade in school that your mother has completed? <input type="checkbox"/> she did not attend school <input type="checkbox"/> between grades 1 and 6 <input type="checkbox"/> between grades 7 and 12 <input type="checkbox"/> graduated from high school <input type="checkbox"/> some college <input type="checkbox"/> graduated from college <input type="checkbox"/> I don't know

Some other tips when evaluating youth outcomes¹⁰⁴:

- Adolescents tend to be task-oriented. Pose evaluation tasks as a challenge and encourage progress.
- Youth respond better to administrators who share characteristics with them (e.g., same age, gender, social class and race).
- Adolescents require privacy when answering any sensitive questions. They will actively seek privacy, and if not provided, they will leave responses missing or answer incorrectly.
- If adolescents were required in some way to be in the program, it is important that the program administrator explain in a friendly manner the importance of the research, yet give youth the opportunity to refuse all or part of the evaluation.
- Avoid long pauses, interruptions and other opportunities for distractions while the respondent is involved in the evaluation process.

Response Options that Work Well with Youth and Children

Response options are important to consider when using multiple-choice questions. They are used to pre-code answers from participants. Pre-coding of the response categories is often done because it eliminates the need for coding results at a later date, makes it easier on the respondent, eliminates bias because it allows respondents to choose the category that best fits their opinion, provides consistency in response across participants and, in some cases, better defines the question.

NRC staff have reviewed hundreds of questions asked of youth including both older and younger adolescents, to assess program outcomes. The response scales most often provided are those most commonly used with adult respondents: four or five option scales using agreement anchors (e.g., strongly agree – strongly disagree), importance anchors (e.g., very important – not important) and frequency anchors (e.g., never – often). While these scales may work very well with older adolescents or youth of higher literacy, they may be too sophisticated or too boring for some youth participants. It is important to construct any evaluation tool at the lowest literacy level that your youth participants and others participating in the evaluation can understand and respond. Response scales used with younger respondents often are based on fewer options (4 options or less) and use graphics or visual cues to demonstrate the relationship of options to each other. Some response scales that have been developed and used successfully for elementary and middle school children are presented in the figure below.

Scales Using Words as Anchors	
Yes	No
YES! yes	no NO!
Like Me	Not Like Me
Agree	Disagree

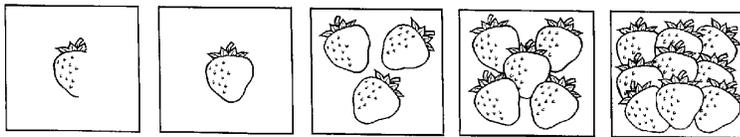
Visual or graphic scales work the best with younger kids and may be more or less culturally biased depending on the population for which they are used. Such scales also tend to be simplistic, which may insult some of your older participants. Examples of visual/graphic scales are presented on the following page. The “smiley face” scale is the most popular scale of this sort.

Scales Using Pictures for Response Options¹⁰⁵:

Example 1: At this program I feel . . .



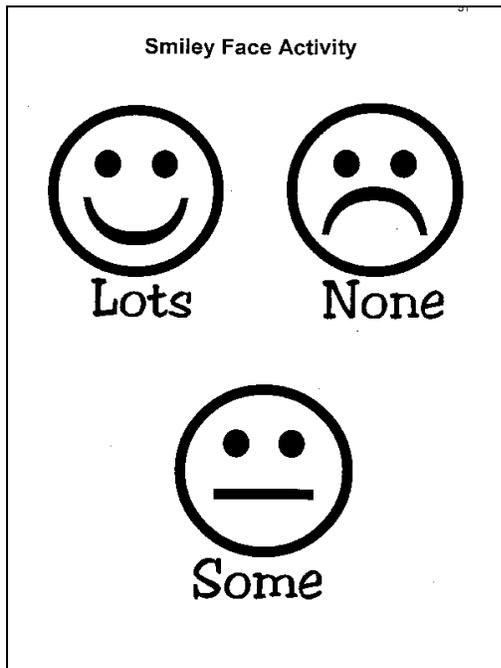
Example 2: At this program I made new friends.



Example 3: I learned a lot from being in this program.



Example 4: I read every day for fun.



Appendix VI. Tips for Conducting Focus Groups with Teens

Teen group dynamics will likely be influenced by many factors: gender, cliques, grade in school, age differences, how well participants know each other or if they know each other. So it is essential to customize focus group planning and moderation to fit the needs of the teen participants. Teens open up best when they feel that someone really values their opinion and when they feel that they are part of a team helping to make something better. Consider the following tips when planning your next focus group with teens.

❖ Conduct gender-alike groups

Planning gender-alike groups can help minimize the impact of peer and environmental influences.

❖ Limit group size

While typical focus groups might have 8 to 12 participants, hold teen group sizes to only around 5 to 7 for a more manageable and productive group.

❖ Select cross-section, homogeneous groups

Generally, focus groups are comprised of individuals who are similar to each other in some way. With teens, commonality may be important as it relates to their age, their ethnicity, their weight or socio-economic status, depending on the focus of the group. However, the less teens in the group know each other, the more likely they may feel comfortable talking about personal issues, so try to select “similar” teen participants from different schools, clubs or programs.

❖ Rapport is critical

Start building rapport with teen participants as soon as the first one walks in the room. Learn their names, where they go to school, chat with them before sitting down to start the group. This builds rapport with the teens right off the bat and teens prefer an atmosphere that is easy going, friendly and focused on them.

❖ Set an informal environment

To establish an informal atmosphere, take a waste paper basket and place it in the middle of some chairs. Top the basket with a large piece of cardboard as a table top and set the tape recorder on it. This setting can help establish a more cohesive group and allows participants eye contact with each other. Since the table is low, it is not a formal barrier the way most tables can appear and teens are likely to soon forget the tape recorder is there, once the conversations begin. Alternatively, everyone can sit on the floor.

❖ **Use different probes than with adults**

Teens are more likely to give you short answers especially at first. Respond to one-word answers directly by repeating the response and asking for more detail. “Tell me more about that.” “What does that mean to you?” “What else did you think about x?” This demonstrates the importance of each participant’s contribution and that the moderator really does not want a “right answer” or a monosyllabic answer. On occasion a teen may not feel comfortable saying more. In this situation, back off, thank the teen for speaking up and then ask for someone else’s opinion.

❖ **Confidentiality is key**

Often focus groups deal with sensitive or highly personal information. Consider creating a participant confidentiality form, pass copies around the room as the group begins and read it aloud; something like “...I promise not to repeat anything that is talked about in this group to anyone else ever...” Then ask each participant to sign it and collect the forms immediately. This act will help set a tone for honesty, openness and respect of the process.

❖ **Call their bluff and call them to action**

Teens want to know you are hearing them and are taking what they say seriously. How do they know you are really writing down what they say? If they are at first in doubt of your sincerity, call their bluff. Tell them that you plan to write down everything they say and then, using a flip chart, write it all down. After you get through some silly or dreadful, even “illegal” comments and once teens realize you are not going to judge or redirect them, they will usually settle down into some real participation. With groups that need this extra encouragement, consider sending your notes back to them later so they can check to make sure you heard them correctly.

❖ **Follow-up and aftercare**

Teens open up best when they feel that someone really values their opinion and when they feel that they are part of a team helping to make something better. In case participants have something more they would like to share with you in a more private setting give out your phone number or email address or that of a neutral party. Also consider giving participants a resource person with whom they can talk if personal feelings or reactions have come up for them.

A publication of National Research Center, Inc. with material adapted from the following sources:

Kiernan, N. E. (2001). Focus Groups: 4-H and Youth Development: Tipsheet 29. University Park, PA: Penn State Cooperative Extension.

Molloy, P. (2004, November 15). Follow up to focus groups. Message posted to EVALTALK@BAMA.UA.EDU

Appendix VII: Designing Age-Appropriate Tools for Older Adults

Older adults may provide unique challenges in evaluation due to the variability in physical health and cognitive abilities. Some of these challenges are presented in the table below, along with recommendations to improve the effectiveness of data collection.

Administration Tips for Older Adults	
Potential Challenge	Recommended Methods
Poor vision	<input type="checkbox"/> Administer instrument on the phone or in person. <input type="checkbox"/> If using a typed instrument use 14-point font. <input type="checkbox"/> Have volunteers or allow family member help elder complete survey.
Poor hearing	<input type="checkbox"/> Use written surveys.
Poor physical health	<input type="checkbox"/> Have volunteers or allow family members to help elder complete survey. <input type="checkbox"/> Keep surveys short and simple. Do not ask complex questions or use complex props.
Raised before baby boom generation	<input type="checkbox"/> Be sensitive to word choices. Do not use slang. <input type="checkbox"/> Be sensitive when asking personal questions. Many of the topics discussed more openly today were not 20 years ago. Anonymity is very important. <input type="checkbox"/> Do not use technology such as computers or a Walkman.
More vulnerable to fraud	<input type="checkbox"/> State clearly the purpose, the civic sponsorship and the importance of the data collection to ensure respondents feel comfortable.

To ensure your instrument is simple, understandable and gets at the right information, make sure to pilot-test it with a subset of older adults from the population you plan to sample.

