

1996

# Measuring Program Outcomes: A Practical Approach

United Way of America

Follow this and additional works at: <https://digitalcommons.unomaha.edu/slceeval>

 Part of the [Service Learning Commons](#)

---

## Recommended Citation

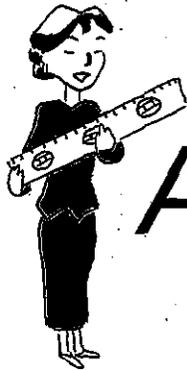
United Way of America, "Measuring Program Outcomes: A Practical Approach" (1996). *Evaluation/Reflection*. 47.  
<https://digitalcommons.unomaha.edu/slceeval/47>

This Report is brought to you for free and open access by the Barbara A. Holland Collection for Service Learning and Community Engagement (SLCE) at DigitalCommons@UNO. It has been accepted for inclusion in Evaluation/Reflection by an authorized administrator of DigitalCommons@UNO. For more information, please contact [unodigitalcommons@unomaha.edu](mailto:unodigitalcommons@unomaha.edu).

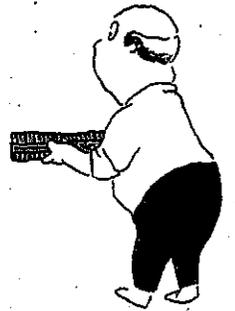




# Measuring Program Outcomes:



*A Practical Approach*



what matters.™



United Way of America

# Measuring Program Outcomes:

*A Practical Approach*



United Way of America

# Acknowledgements

Many individuals and organizations contributed their time, experience, and perspectives to the creation of *Measuring Program Outcomes: A Practical Approach*. Members of United Way of America's Task Force on Impact listed on page v guided the manual's development. Task Force members represent local United Ways, national human service associations, foundations, corporations, and academic and research institutions. The manual reflects their commitment to producing a resource that is both methodologically sound and highly practical.

Of particular value were the detailed comments of Task Force members Sylvia Barsion, James Bell, Stacey Daniels, Gillian Kerr, Larrie Mackie, Myrtis Meyer, Heather Johnston Nicholson, Kathleen Pritchard, James Sanders, Beth Terry, Dan Watanabe, and Joseph Wholey. Their attention enhanced the manual's value as a step-by-step guide and helped assure that its content responds to needs of the human service field.

The Urban Institute was the technical consultant for this project. The outcome measurement expertise of the Urban Institute's project team, Harry P. Hatry and Thérèse van Houten, assured that the manual reflects the state of the art in outcome measurement methodology. Their hands-on experience with diverse service-delivery programs grounded the manual in the day-to-day reality of nonprofit human services.

Program and research directors from 23 national health, human service, and youth- and family-serving organizations, listed on page vii, met with the project's staff and consultants. They shared information and materials related to their program evaluation activities and offered suggestions regarding the manual's content and dissemination. Their input underscored the importance of defining realistic expectations for the implementation of outcome measurement.

Sue Bell interviewed more than 40 local human service program managers in 18 states about their experiences with outcome measurement. Their insights reinforced the manual's emphasis on practical approaches and inclusive processes. The challenges and successes of many of these agencies are highlighted throughout the manual.

Drafts of the manual were distributed to a host of individuals and organizations for review. Thoughtful feedback from reviewers sharpened the final document and helped it maintain a consistent tone and message. Reviewers not listed elsewhere were Tom Curtis, Big Brothers Big Sisters of America, Philadelphia, Pennsylvania; Murray Frost, United Way of the Midlands, Omaha, Nebraska; Jean Gerding, The Gerding Group, Lutherville, Maryland; and George Liacopoulos, St. John's Community Services, Washington, DC.

*Measuring Program Outcomes: A Practical Approach* was written by consultants Harry Hatry and Thérèse van Houten, and by Margaret C. Plantz and Martha Taylor Greenway of United Way of America. Other United Way of America staff, including Elissa Soares, Suzie Gruner, and publishing staff John Libby, Susan Carpenter, Richard Cookson, Beverly McGough, and Norma Reck, provided critical support at many stages. Tamara Patterson of EPI Communications designed the manual and cover. The illustrations were created by Glen Lucas.