Appendix VIII: Common Response Options

Agreement Scales

- Agree/Disagree
- Disagree, Agree
- Strongly Disagree, Disagree, Neither Agree Or Disagree, Agree, Strongly Agree
- Strongly Agree, Slightly Agree, Slightly Disagree, Strongly Disagree
- True/False
- False, True
- Definitely Not True, Mostly Not True, Mostly True, Definitely True
- Less True-The Same-More True
- Not At All True, Hardly Ever True, True A Lot, Almost Always Or Always True

Yes/No

- Yes, No
- Yes, No, Maybe
- No, Maybe, Yes, Strong Yes
- YES, yes, not sure, no, NO
- Yes, Sort of, Not really

Quality Scales

- Poor, Fair, Good, Excellent
- Very bad, Bad, Neither Bad nor Good, Good, Very Good
- Smiley Faces (5 Range)
- Importance Scales
- Not Important, Important
- Not Important At All, Somewhat Important, Very Important, Extremely Important
- Not Important, Somewhat Important, Very Important
- Not Important, Somewhat Important, Moderately Important, Very Important
- Not Important, Somewhat Important, Very Important, Essential
- Not Important, Somewhat Important, Important, Very Important

Frequency Scales

- Never, Sometimes, Usually, Always
- Never, Almost Never, Sometimes, Almost Always, Always
- Never, Once A Year, Once A Month, Once A Week, Everyday
- Never, Once or Twice, More Than Twice
- Never, One or Two Times, About Once A Week, A Few Times A Week, Every Day
- Never, Seldom, Sometimes, Almost Always, Always,
- Not At All, 1 Time, 2 Times, 3 Or More Times

Likelihood Scales

- Definitely Not, Probably Not, Probably Will, Definitely Will
- Definitely Would, Probably Would, Not Sure, Probably Would Not, Definitely Would Not
- Very likely, Somewhat likely, Somewhat unlikely, Very unlikely
- Very likely, Likely, Unlikely, Very unlikely

Amount Scales

- A Great Deal, Quite A Bit, Some, Not Too Much, Very Little
- Less Than Would Like, About Right, More Than Would Like
- Lots!, Some, A Little Bit, None At All
- None, A Few, Some, Most, All
- None, A Little, A Lot
- Too Much, About Right, Too Little

Change Scales

- A Decrease Since I Started, No Difference Since I Started, A Little Increase Since I Started, Noticeably Increased Since I Started
- Much Better, A Little Better, About The Same, A Little Worse, A Lot Worse
- Much Better, A Little Better, No Change, A Little Worse, Much Worse
- Scales to Measure Skill/Knowledge Attainment
- Cannot Do At All, Can Do A Little Bit, Can Do Pretty Well, Can Do Very Well, Do Not Understand
- Confident, Ok, Unsure, Awful
- I'm A Genius!, I'm Pretty Good, It's Kind Of Hard, Not At All
- Know Nothing, Know A Little, Know Some Things, Know Most Things, Know Everything, Does Not Apply
- No Difficulty, A Little, Moderate, Quite A Bit, Extreme
- Not Like Me, Somewhat Like Me, Very Much Like Me
- Very Comfortable, Somewhat Comfortable, Somewhat Uncomfortable, Very Uncomfortable

Appendix IX: Descriptive Analyses Using Excel

If no other statistical analysis resources are available to you, Microsoft Excel can provide simple descriptive analyses of data. In this appendix, we provide an example and some explanation for generating an average rating and a frequency distribution for survey questions below. In writing these explanations, we have assumed that readers will have some basic familiarity with Excel. Readers/users who understand cell referencing and formulas should be able to produce similar analyses with their own data. The entire document can be found at: www.n-r-c.com/excel/handbook.pdf.

Example Analysis

The figure below shows how Excel was used to enter 7 cases from a survey, and then to calculate average ratings and frequency distributions for these results. For this example we used the questionnaire on page 203 as the sample source of data.

	A	B	C	D	E	F	G	H
		q1a 1=poor 2=fair 3=good 4=very good 5=excellent	(formula)	q1b 1=poor 2=fair 3=good 4=very good 5=excellent	q1c 1=poor 2=fair 3=good 4=very good 5=excellent		q3 1=none of my needs met 2=only a few of my needs met 3=most needs met 4=most all of my needs met	
1	ID							
2		1	3		4	5		3
3		2	4		3	4		
4		3	5		5	4		
5		4	3			4		2
6		5	5		5	4		3
7		6	4		4	4		3
8		7	5		4	5		3
9								
10	TOTAL surveys	7	=ROWS(B2:B8)		7	7	TOTAL surveys	7
11								
12	Average Rating	4.14	=AVERAGE(B2:B8)	4.17	4.29		Average Rating	2.80
13								
14	Number Rating "Excellent"	3	=COUNTIF(B\$2:B\$8,5)		2	2	Number Rating "Most All"	0
15	Number Rating "Very Good"	2	=COUNTIF(B\$2:B\$8,4)		3	5	Number Rating "Most"	4
16	Number Rating "Good"	2	=COUNTIF(B\$2:B\$8,3)		1	0	Number Rating "Only a Few"	1
17	Number Rating "Fair"	0	=COUNTIF(B\$2:B\$8,2)		0	0	Number Rating "None"	0
18	Number Rating "Poor"	0	=COUNTIF(B\$2:B\$8,1)		0	0		
19	TOTAL with an opinion	7	=SUM(B14:B18)		6	7	TOTAL with an opinion	5
20								
21	Percent Rating "Excellent"	43%	=B14/B\$19	33%	29%		Percent Rating "Most All"	0%
22	Percent Rating "Very Good"	29%	=B15/B\$19	50%	71%		Percent Rating "Most"	80%
23	Percent Rating "Good"	29%	=B16/B\$19	17%	0%		Percent Rating "Only a Few"	20%
24	Percent Rating "Fair"	0%	=B17/B\$19	0%	0%		Percent Rating "None"	0%
25	Percent Rating "Poor"	0%	=B18/B\$19	0%	0%			
26	TOTAL	100%	=B19/B\$19	100%	100%		TOTAL	100%
27								
28								
29								
30								
31								
32								

The IDs (1 through 7) were placed in Column A, and the results of the first, second and third parts of question 1 were placed in Columns B, D and E, respectively. The results of question three were put in Column G. This was done using the procedures described in the section "Creating an Electronic Dataset" (starting on page 125).

Once the data were entered (in this example, we only entered information from the first seven surveys), analyses could be performed. Column C shows the formulas that were used to produce the analyses.

Using Excel

Cell references

“Cells” in an Excel spreadsheet are referred to by the intersection of the Column and Row in which they appear. In the example on the previous page, the cell that contains the label “ID” is cell A1, because it is in the first column (A) *and* the first row (1). The average rating given by respondents to the first part of question 1 is in cell B12, at the intersection of the second column (B) and the twelfth row.

Cell references are used in formulas (see below for more on formulas), usually to perform calculations or make reference to other cells in the spreadsheet.

Formulas

Formulas are used to perform calculations within a spreadsheet. To insert a formula as opposed to a number or text, type an equals sign (“=”) in the cell where you wish to perform the calculation, and then type in the rest of the formula. A formula can perform mathematical calculations or execute a wide variety of functions (see below for more on functions). To add or subtract: use the plus (+) or minus (-) symbol. To multiply: use an asterisk (*) and to divide use a slash (/). Use parentheses as necessary to indicate the desired order of operations.

For example, if you wanted to know how many seconds there were in three hours, you could type in the formula: =3*60*60. The result displayed in the cell would be 10,800.

You might have a cell somewhere on the page that had a value of “3” to indicate three hours; for the sake of an example, we’ll say that the cell is T21. If you wanted to know how many seconds that represented, you could use the same formula as above, but exchange the cell reference for the “3”. For example: =T21*60*60. If the number of hours in cell T21 changed, the result of the formula would also change.

Absolute versus relative cell references

In a formula, a cell reference can be made in a “relative” or an “absolute” manner. For example, looking at the table below, if you wanted to calculate a percent, you might create a formula in cell C2 that is the proportion of youth served who are 12 to 14 years old. That formula would be: =B2/B5, which would divide the value of B2 (12) by the value of B5 (112).

	A	B	C	
1	Age range	Number served	Percent served	(Formula)
2	12-14 year olds	12	10.7%	=B2/B5
3	15-17 year olds	68		
4	18-20 year olds	32		
5	TOTAL	112		
6				

You may then wish also to calculate the proportion of youth served who are 15 to 17 years old. If you copied the contents of cell C2 to cell C3, and then examined the formula, it would look like this: =B3/B6. This is because in Excel the cell references in your formula are “relative” references; that is, Excel has assumed that because in cell C2 you were dividing the number in the same row and one column to the left by the number three rows below and one column to the left, you want to do the same thing in the cell to which you are copying the formula. However, cell B6 is blank, so an invalid number would be calculated in cell C3 using this formula. You can fix this by changing the formula after you have copied it, so that the denominator refers to B5. But, if you then copy the formula to cell c4, you would again have to manually change the denominator in the formula to refer to the correct cell, which contains the total number of youth served. If you did not make this manual change, the formulas in column C would look like the formulas in column D in the table below.

If, however, you used an “absolute” reference to refer to the row, which contains the total number of youth served, when the formula was copied, the denominator would always refer to row five. The dollar sign (\$) is used to indicate an absolute reference. In this example, it is only used for the row designation, not for the column designation. It can be used for both the row and column designation, or only one or the other. Excel defaults to assuming that all cell references are relative, unless you make the change. Knowing how to use relative and absolute references can greatly speed up creation of spreadsheets in Excel.

	A	B	C	D	E
1	Age range	Number served	Percent served	(Relative formula copied down)	(Absolute formula copied down)
2	12-14 year olds	12	10.7%	=B2/B5	=B2/B\$5
3	15-17 year olds	68		=B3/B6	=B3/B\$5
4	18-20 year olds	32		=B4/B7	=B4/B\$5
5	TOTAL	112			
6					

Formatting cells

As can be seen above, some of the cells are formatted as numbers, and some are formatted as percents. You will want to format the cells appropriately. To format a cell or group of cells, highlight the cells you wish to format, and then choose “Format” from

the menu bar, and then “Cells.” A dialogue box will open, with a number of formatting options. You can format the alignment of the cell contents, the cell shading or border, or the “Number”. If you choose the “Number” tab you will be presented with a list of types of number formats, such as “currency,” “percentage,” etc. Choose the type and then decide how many decimals you want. The highlighted cells will be formatted according to the specifications you choose.

Referring to a range of cells

When using some functions (see below), you will want to refer to a “range” of cells. For example, if you wanted to total the number of youth served in the table above, you could use a formula in cell B5 like this: =B2+B3+B4. Alternatively, you could use the SUM function, and refer to a range of cells to be summed, like this: =SUM(B2:B5). The colon indicates that a range of cells is being referred to, starting with (and including) the cell to the left of the colon, and ending with (and including) the cell to the right of colon. The function “SUM” indicates what is to be done with this range of cells – total all the values together.

Functions

“SUM” is only one of a large number of functions available in Excel. Some of the functions are mathematical, some are logical, some are statistical, and others serve yet more purposes. For the analyses we show in this section, we will concentrate on just a few functions.

All functions begin in a similar fashion: the function, immediately followed by an open parenthesis, the references on which the function should operate each separated by a comma (a different number of references are needed for each function), and a close parenthesis. The functions needed for simple descriptive analyses in Excel are shown on the next page.

Functions and formulas used for simple descriptive analyses in Excel

The table below displays the analyses performed in the example on page 203. The examples all refer to that spreadsheet.

Calculate ...	By ...	Using the function or formula ...	Operators are:	Example:	Value displayed	What it means:
The number of surveys completed	Counting the number of rows of data entered	ROWS	Range of cells for which the number of rows should be counted	=ROWS(D2:D8)	7	7 surveys were completed
The average rating of those who responded	Calculating the average of the ratings given by those answered	AVERAGE	Range of cells containing the values to be averaged	=AVERAGE(D2:D8)	4.17	The average rating for question 1b was 4.17, where 1=poor and 5=excellent.
The number of respondents who gave a specific answer*	Counting the number of responses of a certain type within a range of cells	COUNTIF	1) The range of cells to be examined 2) The value to be counted	=COUNTIF(D\$2:D\$8,5)	2	2 people gave an answer of "excellent" (5) to question 1b
The total number of respondents who answered the question**	Adding the number of people who gave a valid answer to a question	SUM	Range of cells to be totaled	=SUM(D14:D18)	6	6 people answered question 1b
The proportion (percent) of respondents who gave a specific answer	Dividing the number of people who gave a specific answer by the total number of people who gave answered the question	(Division) [Cell reference1]/[cell reference2]	Cell reference1 is the cell reference of the numerator; cell reference2 is the cell reference of the denominator	=D14/D\$19	29%	29% of respondents gave an answer of "excellent" to question 1b

*This is used for each "row" or part of a frequency distribution.

** Or the sum of any list of numbers.

Appendix X: More Advanced Statistics and Statistical Testing

Commonly used inferential statistics are described in this appendix. Please refer to a statistical text for more information on the calculation or interpretation of statistics. References for such texts are provided on page 132 of this handbook.

Definitions of Statistical Terms	
Term	Definition and Interpretation
Inferential Statistics	
Simple Correlation	Correlation or “linear regression” measures the relationship between two variables, and is measured by the “r” statistic, which is bounded from +1 to -1. A score of +1 would indicate a perfect, positive linear relationship (e.g., both variables are increasing or decreasing at the same rate). A score of -1 would indicate a perfect inverse relationship (while one variable is increasing at an equal rate as the other is decreasing). “r” values closer to “0” indicate no relationship between the variables.
Multiple Regression	Similar to simple correlation, multiple regression also summarizes the relationship of variables. Multiple regression however examines the relationship of multiple variables to predict a single outcome (e.g., using age, gender and income status to predict program competence). The statistic is used to test the contribution or relative importance of each socio-demographic or risk factor to an outcome. It also allows the prediction of outcomes based on the intake characteristics of the client.
T-test	A t-test is used to test the null hypothesis that the mean scores of two groups are the same. A statistically significant result ($p < .05$) suggests the means are significantly different or larger than would be expected by chance alone. See also the “p-value” definition below.
Chi-square	The chi-square statistic is used to test the null hypothesis that all of the proportions in a table are the same. A statistically significant test value ($p < .05$) indicates that the proportions are significantly different. See also the “p-value” definition below.
Statistical Significance	
.95 Confidence Interval	1.96 times the standard deviation of the scores divided by the square root of the number of scores. Simply stated, you can be confident that this interval around your mean contains the true population mean because intervals calculated the way you calculated this one will capture the population mean 95 times in 100.
Significance Level	Probability that the results you found could have occurred by chance alone. As the number gets smaller (.05 or less by tradition) the conclusion is that the results are not due to chance but that they are significant. This level should be determined prior to running the tests.
P-value	Same as significance level. “p” is the probability that the results found could not have occurred by chance alone.

Appendix XI: Sample Executive Summary

Sample Executive Summary for a Fictitious Farmers' Market

Fresh Market Merchant Survey: Executive Summary of Evaluation Findings

Background

- The Fresh Market has been in operation since the October 1996. The goals of the market have been to: 1) help community residents gain access to fresh, local produce and 2) help immigrant farmers in the area sell their products and increase their knowledge about farming and agriculture.
- Bonita Rodriguez founded the market after El Mercado grocery store closed, leaving community residents with little access to fresh produce or traditional foods. This market, along with many others, is funded by the U.S. Department of Agriculture to increase community food security in urban areas.
- The Fresh Market operates on the corner of 13th and Lily, on Wednesdays and Saturdays from 10:00 a.m. to 2:00 p.m. The market primarily serves residents in the Dos Hermanos neighborhood. Around 300 to 400 residents visit Fresh Market each week.
- As part of Fresh Market's ongoing evaluation practices, an anonymous survey was administered in September 2003 to those who operate stands at the Fresh Market.
- The survey was conducted to assess farmer satisfaction with the market and the impacts the program has had on their lives.
- Of the 24 farmers asked to participate in the survey, 20 completed the questionnaire, for a response rate of 83%.

Farmers' Evaluation of the Market

- Fresh Market merchants participating in the survey were asked to rate various aspects of selling products at the market. The proportion with an opinion who rated each item as "good" or "excellent" was:
 - Overall quality of products sold at market, 90%
 - Location of market, 75%
 - Hours of operation of market, 70%
 - Advertising for market, 50%
- When asked how they felt about their overall experience with selling at this market, 45% rated it as "excellent," 35% as "good," and 20% as "fair." No respondents rated their experience as "poor."

- Those completing the questionnaire were asked what they liked best about selling at Fresh Market. They could answer the question in their own words. The responses were categorized. What survey participants liked best were:
 - Earning income or making a living in farming, 45%
 - Access to customers, 30%
 - Ability to sell traditional/cultural foods, 15%

Impact of the Market on Farmers

- A number of questions were included on the questionnaire to evaluate the impact of Fresh Market on the lives of farmers participating in it. One of the questions asked survey respondents to what extent they agreed or disagreed with a series of statements about the market. The percent with an opinion who “agreed” or “strongly agreed” with each statement follows:

Because I operate a stand at this market . . .

- I have developed a larger customer base, 100%
- I have learned more about running a small business, 90%
- I feel better about my future in farming, 75%
- I have earned more income from farming, 95%
- I am more able to provide food for my family and myself, 75%
- I have developed new products, 75%
- I have learned new farming skills, 65%
- I have learned more about organic farming, 60%
- About two-thirds of those completing the questionnaire felt that Fresh Market has made a “big difference” in helping them make a living at farming/agriculture.
- When asked how much money they had made at the market the day of the survey, a quarter had made less than \$200, although none of the farmers had made anything.
 - 20% made between \$200 and \$299
 - 25% had made between \$300 and \$399
 - 15% had made between \$400 and \$499
 - 10% had made between \$500 and \$749
 - 5% weren’t sure how much they had made
- Farmers were asked to respond in their own words about the ways in which their lives are different because of this market. Following are the percentages of respondents giving various answers:
 - More income for my family, 65%
 - I have started to grow new products, 45%
 - I am able to farm full-time, 30%
 - I know more about running a farm/managing farm sales, 25%
 - I am using more organic methods/less pesticides and chemicals, 15%

- Those completing the questionnaire were also asked how else the market could help to enhance their work as a farmer. The additional ways mentioned included:
 - More information on value-added products, 35%
 - More advertising in local newspapers, radio and media, 30%
 - More information on purchasing and using EBT machines, 15%
 - More courses on small business financing, 10%.

Characteristics of the Farmers and Their Operation

- A set of questions was asked to find out more about the types of stands operating at the market and about the farmers operating the stands.
- Most of the farmers (80%) operated a stand at least once a week and most commonly sold fruits and vegetables (80%), traditional/cultural foods (40%), and cheese and eggs (35%). Baked goods and other specialty foods were also offered.
- On average, farmers traveled 42 miles to get their goods to the market and farmed less than 6 acres (65%). Two-thirds were employed full-time as a farmer/food producer.
- Most (75%) of the stand operators were between the ages of 25 and 54. Three-quarters of those completing the questionnaire identified their race/ethnicity as Hispanic or Latino; 15% as Asian or Pacific Islander; 5% as American Indian, Eskimo or Aleut; and 5% as "Other."

Next Steps

Fresh Market Staff have identified several areas for improvement and have begun to develop activities to enhance their marketing outreach and education on organic products. Additionally, small groups of participant farmers will be convened to explore how the entire market experience can be strengthened.

Appendix XII: Sample Annotated Evaluation Tool

Sample annotated evaluation tool for a Fictitious Farmers' Market

Fresh Market Merchant Survey: Complete Set of Survey Responses

This survey was administered in September 2003. Of the 24 farmers asked to complete the survey, 20 did, providing a response rate of 83%. All survey responses were given in complete anonymity.