# United Way of America Task Force on Impact

## MEMBERS INVOLVED IN DEVELOPMENT OF MANUAL

James Bell, Co-chair  
James Bell Associates  
Arlington, Virginia

Gary Ostroske, Co-chair  
United Way for the Greater  
New Orleans Area  
New Orleans, Louisiana

Terri Barreiro  
United Way of Minneapolis  
Minneapolis, Minnesota

Sylvia Barsion  
Girl Scouts of the U.S.A.  
New York, New York

David Cordray  
Vanderbilt University  
Nashville, Tennessee

Jackie Cox-New  
Winthrop Rockefeller  
Foundation  
Little Rock, Arkansas

Ann Cramer  
IBM Corporation  
Atlanta, Georgia

Stacey Daniels  
Ewing Marion Kauffman  
Foundation  
Kansas City, Missouri

Dan Duncan  
Mesa United Way  
Mesa, Arizona

Marjorie Gutman  
Robert Wood Johnson  
Foundation  
Princeton, New Jersey

Gillian Kerr  
United Way of Greater Toronto  
Toronto, Canada

Dagmar Edith McGill  
Big Brothers Big Sisters of  
America  
Philadelphia, Pennsylvania

Larrie Mackie  
United Way Services  
Cleveland, Ohio

Astrid Merget  
Syracuse University  
Syracuse, New York

Myrtis Meyer  
YMCA of the USA  
Chicago, Illinois

Ricardo Millett  
W.K. Kellogg Foundation  
Battle Creek, Michigan

Randy Morrison  
Godman Guild  
Columbus, Ohio

Heather Johnston Nicholson  
Girls Incorporated  
Indianapolis, Indiana

Paul Ostergard  
Citicorp Foundation  
New York, New York

Patricia Patrizi  
Consortium for Policy  
Research in Education  
Philadelphia, Pennsylvania

Kathleen Pritchard  
United Way of Greater  
Milwaukee  
Milwaukee, Wisconsin

Tom Rugh  
Indiana Association of United  
Ways  
Indianapolis, Indiana

Howard Rolston  
U.S. Department of Health  
and Human Services  
Washington, DC

James Sanders  
Western Michigan University  
Kalamazoo, Michigan

Betti Sands  
The New United Way  
Fresno, California

Beth Lee Terry  
United Way for the Greater  
New Orleans Area  
New Orleans, Louisiana

Dan Watanabe  
Hawai'i Community Services  
Council  
Honolulu, Hawai'i

Joseph Wholey  
University of Southern  
California—Washington  
Public Affairs Center  
Washington, DC



# NATIONAL HUMAN SERVICE ORGANIZATION REPRESENTATIVES MEETING TO ADVISE PROJECT STAFF AND CONSULTANTS February 13, 1996

<p><b>American Foundation for the Blind</b> Corinne Kirchner, Director of Programs &amp; Policy Research</p>	<p><b>Boy Scouts of America</b> C. Michael Hoover, Jr., Assistant Chief Scout Executive Stephen Medlicott, Associate Director of External Communications</p>	<p><b>Girls Scouts of the U.S.A.</b> Sharon Hussey, National Director of Membership &amp; Program</p>	<p><b>National Urban League</b> Curtiss E. Porter, Vice President Affiliate Development Department Darlene Hasselbring, Director of Evaluation Assessment</p>
<p><b>American Red Cross</b> Allen Berkowitz, Director, Corporate Planning &amp; Evaluation Sharon Silva, Evaluation Associate, Public Policy &amp; Planning</p>	<p><b>Catholic Charities USA</b> Jane Stenson, Program Director for Community Services John Keightley, Director-Development</p>	<p><b>Girls Incorporated</b> Janet Greenberg, Director of National Programs</p>	<p><b>The Salvation Army</b> Clarence W. Harvey, Secretary, National Community Relations &amp; Development Department</p>
<p><b>Association of Jewish Family &amp; Children's Agencies</b> Sherry Rosen, Senior Staff Associate</p>	<p><b>Child Welfare League of America</b> Lisa Lunghofer, Director, Odyssey Project</p>	<p><b>Goodwill Industries International</b> Ken Shaw, Division Director, Program Development &amp; Research Terri Guengerich, Research Director</p>	<p><b>United Cerebral Palsy Associations</b> Elizabeth Matthews, Administrative Assistant, Affiliate Services Division</p>
<p><b>Big Brothers Big Sisters of America</b> Tom Curtis, Assistant Director, Training &amp; Education Lynda Long, Assistant Director, Program Development</p>	<p><b>Council of Jewish Federations, Inc.</b> Joan F. Strauss, Associate Director of Community Planning</p>	<p><b>National Assembly of National Voluntary Health and Social Welfare Organizations</b> Kim Helfgott, Director, Human Resource Development</p>	<p><b>United Neighborhood Centers of America</b> David Anderson, Agency Director, Pillsbury Neighborhood Services</p>
<p><b>Boys &amp; Girls Clubs of America</b> Judith J. Carter, Senior Vice President, Program Services</p>	<p><b>Epilepsy Foundation of America</b> Jim Troxell, Senior Director of Research &amp; Professional Education, &amp; Employee &amp; Training Services</p>	<p><b>National Coalition of Hispanic Health and Human Services Organizations (COSSMHO)</b> Ana Rivas Bertolo, Senior Policy Advisor</p>	<p><b>Volunteers of America</b> Margaret Ratcliff, Vice President of Service Development</p>
	<p><b>Family Service America</b> Paul Mattessich, Director of Research, Wilder Foundation</p>	<p><b>National Council of La Raza</b> Tita Nieves, Technical Assistance Specialist</p>	<p><b>YMCA of the USA</b> Carmelita Gallo, Director of Community Resources Lynne G. Vaughan, Director of Program Development</p>