1. How often do you operate a stand at this market?

35% Twice a week 5% Two times a month 0% Once a month
 45% Once a week 15% Three times a month 0% Less than once a month

2. Please check the types of products you are selling today. (Please check all that apply.)

80% Fruits and vegetables 40% Traditional/cultural foods 35% Cheese and eggs
 25% Other: fish, herbs, spices, jellies, jams 20% Baked goods

3. About how far did you travel to get your goods to market today?

Mean 42 miles, median 35 miles (one-way)

4. What do you like best about selling at this market? (Open-ended responses coded during analysis.)

45% Earning Income/making a living in farming
 30% Access to customers
 15% Ability to sell traditional/cultural foods
 10% Other

5. Please rate each of the following aspects of selling products at this market.

	Poor	Fair	Good	Excellent	Don't Know
a. Location of market	5%	15%	60%	15%	5%
b. Hours of operation of market	20%	10%	50%	20%	0%
c. Advertising for market.....	20%	25%	25%	25%	5%
d. Overall quality of products sold at market.....	0%	10%	20%	70%	0%

6. How would you rate your overall experiences selling at this market?

0% Poor 20% Fair 35% Good 45% Excellent

7. In what ways do you think your life is different because of this market? (Open-ended responses coded during analysis. Respondents could give more than one answer.)

- 65% More income for my family
- 30% I am able to farm full-time
- 45% I have started to grow new products
- 25% I know more about running a farm/managing farm sales
- 15% I am using more organic methods/less pesticides and chemicals
- 20% Other

8. The following list contains some changes you may or may not have experienced because of participating in this market. Please indicate how much you agree or disagree with following statements as a result of operating a stand here.

Because I operate a stand at this market...

	Strongly Disagree	Disagree	Neither Nor Agree	Disagree Agree	Strongly Agree	Don't know
a. I have developed new products.....	0%	25%	0%	20%	55%	0%
b. I have learned new farming skills.....	0%	10%	25%	35%	30%	0%
c. I have learned more about organic farming.....	5%	15%	15%	20%	40%	5%
d. I have learned more about running a small business.....	0%	0%	0%	30%	60%	10%
e. I have earned more income from farming...	0%	0%	5%	15%	80%	0%
f. I feel better about my future in farming...	0%	15%	5%	15%	60%	5%
g. I have developed a larger customer base	0%	0%	0%	10%	80%	10%
h. I am more able to provide food for my family and myself.....	0%	0%	15%	25%	50%	10%

9. To what extent has this market helped you make a living at farming/agriculture? Does it make ...

0% No difference 10% Small difference 25% Moderate difference 65% Big difference

10. About how much did you make today at the market? Please check the appropriate range.

- 0% \$0
- 10% \$1-\$99
- 15% \$100-\$199
- 20% \$200-\$299
- 25% \$300-\$399
- 15% \$400-\$499
- 10% \$500-\$749
- 0% \$750 or more
- 5% Don't Know

11. Are you able to accept WIC or senior vouchers? 45% Yes 55% No

My last questions are about you. The information will be used to help classify your answers.

12. What is your age?

- 0% Under 18
- 15% 18 - 24
- 30% 25 - 34
- 20% 35- 44
- 25% 45 - 54
- 5% 55 - 64
- 5% 65 or older

13. What best describes your race/ethnicity? (Check all that apply.)

5% American Indian, Eskimo or Aleut 0% Black or African American 75% Hispanic or Latino
15% Asian or Pacific Islander 0% White or Caucasian 5% Other_____

14. How many acres do you farm?

25% 3 acres or less 40% 4 to 6 acres 25% 7 to 9 acres
10% 10 acres or more 0% Don't know

15. Are you employed full-time or part-time as a farmer/food producer?

65% Full time 35% Part Time

16. What is your home zip code? _____

17. Do you have any additional ways this market could help to enhance your work as a farmer? (Open-ended responses coded during analysis. Respondents could give more than one answer.)

35% More information on value-added products
10% More courses on small business finance
30% More advertising in local newspapers, radio and media
30% Other
15% More information on purchasing and using EBT machines

Appendix XIII: Communication Materials

- Press Release
- Newsletters
- Annual Report
- Grant Proposal

Sample of a press release for a Fictitious Farmers' Market

For Immediate Release Contact: Juana Venegas, Fresh Market, (310) 494-1700

FRESH MARKET HELPS LOCAL FARMERS –
SURVEY SHOWS TWICE WEEKLY MARKET PROVIDES INCOME NEEDED FOR FARMERS TO SUSTAIN

ANYTOWN – July 14, 2003 – The Saturday market isn't just a great place to stroll and collect fresh produce for dinner. A new survey of farmers who sell their products at the market shows that the Fresh Market is helping to keep them going in these tough economic times. Two-thirds of farmers report that the Fresh Market provides important income for their families – 65% say the market makes a “big difference” in helping them make a living as a farmer. Most of the farmers make over \$300 on a market day.

But farmers get more than money from marketing their goods at Fresh Market. Over half of the farmers report that they have developed new products and broadened their customer base. Even more important, perhaps, about 75% of the farmers report that they feel better about their future as farmers.

Most Fresh Market farmers tend small fields – between four and ten acres – trucking their goods about 40 miles to the market. They sell mostly fruits and vegetables, but their harvest also includes cheese, breads and other baked goods, jellies, herbs, spices and cultural foods.

The survey is part of a comprehensive evaluation program being undertaken by Fresh Market to assess and improve the quality of the services it provides to farmers and their clients and the goods sold at the twice-a-week market.

“Fresh Market is a business just like any other,” said Venicio Machado, executive director. “We don't just set up our stands and hope for the best. Our market research gives us the information we need to provide great customer service,” he added.

Samples of newsletters for a Fictitious Farmers' Market

On the following pages are examples of newsletters for Fresh Market farmers and for the Fresh Market community. All examples are fictional.

Fresh Market Survey Results

A Report to Fresh Market Farmers

July 2003



Fresh Market

Thank you for your participation!

The Fresh Market would like to thank you for your participation in the merchant survey. Your feedback is valuable and we appreciate you taking the time to help. With your input, we hope to better serve the community and the visitors of the market, as well as provide you with the knowledge and opportunities you require to produce and sell quality products at Fresh Market. Here is a summary of the survey results.



Survey Findings By Juana Venegas

In a recent survey of Fresh Market Farmers, **80% of the respondents "strongly agree"** that, because they operate a stand at this market, **they have earned more income from farming** and have been able to develop a larger customer base.

The majority reported that the market helps to bring in more than \$300 every day. (See graph at right.) Almost two-thirds of our farmers said the Fresh Market has made a "big difference" in helping them make a living at farming and/or agriculture.

Farmers who operate stands said they bring their produce to the market once a week. Mainly selling fruits and vegetables, some also sell traditional or cultural foods, cheese, eggs and baked goods.

The market has also helped our farmers further develop their skills, knowledge and products. Forty-five percent of our respondents have started to grow new products, while one-quarter said they now know more about running a farm and managing farm sales.

Respondents also told us that they are using more organic methods of farming, including using less pesticides and chemicals due to customer requests.

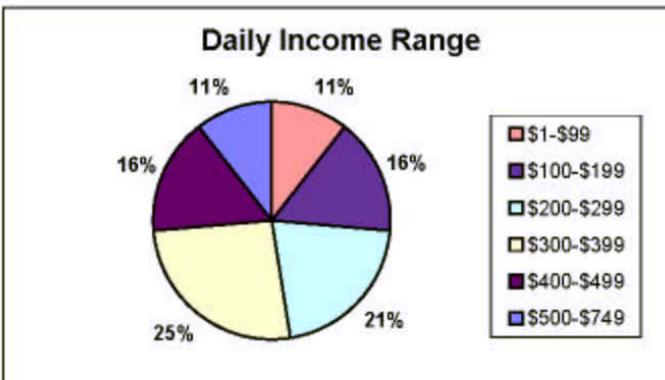
By gaining this knowledge and the income, **60% "strongly agree" that they "feel better about their future in farming."** Half of our farmers feel they are "more able to provide food for their families and themselves."

Farmers who operate stands at the market mainly tend to smaller fields - about 4 to 6 acres in size. Sixty-five percent of our respon-

dents are full-time food producers. We found they travel, on average, 84 miles round-trip to bring their products to the market. Although farmers at our market have a fairly long drive, they told us they feel that one of the best things about selling at this market is the accessibility to their customers. (See graph on page 2.)

Overall, 70% rated the quality of products sold at the market as "excellent." The majority of our respondents also feel their experiences selling at the Fresh Market

(Continued on page 2)



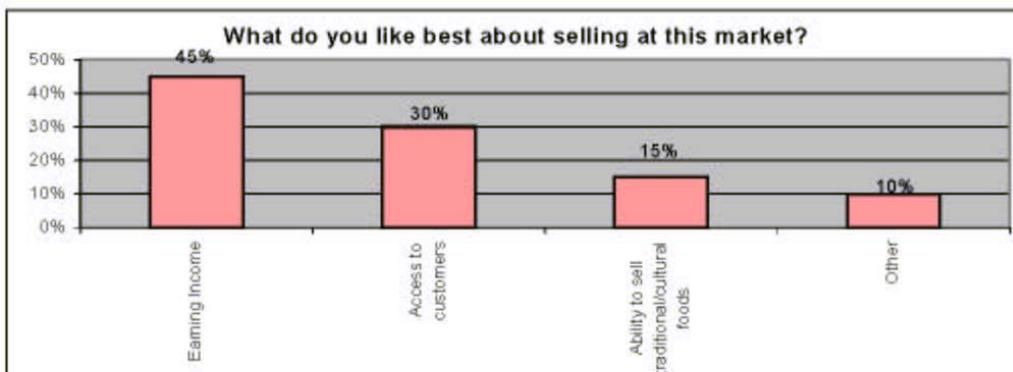
Issue 1

July 2003

(Continued from page 1)
have been "excellent."



The survey is part of a comprehensive evaluation program being undertaken by Fresh Market to assess and improve the quality of the services it provides to farmers and their clients, and the goods sold at the twice-a-week market.



Additional Resources for Merchants and Farmers

When surveyed at the Fresh Market, merchants were asked if they would suggest "any additional ways this market could help to enhance your farming career." Thirty-five percent of the respondents would like more information on value-added products, 30% wanted more advertising in local media, 15% wanted more information on purchasing and EBT machines and 10% wanted more courses on small business finance. The following list of materials provides more information on each of these topics.

- **The Agricultural Marketing Resource Center (AgMRC)** (www.agmrc.org) - This is an electronically based center to create and present information about value-added agriculture. It provides information on products and commodities, value-added markets and industries, creat-

ing, developing and operating a value-added agricultural business, along with other directories and resources. For more information on value-added agriculture in the state of New York, contact: Jessica A. Chittenden, Director of Communications, NYS Department of Agriculture and Markets. One Winners Circle, Albany, NY 12235. Phone: (518) 457-3136; Fax: (518) 457-3087; Email: JessicaChittenden@agmkt.stat.ny.us.

- **AgriSurf** (www.agrisurf.com) - This is a great resource for tracking down agricultural-based Web sites. This site contains a database with over 20,000 agricultural sites. This site also provides a weekly mail-cast publication, *The Agrisurfer*, that will keep you up-to-date on things that are happening on the Web agriculturally.

- **The Library of Congress, The Entrepreneur's Reference Guide to Small Business Information** (www.loc.gov/tr/business/guide/guide2.html) - This is an on-line guide that provides "how to" books, reference books and other useful tools that would assist in starting up a new business. Many of the books, references and journals are available at your local public or university library, or local book dealers.

- **KGLS, Connersville** (www.kgls.com) - For advertising, contact Elle Daniels at (310) 555-9375.

- **The Connersville Cannon** (www.connersvillecannon.com) - Contact Joseph Hernandez, Advertising Department, at (310) 555-5839.



This newsletter is edited and produced by:
Fresh Market · 1234 Lily Lane · Connersville, New York 54321
Phone: (310) 494-1700



Fresh Market

Fresh Market Update

Issue 1 A Report to Community Members July 2003

The Facts on Fresh Market By Venicio Machado

The Fresh Market has been in operation since the October 1996. The goals of the market have been to: 1) help community residents gain access to fresh, local produce, and 2) help immigrant farmers in the area sell their products and increase their knowledge about farming and agriculture. Bonita Rodriguez founded the market after *El Mercado* grocery store closed, leaving community residents with little access to fresh produce or traditional foods. This market, along with many others, is funded by the U.S. Department of Agriculture to increase community food security in urban areas.



The Fresh Market operates at the corner of 13th and Lily, Wednesdays and Saturdays from 10:00 a.m. to 2:00 p.m. The market primarily serves residents in the Dos Hermanos neighborhood. Around 300 to 400 residents visit Fresh Market each week.

Market Provides Extra Income for Local Farmers By Juana Venegas

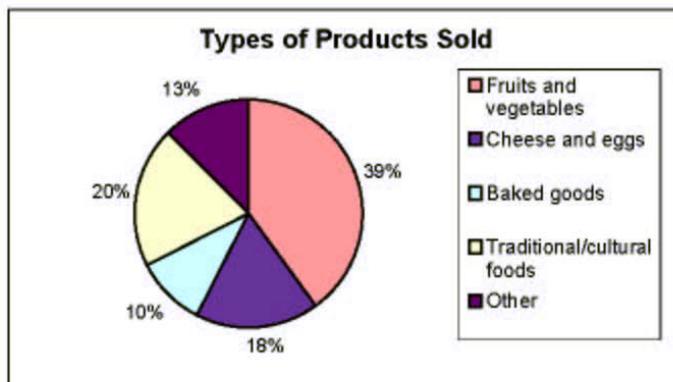
The Fresh Market has helped local farmers during these tough economic times. In a recent survey, **80% for farmers selling at the market said they "strongly agree"** that, because they operate a stand at this market, **they have earned more income from farming** and have been able to develop a larger customer base.

The majority of farmers reported that the market helps them bring

in more than \$300 every day. Almost two-thirds of the farmers who responded to the survey said the Fresh Market has made a "big difference" in helping them make a living at farming and/or agriculture.

The merchants mainly sell fruits and vegetables but they also sell traditional or cultural foods, cheese, eggs and baked goods. (See graph below.)

The market has also helped the farmers to further develop their skills, knowledge and products. Forty-five percent of farmers have started to grow new products, while one-quarter of the farmers said they now know more about running a farm and managing farm sales. Merchants at the market also said they are using more organic methods of farming, which includes using less pesticides and chemicals.



By gaining both knowledge and income, **60% of the merchants "strongly agree" that they "feel better about their future in farming."** Half of the farmers feel they are "more able to provide food for their families and themselves."

Farmers who operate stands at the market mainly tend to smaller fields - about 4 to 6 acres in size. Sixty-five percent of farmers at

(Continued on page 2)

(Continued from page 1)

the Fresh Market are full-time food producers. They travel, on average, 84 miles round-trip to bring their products to the market. Although they have a fairly long drive, the farmers feel that one of the best things about selling at this market is the accessibility to customers.

Overall, 70% of the farmers rated the quality of products sold at the market as "excellent." The majority also feels their experiences selling at the Fresh Market have been excellent.

The survey is part of a comprehensive evaluation program being undertaken by Fresh Market to assess and improve the quality of the services it provides to farmers and their clients, and the goods sold at the twice-a-week market. Survey responses are based on 20 out of 24 farmers who sell at the market.



Wide varieties of fruits and vegetables can be seen at the Fresh Market throughout the year.

Additional Resources

The following resources provide more information regarding open-air and farmers' markets, the USDA, and general agricultural information.

If you have questions regarding any of the information given, or would like more information, please contact Juana Venegas at (310)494-1700.

- **The United States Department of Agriculture (USDA), Food and Nutrition Service** (<http://www.fns.usda.gov/>)

Fresh Market Information:

- The Fresh Market hours of operation are Wednesday and Saturday: 10:00 a.m. to 2:00 p.m. (WIC and senior vouchers accepted.)
- Located on the corner of 13th and Lily downtown.
- For questions concerning the market or for more information, please contact Juana Venegas at (310) 494-1700.

fsec/) - The USDA defines food security as "access by all people at all times to enough nutritious food for an active, healthy life." This site will give you more information about funding for programs like the Fresh Market has received. This site will help you take action on food security issues.

- **Community Food Security Coalition (CFSC)** (<http://www.foodsecurity.org/>) - CFSC is a North American organization dedicated to building strong, sustainable, local and regional food systems that ensure access to affordable, nutritious, and culturally appropriate food for all people at all times. They seek to develop self-reliance among all communities in obtaining their food and to create a system of growing, manufacturing, processing, making available, and selling food that is regionally based and grounded in the principles of justice, democracy, and sustainability. The Web site provides links to community gardens/urban agriculture, nutrition, sustainable agriculture, community development, farmers' markets, food access and anti-hunger Web sites.

- **Farmers' Markets** (<http://www.ams.usda.gov/farmersmarkets/map.htm>) - This site contains a map of the United States. To get a list of farmers' markets in your state, simply click on the state.
- **The Food Guide Pyramid** (<http://www.usda.gov/cnpp/pyrabklt.pdf>) - This booklet discusses the Food Pyramid and the research guided its development. It also suggests ways to apply these dietary guidelines in everyday life.



Farmers in up-state New York load up their trucks as they prepare for the market.

This newsletter is edited and produced by:
 Fresh Market . 1234 Lily Lane . Connersville, New York 54321
 Phone: (310) 494-1700

Sample of an annual report excerpt for a Fictitious Farmers' Market

Fresh Market

Evaluation Summary for Annual Progress Report to J.K. Funder Foundation

Our preliminary evaluation efforts demonstrate the success of our Fresh Market project to build self-reliance among our local farming community. As our original goals, the Fresh Market project set out to “help immigrant farmers in the area sell their products and increase their knowledge about farming and agriculture.”

To address these goals, our Fresh Market project staff administered a survey in September 2003 to immigrant farmers who operate stands at the Fresh Market. (A full report of our survey results is provided in this brief.) By analyzing data from these surveys, we found that we were successfully meeting our goals. In addition, we have solid information now to direct improvements to our project.

All of the farmers we surveyed (n=20) reported developing a larger customer base because of our Fresh Market project while 89% said they felt better about their future in farming because of their involvement with Fresh Market. Individual immigrant farming families also prospered. 85% of the farmers surveyed said they increased their income because of the Market (half reported making over \$300 per day of operation) and 83% said that they have a greater ability now to provide food for their families.

The continuing educational component of our Market was also a hit: 100% of farmers surveyed said they learned more about running a small business; 75% said they have developed new products, 65% have learned new farming skills and 63% learned more about organic farming.

The results of our Fresh Market survey demonstrate that this project not only achieves its individual goals, it also supports the goals of the United States Department of Agriculture (USDA) by supporting entrepreneurial development and promoting comprehensive food, farm and nutrition responses. As our program continues to succeed, staff want to additionally demonstrate how Fresh Market is representative of best practices in community food security and hope to expand the evaluation to further show this success. Further, in our survey, farmers identified a variety of ways they would benefit from additional Market programs. Farmer respondents said they would like more information on value-added products; more advertising in local newspapers, radio and media about the market; more information on purchasing and using EBT machines; and more courses on small business financing. We hope the J.K. Funder Foundation and the USDA will help us achieve these additional goals. Sustaining the positive influence of our program while attending to the identified needs of our community farmers will also help the USDA attain its national vision and goals.