## Moderators

Sylvia Barsion, National Director of Research & Evaluation  
Girl Scouts of the U.S.A.

Heather Johnston Nicholson, Associate Executive  
Director for Research, Program & Training  
Girls Incorporated

Myrtis Meyer, Director of Research  
YMCA of the USA

Daneen Peterson, Director of Research & Evaluation  
Big Brothers Big Sisters of America



# Contents

<b>Acknowledgements</b> .....	iii
<b>United Way of America Task Force on Impact: Members Involved in Development of Manual</b> .....	v
<b>National Human Service Organization Representatives Meeting to Advise Project Staff and Consultants</b> .....	vii
<b>Glossary of Selected Outcome Measurement Terms</b> .....	xv
<b>Introduction to Outcome Measurement</b> .....	1
Why Measure Outcomes? .....	4
About This Manual .....	5
Notes About Terminology .....	7
In Conclusion .....	8
<b>Exhibits</b>	
A: Summary of Program Outcome Model .....	3
B: Measuring Program Outcomes: Eight Steps to Success .....	6
<b>Step 1: Get Ready</b> .....	11
Task 1: Assemble and Orient an Outcome Measurement Work Group .....	12
Work Group Responsibilities .....	13
Work Group Membership .....	14
Initial Discussions .....	16
Definition of Terms .....	16
Value of Outcome Measurement .....	16
Limitations and Potential Problems of Outcome Measurement .....	21
Expectations of Agency Director, Board, Other Key Players .....	24
Available Resource Materials .....	25
Task 2: Decide Which Program(s) to Start With .....	25
Task 3: Develop a Timeline .....	27
Task 4: Distribute Your Game Plan to Key Players .....	30
<b>Exhibits</b>	
1-A: Responsibilities of Outcome Measurement Work Group .....	15
1-B: The Traditional Service Program Model .....	17
1-C: The Program Outcome Model .....	18
1-D: Hints as to What Are and Are Not Outcomes .....	19
1-E: Examples of Diverse Programs and Possible Outcomes .....	20
1-F: Examples of Potential Benefits of Outcome Measurement .....	22
1-G: Guidelines for Incorporating Outcome Data in Funding Decisions .....	26
1-H: Sample Timeline for Planning and Implementing Outcome Measurement in a Program .....	28
<b>Worksheet</b>	
1: Timeline for Planning and Implementing Outcome Measurement .....	29