Sample of a grant proposal excerpt for a Fictitious Farmers' Market

Fresh Market Summary of Evaluation Findings for New Grant Proposal

Evaluation of Program to Date

Evaluation has been an important tool in the design and the operations of our program. We also use the data to demonstrate to funders that their dollars are well spent. Our most recent survey of farmers (September 2003) was used to help measure our success at meeting the goal of "helping immigrant farmers in the area sell their products and increase their knowledge about farming and agriculture". Selected results from a recent evaluation of the Fresh Market are submitted in support of this report.

Impact of the Market on Farmers

Survey respondents (n=20) who operate a stand at this market:

- Develop a larger customer base (100%)
- Learn more about running a small business (100%)
- Feel better about my future in farming (89%)
- Earn more income from farming (85%)
- Are able to provide food for their families (83%)
- Have developed new products (75%)
- Have learned new farming skills (65%)
- Have learned more about organic farming (63%)

In addition, more than 50% of the farmers earned at least \$300 per day at the market. Thus, our participants report that Fresh Market has helped to develop self-reliance as they grow, manufacture, process, and market food that is locally grown.

The survey was also used to assess unmet needs of the participants. Farmers identified ways in which the market could further enhance their work in farming, including:

- More information on value-added products
- More advertising in local newspapers, radio and media
- More information on purchasing and using EBT machines
- More courses on small business financing

As we discussed earlier in the proposal, we are seeking additional funds to not only continue our current work with the farmers but also enhance the program's marketing and outreach as well as provide additional training opportunities to farmers. These data are instrumental in determining how to enhance our marketing and education workshops.

With the increasing diversity in our immigrant communities, building strong, sustainable, food system that can ensure access to affordable, nutritious and culturally appropriate food for all people has become a vital component of the local economy. Projects like the Fresh Market are essential if the goal of developing new food systems is going to be realized. Fresh Market strives to achieve the USDA's goals locally by

- Meeting the food needs of low-income people
- Increasing community self-reliance
- Promoting comprehensive food, farm and nutrition responses
- Developing innovative links (between farmers and their community)
- Supporting entrepreneurial development, and
- Achieving project self-sufficiency

More research and evaluation about the development of sustainable food systems is needed. We are requesting financial support to continue the positive influence of this program, attend to emerging needs and expand our evaluation component for the project.

Appendix XIV: Glossary

Activities – What the program does with its resources to fulfill its mission

Analysis Plan – Usually contains the evaluation question(s) to be answered, the sources of data that will be used to answer those questions, and the analysis to be performed to answer the questions.

Assumed Causes – Assumptions about the factors contributing to the community need.

Bias – Any influence that distorts the results of a research study.

Codebook – Provides a layout of how your data will be entered into a file created by the software program. It serves as a reference guide for understanding your data printout.

Community Context – The conditions or events in the program, community or target population that may limit or expand the extent to which the program achieves its desired outputs or outcomes.

Community Need – Statement describing the community need your program will address. Needs may also be defined in terms of assets to be strengthened.

Composite Description – A statement that describes the responses in each of the categories used to classify qualitative data.

Control – Processes employed to hold the conditions under which an investigation is carried out uniform or constant. In a true experimental design, the control group is the group that does not receive the project's services. The outcomes are then compared for the control and the group receiving the project's services. In other experimental designs, this group may be referred to as the comparison group.

Descriptive Statistics – A numerical description used to summarize a large amount of data.

Emergent Categories – Categories used to classify qualitative data that are not identified in advance and which become apparent after reading through the raw data.

Focus Group – An interview conducted with a small group of people to explore their ideas on a particular topic.

Formative Evaluation – Evaluations that describe how a program's activities might be improved are called formative evaluations. Formative evaluations ask "What is it?" and "How does it work?" They often occur during early stages of a program because they provide feedback and allow for changes in the program.

Frequency – The percent of scores falling into each response category.

Inferential Statistics – Inferential statistics are used to test the hypothesis(es). They provide conclusions that extend beyond the data. Inferential statistics make inferences from the sample about the population from which it was drawn.

Informed Consent – The process of obtaining voluntary participation of individuals in research based on a full understanding of the possible benefits and risks.

Interrater Reliability (inter-observer reliability) – A measure of the consistency between the ratings/values assigned to an attribute that is being rated or observed; it is usually expressed as a percentage of agreement between two raters/observers or as a

coefficient of agreement that may then be expressed as a probability. Usually employed by researchers using structured observation techniques.

Likert Scale – A method used to measure attitudes, which involves respondents indicating their degree of agreement or disagreement with a series of statements. Scores are summed to give a composite measure of attitudes.

Long-Term Outcomes – Changes in individual or group behavior of community conditions that a program hopes to achieve over time. They indicate a measurable change in participant knowledge, attitudes or behavior.

Mean – The sum of all scores divided by the number of scores summed.

Median – The score that is halfway between the lowest and highest value when all the scores are listed in ascending order.

Method to Address Need – The program, its mission and its goals.

Outcome Indicator – An indicator is the specific information that will determine how well the program is doing at meeting its outcome goal.

Outputs – The direct products of program activities (the quantification of activities).

Percentile – The value that falls below a given percent of the scores.

Performance Standard – The level or amount of change that is expected to be achieved in an indicator.

Population – A well-defined group or set that has certain specified properties (e.g., all adult Latina women enrolled in project activities).

Predetermined Categories – Categories defined in advance of classifying qualitative data.

Program Support Activities – Activities that provide the infrastructure necessary to provide quality activities.

Qualitative Data – Information gathered in narrative (nonnumeric) form (e.g., a transcript of an unstructured interview).

Quality Outputs – Outputs produced at a level reflecting “quality” or “efficiency.”

Quantitative Data – Information gathered in numeric form.

Randomization – The random assignment of participants to the experimental project activity and control groups (i.e., the allocation to groups is determined by chance).

Randomized Controlled Trial (RCT) – In a RCT, participants are randomly assigned either to your project activity (e.g., a program on leadership skills) or to a control group (e.g., no program). Both groups are followed up over a specified period of time and the effects of the program on specific outcomes defined at the outset are analyzed (e.g., increased leadership skills).

Reliability – Reliability is concerned with the consistency and dependability of a measuring instrument, i.e., it is an indication of the degree to which it gives the same answers over time, across similar groups and irrespective of who administers it. A reliable measuring instrument will always give the same result on different occasions assuming that what is being measured has not changed during the intervening period.

Research Method – Specific procedures used to gather and analyze research data.

Research Methodology – Different approaches to systematic inquiry developed within a particular paradigm with associated epistemological assumptions (e.g., experimental research, grounded theory, ethn methodology).

Resources – Resources dedicated to or consumed by the program.

Response Rate – The proportion (percentage) of those invited to participate in a research study that actually do so.

Sampling – There are several types of sampling, including:

Random - A process of selecting a sample whereby each member of the population has an equal chance of being included.

Convenience - A non-probability sampling strategy that uses the most easily accessible people (or objects) to participate in a study.

Systematic - A probability sampling strategy involving the selection of participants randomly drawn from a population at fixed intervals (e.g., every 20th name from a sampling frame).

Sampling Bias – Distortion that occurs when a sample is not representative of the population from which it was drawn.

Secondary Data – Data collected by others.

Short-Term Outcomes – The direct result of your program activities. They indicate a measurable change in participant knowledge, attitudes or behavior.

Standard deviation – The square root of the sum of all squared deviations around the mean divided by the number of deviations summed minus one.

Statistic – An estimate of a parameter calculated from a set of data gathered from a sample.

Statistical Analysis – Most statistical analysis is based on the principle of gathering data from a sample of individuals and using those data to make inferences about the wider population from which the sample was drawn.

Structured Interview – The interviewer asks the respondents the same questions using an interview schedule – a formal instrument that specifies the precise wording and ordering of all the questions to be asked of each respondent.

Subjects – A term most often used in experimental research to describe those who participate in research and provide the data.

Summative Evaluation - Evaluations that demonstrate what a program has accomplished are called summative evaluations. Summative evaluations ask, “Does it work?” They often take place once a program is fully in place.

Survey Research – A research approach designed to systematically collect descriptions of existing phenomena in order to describe or explain what is going on; data are obtained through direct questioning of a sample of respondents.

Target Population – The characteristics of the participant population you serve or intend to serve.

Test-retest Reliability – A means of assessing the stability of a research instrument by calculating the correlation between scores obtained on repeated administrations.

Unstructured Interview – The researcher asks open-ended questions which give the respondent considerable freedom to talk freely on the topic and to influence the

direction of the interview since there is no predetermined plan about the specific information to be gathered from those being interviewed.

Validity – In research terms, validity refers to the accuracy and truth of the data and findings that are produced. It refers to the concepts that are being investigated; the people or objects that are being studied; the methods by which data are collected; and the findings that are produced.

Endnotes

- 1 Rossi, P. H., and H. E. Freeman (1993). *Evaluation: A Systematic Approach*. Thousand Oaks, CA: Sage Publications, Inc. To order this textbook on evaluation, visit: www.sagepub.com.
- 2 *Using Social Research in Public Policy Making*, edited by Weiss. C. A. Lexington, MA: D.C. Heath. As quoted in *Utilization-Focused Evaluation*. (1997), 7. Patton, M.Q. Thousand Oaks, CA: Sage Publications, Inc.
- 3 US Department of Health and Human Services. (2002). *Physical Activity Evaluation Handbook* (p. 6). Atlanta, GA:
US Department of Health and Human Services, Centers for Disease Control and Prevention.
(<http://www.cdc.gov/nccdphp/dnpa/physical/handbook/pdf/handbook.pdf>)
- 4 Fetterman, David. (2001). Welcome to my virtual office. American Evaluation Association. Retrieved October 2002, from the World Wide Web: <http://www.stanford.edu/~davidf/#Collaborative>.
- 5 Bureau of Justice Assistance Evaluation. (December 13, 2002). BJA Evaluation Web Site. Bureau of Justice Assistance Evaluation. Retrieved October 2002, from the World Wide Web:
http://www.bja.evaluationwebsite.org/html/roadmap/beyond/impact_evaluation/what_is_impact.html
- 6 The Joint Committee on Standards for Educational Evaluation. (1994). *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.
- 7 Coffman J. (1999). *Learning from Logic Models: An Example of a Family/School Partnership Program*. Cambridge, MA: Harvard Family Research Program.
- 8 Logic model guide. (October 2000). W.K. Kellogg Foundation.
- 9 McLaughlin JA and Jordan GB. (July 1998). *A Tool for Telling Your Program's Performance Story*. US Department of Energy, Office of Energy Efficiency and Renewable Energy.
- 10 McLaughlin JA and Jordan GB. (July 1998). *A Tool for Telling Your Program's Performance Story*. US Department of Energy, Office of Energy Efficiency and Renewable Energy.
- 11 *Building a Successful Prevention Program*. Western Regional Center for the Application of Prevention Technologies. Center for Substance Abuse Prevention.
- 12 *Measuring Program Outcomes: A Practical Approach*. (1996). United Way of America,
- 13 United Way breaks outcomes into 3 categories: initial, intermediate and long-term.
- 14 Adapted from: Logic model guide. (October 2000). W.K. Kellogg Foundation.
- 15 Adapted from: Logic model guide. (October 2000) W.K. Kellogg Foundation.
- 16 Focusing on the assets of a community, many CFPs work from a model known as asset-based community development (ABCD). This model approaches community problem solving by identifying the community's assets rather than its needs. The ABCD model can be incorporated into the development of a logic model by including the community's assets within the "resources" section of the logic model as well as within the "Community Context" and "Program Support Activities" sections. For more information on the ABCD model, see <http://www.northwestern.edu/ipr/abcd.html>.
- 17 Fetterman, David. (2001). Welcome to my virtual office. Retrieved October 2002, from the World Wide Web: <http://www.stanford.edu/~davidf>.
- 18 Cohen, Barbara. (July 2002). U.S. Department of Agriculture. Economic Research Services. USDA Community Food Security Assessment Toolkit. (page 3). Retrieved November 2002, from <http://www.ers.usda.gov/publications/efan02013/>.
- 19 Harrel, A. *Evaluation Strategies for Human Service Programs*. Washington, DC: The Urban Institute.
- 20 Martin, L.L. & Kettner, P.M. (1996). *Measuring the Performance of Human Service Programs*. California: Sage Publications, Inc.
- 21 Coffman J. (1999). *Learning from Logic Models: An Example of a Family/School Partnership Program*. Cambridge, MA: Harvard Family Research Program.
- 22 *Measuring Program Outcomes: A Practical Approach*. (1996). United Way of America.
- 23 Adapted from: Logic model guide. (October 2000). W.K. Kellogg Foundation.
- 24 Patton, M.Q. (1997). *Utilization-Focused Evaluation*. Thousand Oaks, CA: Sage Publications, Inc.
- 25 N. K. Young, S. L. Gardner, & S. M. Coley. (1994). *Getting to Outcomes in Integrated Service Delivery Models*. In *Making a Difference: Moving to Outcome-Based Accountability for Comprehensive Service Reforms*. Falls Church, VA: National Center for Service Integration Information Clearinghouse.
- 26 *Report on Lessons Learned in the Pilot Phase of the United Way Outcomes Project*. (1995). United Way of Greater Milwaukee, Inc. (pp 6-13).
- 27 Adapted from (Donobedian, 1992)
- 28 Adapted from Logic model guide. (October 2000). W.K. Kellogg Foundation.
- 29 Prochaska, J.O. Norcross, J.C. & DiClemente, C.C. (1995). *Changing for good*. Avon Books.

- 30 Cohen, Barbara. (July 2002). U.S. Department of Agriculture. Economic Research Services. USDA Community Food Security Assessment Toolkit. (page 3). Retrieved November 2002, from <http://www.ers.usda.gov/publications/efan02013/>.
- 31 McLeroy K, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. *Health Education Quarterly* 1998;15:351-377.
- 32 Adapted from Logic model guide. (October 2000). W.K. Kellogg Foundation.
- 33 Although a common goal of food security program is to change the life conditions of a larger community, services are generally applied to specific individuals within a community. Unless a program significantly impacts each member of the community, it is likely that program impacts will be weakened because “untreated” people will be thrown into the mix with others receiving your services.
- 34 Community level outcomes or indicators which are better covered in the Community Food Assessment Toolkit include the broader concepts of increasing community food security, improving the farm to market service delivery, creating local jobs, strengthening the food system, increasing the number of community supported stores, etc.
- 35 Augmentation of Bill Phillips in Report on Lessons Learned in the Pilot Phase of the United Way Outcomes Project. (June 1995). United Way of Greater Milwaukee, Inc., 5.
- 36 Adapted from *Creating a Culture of Inquiry*. (July 2001). Hernandez, G. & Visher, M.G. The James Irvine Foundation.
- 37 Adapted from: US Department of Health and Human Services. (2002). *Physical Activity Evaluation Handbook*. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention.
- 38 US Department of Health and Human Services. (2002). *Physical Activity Evaluation Handbook*. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention.
- 39 McDonald, D, Kutara, P, et al. (December 2004). Culturally Respectful Evaluation. In *The Forum for Family and Consumer Issues* 9(3). NC State University. Retrieved online on December 5, 2005: http://www.ces.ncsu.edu/depts/fcs/pub/9_3/cultural.html
- 40 McClure, C. (1996). *Assessing the Academic Networked Environment: Strategies and Options*. School of Information Studies, Syracuse University. Retrieved from http://istweb.syr.edu/~mcclure/network/Part_II.html
- 41 It should be noted that there is significant variation in opinion in the field when it comes to describing qualitative methods, designs and purposes and the distinctions between qualitative and quantitative methods.
- 42 This table adapted from the work of J. Shcheurich, Department of Educational Administration, College of Education, University of Texas at Austin <http://www.edb.utexas.edu/faculty/scheurich/proj1/inquiry.html>. The Royal Windsor Society for Nursing Research http://www.windsor.igs.net/~nhodgins/quant_qual.html. Adapted from Burns and Grove, 1993; Streubert and Carpenter 1995; D. Ratcliff, Department of Psychology, Biola University <http://don.ratcliff.net/qual/qualquant.html>, adapted from Cook and Reichardt, 1979; R. Anderson, Department of Sociology, University of Oxford, <http://www.crest.ox.ac.uk/andersen/methods/methods3.pdf>.
- 43 McClure C. (1996). *Assessing the Academic Networked Environment: Strategies and options*. School of Information Studies, Syracuse University. Retrieved November 2002 from the World Wide Web: http://istweb.syr.edu/~mcclure/network/Part_II.html
- 44 It is important to understand that many facets influence research results. Thus, the word “prove” is in quotations because professional researchers generally agree that regardless of the sophistication and quality of the experiment, results are never unequivocally considered “proven.”
- 45 Frechting J. & Sharp L (Eds.). (1997). *User Friendly Handbook for Mixed Method Evaluations*. Division of Research, Evaluation and Communication, National Science Foundation. Retrieved March 2002, from the World Wide Web: http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/CHAP_3.HTM.
- 46 Frechting J. & Sharp L. (Eds.). (1997). *User Friendly Handbook for Mixed Method Evaluations* Division of Research, Evaluation and Communication, National Science Foundation. Retrieved March 2002. from the World Wide Web: http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/CHAP_3.HTM.
- 47 The figure below is based on Frechting J. & Sharp L. (Eds.). (1997). *User Friendly Handbook for Mixed Method Evaluations* Division of Research, Evaluation and Communication, National Science Foundation. Retrieved March 2002, from the World Wide Web: http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/CHAP_3.
- 48 Smith, R. (1996). Capturing Sensitive Data from Young People in a Household Setting. *Journal of the Market Research Society*, 38(2), 177-183.
- 49 For more information on the asset-based community development (ABCD) model see <http://www.northwestern.edu/ipr/abcd.html>.
- 50 Krueger, R. A. and Casey, M. A. *Focus Groups: A Practical Guide for Applied Research*, 3rd Edition. Sage Publications. Thousand Oaks, CA, 2000.
- 51 Frechting J. & Sharp L. (Eds.). (1997). *User Friendly Handbook for Mixed Method Evaluations* Division of Research, Evaluation and Communication, National Science Foundation. Retrieved March 2002, from the World Wide Web: http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/CHAP_3.
- 52 Professionally conducted focus groups are held in a focus group facility that includes a one-way mirror for observation. The sessions are usually filmed as well as recorded. Participants are often paid for attendance and provided with a meal or snack.