<b>Step 2: Choose the Outcomes You Want to Measure</b> .....	31
Levels of Outcomes .....	32
Task 1: Gather Ideas for What Your Program's Outcomes Are from a Variety of Sources .....	33
Task 2: Construct a Logic Model for Your Program .....	38
Issues in Defining Outcomes .....	49
Preparing your Logic Model .....	52
Task 3: Select the Outcomes That Are Important to Measure .....	53
Task 4: Get Feedback on Your Logic Model and the Outcomes Selected for Measurement .....	56
<b>Exhibits</b>	
2-A: Inputs Through Outcomes: The Conceptual Chain .....	34
2-B: Conducting Focus Groups .....	36
2-C: Examples of Outcomes Based on Observations of a Neighborhood After-school Program .....	39
2-D: Southside Children's Agency .....	41
2-E: Example Logic Models for Four Programs of Southside Children's Agency (horizontal format) .....	42
2-F: Example Logic Models for Four Programs of Southside Children's Agency (vertical format) .....	45
2-G: Questions Regarding Proposed Program Logic Model and Outcomes .....	57
<b>Worksheets</b>	
2: Program Logic Model .....	44
3: Assessing Possible Outcomes .....	54
<b>Step 3: Specify Indicators for Your Outcomes</b> .....	59
Task 1: Specify One or More Indicators for Each Outcome .....	59
Task 2: Decide What Factors Could Influence Participant Outcomes .....	68
Special Problems with Hard-to-Measure Outcomes .....	74
<b>Exhibits</b>	
3-A: Sample Outcome Indicators: Southside Children's Agency .....	64
3-B: Example Outcomes and Outcome Indicators for Various Programs .....	66
3-C: Example of Value of Outcome Data Broken Out by Influencing Factors .....	69
3-D: Participant and Program Characteristics That May Influence Outcomes .....	70
3-E: Factors That May Influence Participant Outcomes: Southside Children's Agency .....	72
<b>Worksheets</b>	
4: Outcome Measurement Framework .....	60
5: Outcome Measurement Framework: Factors That May Influence Participant Outcomes .....	75

<b>Step 4: Prepare to Collect Data on Your Indicators</b>	.81
Task 1: Identify Data Sources for Your Indicators	.81
Task 2: Design Data Collection Methods	.87
Preparing Data Collection Instruments	.89
Developing Data Collection Procedures	.95
At What Milestones Should Data Be Collected?	.96
Who Is Considered a Participant?	.97
Include All Participants or Only a Sample?	.97
Who Will Collect the Data?	.98
How Will Confidentiality Be Protected?	.98
How Will Participants Be Informed About the Data Collection Process?	.100
Task 3: Pretest Your Data Collection Instruments and Procedures	.102
<b>Exhibits</b>	
4-A: Advantages and Disadvantages of Data Sources for Outcome Indicators	.86
4-B: Comparison of Major Data Collection Methods	.88
4-C: Sample Outcome Measurement Framework: Southside Children's Agency	.90
4-D: Considering Cultural Issues in Instrument Development	.94
4-E: Sample Pledge of Confidentiality for Data Collectors	.99
4-F: Sample Consent Form	.101
<b>Step 5: Try Out Your Outcome Measurement System</b>	.105
Task 1: Develop a Trial Strategy	.106
Task 2: Prepare the Data Collectors	.107
Task 3: Track and Collect Outcome Data	.107
Task 4: Monitor the Outcome Measurement Process	.110
<b>Exhibits</b>	
5-A: Sample Master List of Data Elements Needed for Each Participant	.107
5-B: Sample Data Collection Log	.109
<b>Step 6: Analyze and Report Your Findings</b>	.113
Task 1: Enter the Data and Check for Errors	.114
Task 2: Tabulate the Data	.115
Task 3: Analyze the Data Broken Out by Key Characteristics	.116
Task 4: Provide Explanatory Information Related to Your Findings	.118
Task 5: Present Your Data in Clear and Understandable Form	.118
<b>Exhibits</b>	
6-A: Sample Spreadsheet for Programs Without Computers	.115
6-B: Outcome of Delivering Healthy Babies, by Age of Mother	.117
6-C: Sample Outcome Data Table: Southside Children's Agency Teen Mother Parenting Education Program	.119
6-D: Percent of Participants with Pre- to Post-Test Increases of Six Points or More	.120
6-E: Changes in Youths' Conscientiousness About Homework	.121

<b>Step 7: Improve Your Outcome Measurement System</b> .....	125
Task 1: Review Your Trial-run Experience, Make Necessary Adjustments, and Start Full-scale Implementation .....	126
Task 2: Monitor and Review Your System Periodically .....	127
 <b>Step 8: Use Your Findings</b> .....	129
Internal Uses of Outcome Findings .....	130
Provide Direction for Staff .....	130
Identify Training and Technical Assistance Needs .....	131
Identify Program Improvement Needs and Strategies .....	131
Support Annual and Long-range Planning .....	140
Guide Budgets and Justify Resource Allocations .....	140
Suggest Outcome Targets .....	140
Focus Volunteers' Attention on Policy and Programmatic Issues .....	141
External Uses of Outcome Findings .....	142
Recruit Talented Staff and Volunteers .....	142
Promote Your Program to Potential Participants and Referral Sources .....	143
Identify Partners for Collaborations .....	143
Enhance Your Program's Public Image .....	144
Retain and Increase Funding .....	145
 <b>Exhibits</b>	
8-A: Outcome Findings by Program Unit and Level of Difficulty of Participants' Problems .....	133
8-B: Comparative Findings for Two Youth Services Approaches: Team vs. Individual Activity Emphasis .....	134
8-C: Percent of Participants with Pre- to Post-Test Increases of Six Points or More, 1995 and 1996 .....	136
8-D: Southside Neighborhood Children's Agency Child Abuse Public Awareness and Prevention Program: Actual Outcomes vs. Targets .....	138
8-E: Burglaries in Southside Neighborhood Before and After Implementation of Neighborhood Organizing Program .....	139

<b>Appendix: Issues in Developing Data Collection Instruments and Procedures</b>	.147
Record Extraction Forms	.147
Instrument Development	.147
Procedural Issues Specific to Record Extraction	.149
Questionnaires	.150
Instrument Development	.150
Content	.150
Wording	.153
Format and Style	.155
Procedural Issues Specific to Questionnaires	.156
Administration	.156
Voluntary Completion	.159
Participant Confidentiality	.159
Anonymity	.159
Times for Contacting Potential Interview Respondents	.160
Response Rate	.160
Costs	.160
Trained Observer Ratings	.162
Instrument Development	.162
Procedural Issues Specific to Trained Observer Ratings	.162
<b>Exhibits</b>	
1: Example of a Record Extraction Form	.148
2: Survey Questions on Family Counseling	.152
3: Examples of Fixed-choice Response Sets	.154
4: Summary Comparison of Questionnaire Administration Options	.158
5: Tips for Improving Mail Survey Response Rates	.161
6: Sample Rating Scale for Trained Observer Rating of a Playground for Preschool Children	.163
<b>Bibliography</b>	.165
Performance Measurement	.165
Data Collection	.166
Performance Indicators	.168



# Glossary of Selected Outcome Measurement Terms

**Inputs** are resources a program uses to achieve program objectives. Examples are staff, volunteers, facilities, equipment, curricula, and money. A program uses *inputs* to support *activities*.

**Activities** are what a program does with its inputs—the services it provides—to fulfill its mission. Examples are sheltering homeless families, educating the public about signs of child abuse, and providing adult mentors for youth. Program *activities* result in *outputs*.

**Outputs** are products of a program's activities, such as the number of meals provided, classes taught, brochures distributed, or participants served. Another term for "outputs" is "units of service." A program's *outputs* should produce desired *outcomes* for the program's participants.

**Outcomes** are benefits for participants during or after their involvement with a program. Outcomes may relate to knowledge, skills, attitudes, values, behavior, condition, or status. Examples of outcomes include greater knowledge of nutritional needs, improved reading skills, more effective responses to conflict, getting a job, and having greater financial stability.

For a particular program, there can be various "levels" of outcomes, with initial outcomes leading to longer-term ones. For example, a youth in a mentoring program who receives one-to-one encouragement to improve academic performance may attend school more regularly, which can lead to getting better grades, which can lead to graduating.

**Outcome indicators** are the specific items of information that track a program's success on outcomes. They describe observable, measurable characteristics or changes that represent achievement of an outcome. For example, a program whose desired outcome is that participants pursue a healthy lifestyle could define "healthy lifestyle" as not smoking; maintaining a recommended weight, blood pressure, and cholesterol level; getting at least two hours of exercise each week; and wearing seat belts consistently. The number and percent of program participants who demonstrate these behaviors then is an *indicator* of how well the program is doing with respect to the outcome.

**Outcome targets** are numerical objectives for a program's level of achievement on its outcomes. After a program has had experience with measuring outcomes, it can use its findings to set targets for the number and percent of participants expected to achieve desired outcomes in the next reporting period. It also can set targets for the amount of change it expects participants to experience.

**Benchmarks** are performance data that are used for comparative purposes. A program can use its own data as a baseline benchmark against which to compare future performance. It also can use data from another program as a benchmark. In the latter case, the other program often is chosen because it is exemplary and its data are used as a target to strive for, rather than as a baseline.