- 53 Frechting J. & Sharp L. (Eds.). (1997). User Friendly Handbook for Mixed Method Evaluations Division of Research, Evaluation and Communication, National Science Foundation. Retrieved March 2002, from the World Wide Web: http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/CHAP_3.HTM.
- 54 Guba EG. & Lincoln YS. (1981). Effective Evaluation. San Francisco: Jossey-Bass.
- 55 Lincoln, YS. & Guba EG. (1985). Naturalistic Inquiry. Thousand Oaks, CA: Sage Publications, Inc.
- 56 Frechting J. & Sharp L. (Eds.) (1997). User Friendly Handbook for Mixed Method Evaluations Division of Research, Evaluation and Communication, National Science Foundation. Retrieved March 2002, from the World Wide Web: http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/CHAP_3.HTM.
- 57 A 'rubric' is a set of rules or guidelines that helps set criteria for making judgments across subjects.
- 58 In fact, studies have shown that kids prefer taking surveys on the computer when compared to paper and pencil tests. Vispoel WP, Boo J. & T Bleiler (2001). Computerized and paper-and pencil versions of the Rosenberg self-esteem scale: a comparison of psychometric features and respondent preferences. Educational and Psychological Measurement, 61, 461-474 and Supple AJ, Aquilino, WS. & DL Wright (1999). Collecting sensitive self-report data with laptop computers: impact on the response tendencies of adolescents in a home interview. Journal of Research on Adolescence, 9, 467-488.
- 59 Issues with literacy and language can be addressed in self-administered questionnaires by producing multiple versions/translations.
- 60 Public records.
- 61 Personal records.
- 62 Chin, JL. (1999, October 25). Cultural Competence in a Multicultural Society, A Checklist for Cultural Competence Self-Assessment). Retrieved April 2002, from the World Wide Web: <http://shpm.com.com/articles/cultural/culture.html>.
- 63 LaFrance, J. (2001, November). Non-Western Evaluation: Alternative Thoughts and Practices. Presented at the meeting of the American Evaluation Association, St. Louis, MO.
- 64 For more information on validity or reliability, refer to any research methods textbook such as: Stanley, J. & Hopkins, K. (1972). Educational and Psychological Measurement and Evaluation. New Jersey: Prentice Hall, 1972.
- You may also consult the following Web-based documents and sites:
[Bill Trochim's Center for Social Research Methods](http://trochim.human.cornell.edu/). <http://trochim.human.cornell.edu/>
- Rudner, L.M. (1994). Questions to ask when evaluating tests. Practical Assessment, Research & evaluation, 4(2). Available online: <http://ericae.net/pare/getvn.asp?v=4&n=2>
- The Evaluation Center at Western Michigan University's Glossary of Evaluation Terms, <http://ec.wmich.edu/glossary/>
- 65 Guba EG., & Lincoln, Y S. (1989). Fourth Generation Evaluation. Newbury Park, CA: Sage Publications, Inc.
- 66 Sedlacek, W. & Kim S. (1995, January 1). Multicultural Assessment in ERIC Digest, ED391112.
- 67 In other words, when validity and reliability are established using one group of people, the instrument may not be reliable or valid for use with another group of people.
- 68 American Counseling Association Code of Ethics. (1995). Retrieved on January 8, 2002, from the World Wide Web: <http://www.counseling.org/resources/codeofethics.htm#ce>.
- 69 Sedlacek, W. & Kim S. (1995, January 1). Multicultural Assessment in ERIC Digest, ED391112.
- 70 Caller, S., Betts, S., et.al. (2000). Community-Based Project Evaluation Guide. Retrieved January 10, 2002 from The University of Arizona Institute for Children, Youth & Families Web site: <http://ag.arizona.edu/fcr/fs/cyfar/evalgde.htm>.
- 71 Sedlacek, W. & Kim S. (1995, January 1). Multicultural Assessment in ERIC Digest, ED391112.
- 72 Johnson, E. (2001, November). Walking the Path Toward Multiculturally Competent Evaluation: What Does It Take? Paper presented at the meeting of the American Evaluation Association, St. Louis, MO.
- 73 Johnson, E. (2001, November). Walking the Path Toward Multiculturally Competent Evaluation: What Does It Take? Paper presented at the meeting of the American Evaluation Association, St. Louis, MO.
- 74 McDonald, D, Kutara, P, et al. (December 2004). Culturally Respectful Evaluation. In The Forum for Family and Consumer Issues 9(3). NC State University. Retrieved online on December 5, 2005: http://www.ces.ncsu.edu/depts/fcs/pub/9_3/cultural.html
- 75 Sedlacek, W. & Kim S. (1995, January 1). Multicultural Assessment in ERIC Digest, ED391112.
- 76 Chang, C. (2001, May 3). Cross-Cultural Assessment: A Call for Test Adaptation. Retrieved on January 9, 2002, from the World Wide Web: Association for Assessment in Counseling website, <http://aac.ncat.edu/newsnotes/y99sum1.html>.
- 77 Marin G., & Marin. B.V. (1991). Research with Hispanic populations. Newbury Park, CA: Sage Publications, Inc.
- 78 This table is based on information adapted from: Miller, T.I. & M.A. Miller. (2002, 2nd ed.). Citizen Surveys: How to Do Them, How to Use Them, What They Mean, 80-88. ICMA: Washington DC.
- 79 Adapted from Earthman, E., Richmond, L. S. et al. (October 1999). Adapting Evaluation Measures for 'Hard to Reach' Audiences in CYFERNet. Tucson, AZ: Children, Youth and Families Education and Research Network, The University of Arizona.

- 80 Some of this information was adapted from JSI Research & Training Institute, Inc. (date unknown). SAMSHA Prevention Project Survey Administration Manual, p. 4 - 9. Retrieved on November 6, 2003 from http://www.health.org/sa-hiv/SAMHSA_survey_protocol.doc.
- 81 McDonald, D, Kutara, P, et al. (December 2004). Culturally Respectful Evaluation. In The Forum for Family and Consumer Issues 9(3). NC State University. Retrieved online on December 5, 2005: http://www.ces.ncsu.edu/depts/fcs/pub/9_3/cultural.html
- 82 Some of this information was adapted from ethics guidelines from the American Psychological Association and from this web link provided by the Research Council from Norway: <http://www.etikkom.no/NESH/eretn.htm>.
- 83 Making a Difference: Moving to Outcome-Based Accountability for Comprehensive Service Reform. Resource Brief 7, by Young, Nancy et al. National Center for Service Integration, Falls Church, VA, 1994
- 84 Augmentation of Bill Phillips in Report on Lessons Learned in the Pilot Phase of the United Way Outcomes Project. (June 1995). United Way of Greater Milwaukee, Inc., 5.
- 85 N. K. Young, S. L. Gardner, and S. M. Coley. Getting to Outcomes in Integrated Service Delivery Models. In Making a Difference: Moving to Outcome-Based Accountability for Comprehensive Service Reforms. Falls Church, VA: National Center for Service Integration Information Clearinghouse, 1994.
- 86 Report on Lessons Learned in the Pilot Phase of the United Way Outcomes Project, United Way of Greater Milwaukee, Inc., June 1995.p 6-13.
- 87 A. Donabedian. The Role of Outcomes in Quality Assessment and Assurance. QRB, November 1992, 356-360.
- 88 Corporation for National Service. Toolkit. A User's Guide to Evaluation for National Service Programs.
- 89 Making a Difference: Moving to Outcome-Based Accountability for Comprehensive Service Reform. Resource Brief 7, by Young, Nancy et al. National Center for Service Integration, Falls Church, VA, 1994
- 90 Augmentation of Bill Phillips in Report on Lessons Learned in the Pilot Phase of the United Way Outcomes Project. (June 1995). United Way of Greater Milwaukee, Inc., 5.
- 91 N. K. Young, S. L. Gardner, and S. M. Coley. Getting to Outcomes in Integrated Service Delivery Models. In Making a Difference: Moving to Outcome-Based Accountability for Comprehensive Service Reforms. Falls Church, VA: National Center for Service Integration Information Clearinghouse, 1994.
- 92 Report on Lessons Learned in the Pilot Phase of the United Way Outcomes Project, United Way of Greater Milwaukee, Inc., June 1995.p 6-13.
- 93 A. Donabedian. The Role of Outcomes in Quality Assessment and Assurance. QRB, November 1992, 356-360.
- 94 Adapted from W.K. Kellogg Foundation Evaluation Handbook (January 1998). W.K. Kellogg Foundation.
- 95 For more information on communicating the results, and the types of reporting that can be done or may be appropriate, please see the section in this chapter "Communicating the Results"
- 96 Adapted from: Utah State University Student Support Services Plan to Use Evaluation Results (<http://milo.usu.edu/ssss/eval.htm>).
- 97 Horizon Research, Inc. Taking Stock. A Practical Guide to Evaluating Your Own Programs. Chapel Hill, NC: 1997.
- 98 Center for Substance Abuse and Prevention, Western Center for the Application of Prevention Technologies. Building A Successful Prevention Program. 2002.
- 99 Most resources excerpted from a very comprehensive resource list compiled by Catherine Callow Elwell, see: www.siu.edu/~ritzel/courses/526/evaluation.htm.
- 100 College of Education, San Diego State University. Retrieved January 2003 at http://edweb.sdsu.edu/triton/july/rubrics/Rubric_Template.html,
- 101 Adapted from Behavior Rubric Generator, http://teach-nology.com/web_tools/rubrics/behavior/
- 102 Adapted from the Collaboration Rubric of Cabrillo Tidal Pool Study, Retrieved January 2003 at <http://edweb.sdsu.edu/triton/tidepoolunit/Rubrics/collrubric.html>
- 103 Adapted from Rubenstein, E. (1991, April). An Overview of Adolescent Development, Behavior, and Clinical Intervention. Families in Society: The Journal of Contemporary Human Services (pp220-22).
- 104 Based on articles by Amato, P.A. & Ochilree, G. (1987). Interviewing Children about Their Families: A Note on Data Quality. Journal of Marriage and the Family. 49, (pp 669-675) and Weber LR., Miracle A. & T Skehan, (1994). Interviewing Early Adolescents: Some Methodological Considerations in Human Organization. 53 (1), (pp42-47).
- 105 The first three graphics (smiley faces, strawberries and kids) are from http://www.cbr.com.au/Method_KidsResearch.htm. Example 4 was retrieved April 2002 from <http://www.ces.ncsu.edu/AboutCES/smp/19/stairxo.htm>. Example 5 was retrieved April 2002 from <http://www.ces.ncsu.edu/AboutCES/smp/19/smlots.htm>.