"The reason we got into this business was to change lives. Now we have all the staff thinking in this orientation. Outcome measurement creates focus in a way no other management tool can do. We used to have forms we filled out that were meaningless. Now we have the same number of forms, but we get real information. We used to count the number of things we did. Now we count results."

*Jan Berry, Executive Director, Freeport West, Minneapolis, Minnesota*

"Outcome measurement provides very important internal benefits. If we find something we're doing that is giving less than satisfactory results, we can amend the program and make improvements. The process also has helped us actually discontinue one of our programs. We were concerned about it before, and our outcome assessment helped make the decision to discontinue it."

*Natalie R. Peterson, Executive Director, Family Service Association of San Antonio Inc., San Antonio, Texas*

"Outcome information is a real morale builder for the staff. They take it hard when teens get pregnant again, but I was able to show them that only 7 percent get pregnant while in the program, compared to 52 percent in the county and 44 percent in the state. It gives the staff a sense of pride to understand that these girls would be part of the 52 percent if we didn't work with them."

*Peggy Ann Dorrian, Talking, Listening, Caring Program Supervisor, Catholic Family and Community Services, Schenectady, New York*

"In our parenting skills class, the parents see the assessment as an empowering experience. People traditionally have a negative mind-set about tests, but these parents find the process helpful. It puts bookends on what it means to be a good parent. They can see they have specific skills. It's concrete. They go down the list and say 'I'm good at that.' They feel in control. The program participants have been so positive that it's rubbed off on the staff."

*David L. Hoffman, President, Family Service of Milwaukee, Milwaukee, Wisconsin*

"Outcome data gives our volunteer fund raisers a greater level of comfort to go out and represent us. Before, they were just able to tell anecdotal stories."

*Patricia C. Howard, Executive Director, Girls Incorporated of Memphis, Memphis, Tennessee*

"Our outcome measurement system has affected how we are viewed in the community. We can talk to Mr. and Mrs. John Doe in terms of cost-effectiveness for their dollars. It has given staff a sense of pride and accomplishment. We no longer have the turnover we had seven years ago. It also has affected our voluntarism. We had a 313 percent increase in one year."

*Kay Coffin, Executive Director, Big Brothers Big Sisters of Fresno County, Fresno, California*

"For us it was definitely worth the time we put into it. Definitely. Two years ago I would not have told you that. It has helped us recognize inefficiencies and helped us reorganize and improve services. I tell the staff over and over that it's not just a chore being mandated by our funder. It's a good thing to be doing anyway. People are beginning to understand."

*Michelle Gaynier, Director of Communications and Marketing, The Detroit Institute for Children, Detroit, Michigan*

"Kingsley House has been more successful in maintaining funding and getting supplemental funding because we are able to demonstrate the impact of dollars. Corporations particularly are quite impressed with our ability to show impact—the effect on kids and our ability to keep adults out of a nursing home."

*Michael Moreau, Executive Director, Kingsley House and New Orleans Day Nursery Association, New Orleans, Louisiana*

"We really feel we're going in the right direction. It's a lot of work, but well worth it. It provides a clarity to the staff, insures a continuation of funding and from a management perspective, you really need it. Now that we've done it I don't know how people make it without it."

*Nina Waters, Executive Director, Practical and Cultural Education (PACE) Center for Girls, Jacksonville, Florida*

"If United Way stopped requiring this today, we would keep doing it."

*Marilyn Lange, Executive Director, Village Adult Day Center, Inc., Milwaukee, Wisconsin*

# Introduction to Outcome Measurement

If yours is like most human service agencies or youth- and family-serving organizations, you regularly monitor and report on how much money you receive, how many staff and volunteers you have, and what



they do in your programs. You know how many individuals participate in your programs, how many hours you spend serving them, and how many brochures or classes or counseling sessions you produce. In other words, you document program *inputs, activities, and outputs*.

- **Inputs** include resources dedicated to or consumed by the program. Examples are money, staff and staff time, volunteers and volunteer time, facilities, equipment, and supplies. For instance, inputs for a parent education class include the hours of staff time spent designing and delivering the program. Inputs also include constraints on the program, such as laws, regulations, and requirements for receipt of funding.
- **Activities** are what the program does with the inputs to fulfill its mission. Activities include the strategies, techniques, and types of treatment that comprise the program's service methodology. For instance, sheltering and feeding homeless families are program activities, as are training and counseling homeless adults to help them prepare for and find jobs.
- **Outputs** are the direct products of program activities and usually are measured in terms of the volume of work accomplished—for example, the numbers of classes taught, counseling sessions conducted, educational materials distributed, and participants served. Outputs have little inherent value in themselves. They are important because they are intended to lead to a desired benefit for participants or target populations.

If given enough resources, managers can control output levels. In a parent education class, for example, the number of classes held and the number of parents served are outputs. With enough staff and supplies, the program could double its output of classes and participants.

If yours is like most human service organizations, you do *not* consistently track what happens to participants after they receive your services. You cannot report, for example, that 55 percent of your participants used more appropriate approaches to conflict management after your youth development program conducted sessions on that skill, or that your public awareness program was followed by a 20 percent increase in the number of low-income parents getting their children immunized. In other words, you do not have much information on your program's *outcomes*.

- **Outcomes** are benefits or changes for individuals or populations during or after participating in program activities. They are influenced by a program's outputs. Outcomes may relate to behavior, skills, knowledge, attitudes, values, condition, or other attributes. They are what participants know, think, or can do; or how they behave; or what their condition is, that is different following the program.

For example, in a program to counsel families on financial management, outputs—what the service produces—include the number of financial planning sessions and the number of families seen. The desired outcomes—the changes sought in participants' behavior or status—can include their developing and living within a budget, making monthly additions to a savings account, and having increased financial stability.

In another example, outputs of a neighborhood clean-up campaign can be the number of organizing meetings held and the number of weekends dedicated to the clean-up effort. Outcomes—benefits to the target population—might include reduced exposure to safety hazards and increased feelings of neighborhood pride. Exhibit A depicts the relationship between inputs, activities, outputs, and outcomes.

"We didn't have this information before. We knew the number of people we'd seen, but didn't really know if the services had made a difference for the client. Today, no one pays you to 'do good.' They pay for effectiveness."

*Ellie Gersten, Director,  
East Valley Catholic Social  
Services, Mesa, Arizona*

---

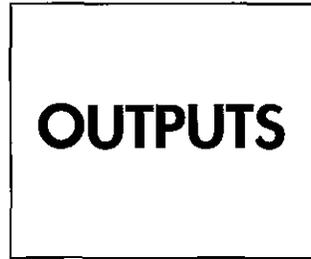
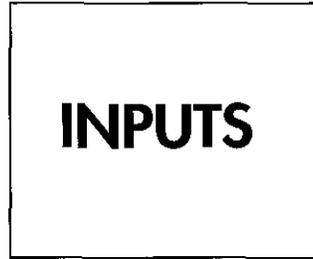
**Note:** *Outcomes* sometimes are confused with outcome *indicators*, specific items of data that are tracked to measure how well a program is achieving an outcome, and with outcome *targets*, which are objectives for a program's level of achievement.

For example, in a youth development program that creates internship opportunities for high school youth, an *outcome* might be that participants develop expanded views of their career options. An *indicator* of how well the program is succeeding on this outcome could be the number and percent of participants who list more careers of interest to them at the end of the program than they did at the beginning of the program. A *target* might be that *40 percent of participants list at least two more careers* after completing the program than they did when they started it. Outcome indicators and targets are discussed in Steps 3 and 8, respectively.

---

Exhibit A

**Summary of Program Outcome Model**



Resources dedicated to or consumed by the program

e.g.

- money
- staff and staff time
- volunteers and volunteer time
- facilities
- equipment and supplies

Constraints on the program

e.g.

- laws
- regulations
- funders' requirements

What the program does with the inputs to fulfill its mission

e.g.

- feed and shelter homeless families
- provide job training
- educate the public about signs of child abuse
- counsel pregnant women
- create mentoring relationships for youth

The direct products of program activities

e.g.

- number of classes taught
- number of counseling sessions conducted
- number of educational materials distributed
- number of hours of service delivered
- number of participants served

Benefits for participants during and after program activities

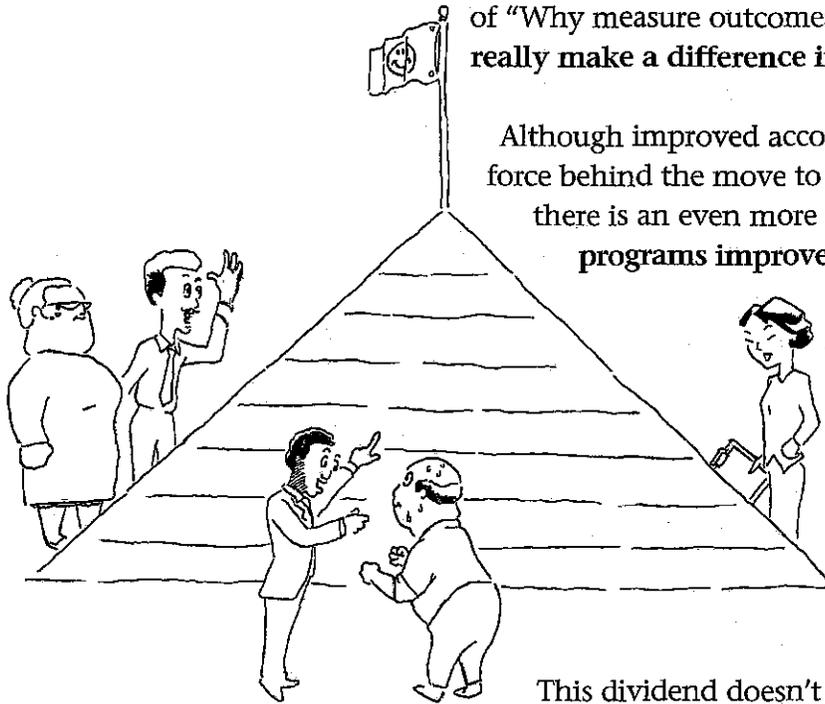
e.g.

- new knowledge
- increased skills
- changed attitudes or values
- ↓
- modified behavior
- ↓
- improved condition
- altered status

## Why Measure Outcomes?

In growing numbers, service providers, governments, other funders, and the public are calling for clearer evidence that the resources they expend actually produce benefits for people. Consumers of services and volunteers who provide services want to know that programs to which they devote their time really make a difference. That is, they want better accountability for the use of resources. One clear and compelling answer to the question

of "Why measure outcomes?" is: **To see if programs really make a difference in the lives of people.**



Although improved accountability has been a major force behind the move to outcome measurement, there is an even more important reason: **To help programs improve services.** Outcome measure-

ment provides a learning loop that feeds information back into programs on how well they are doing. It offers findings they can use to adapt, improve, and become more effective.

This dividend doesn't take years to occur. It often starts appearing early in the process of setting up an outcome measurement system. Just the process of focusing on outcomes—on why the program is doing what it's doing and how it thinks participants will be better off—gives program managers and staff a clearer picture of the purpose of their efforts. That clarification alone frequently leads to more focused and productive service delivery.

Down the road, being able to demonstrate that their efforts are making a difference for people pays important dividends for programs. It can, for example, help programs:

- Recruit and retain talented staff.
- Enlist and motivate able volunteers.
- Attract new participants.
- Engage collaborators.
- Garner support for innovative efforts.
- Win designation as a model or demonstration site.
- Retain or increase funding.
- Gain favorable public recognition.

Results of outcome measurement show not only where services are being effective for participants, but also where outcomes are not as expected. Program managers can use outcome data to:

- Strengthen existing services.
- Target effective services for expansion.
- Identify staff and volunteer training needs.
- Develop and justify budgets.
- Prepare long-range plans.
- Focus board members' attention on programmatic issues.

To increase its internal efficiency, a program needs to track its inputs and outputs. To assess compliance with service delivery standards, a program needs to monitor activities and outputs. But to improve its effectiveness in helping participants, to assure potential participants and funders that its programs produce results, and to show the general public that it produces benefits that merit support, an agency needs to measure its outcomes.

These and other benefits of outcome measurement are not just theoretical. Scores of human service providers across the country attest to the difference it has made—for their staff, their volunteers, their decision makers, their financial situation, their reputation, and, most important, for the public they serve. The rewards have been impressive enough to lead governments and private funders to pick up the idea. The 1993 Government Performance and Results Act (GPRA), for example, requires federal programs to identify and measure their outcomes. Many foundations now require programs they fund to measure and report on outcomes.

"Outcome measurement has been a tremendous benefit to our agency by clarifying the services we provide. It has helped the staff, clients, board, constituents, and funders. We have a whole list of topics to talk about and we are all on the same page. This is concrete data. We can compare success versus instances where there was not success and draw parallels."

*Nina Waters, Executive  
Director, Practical and  
Cultural Education (PACE)  
Center for Girls,  
Jacksonville, Florida*

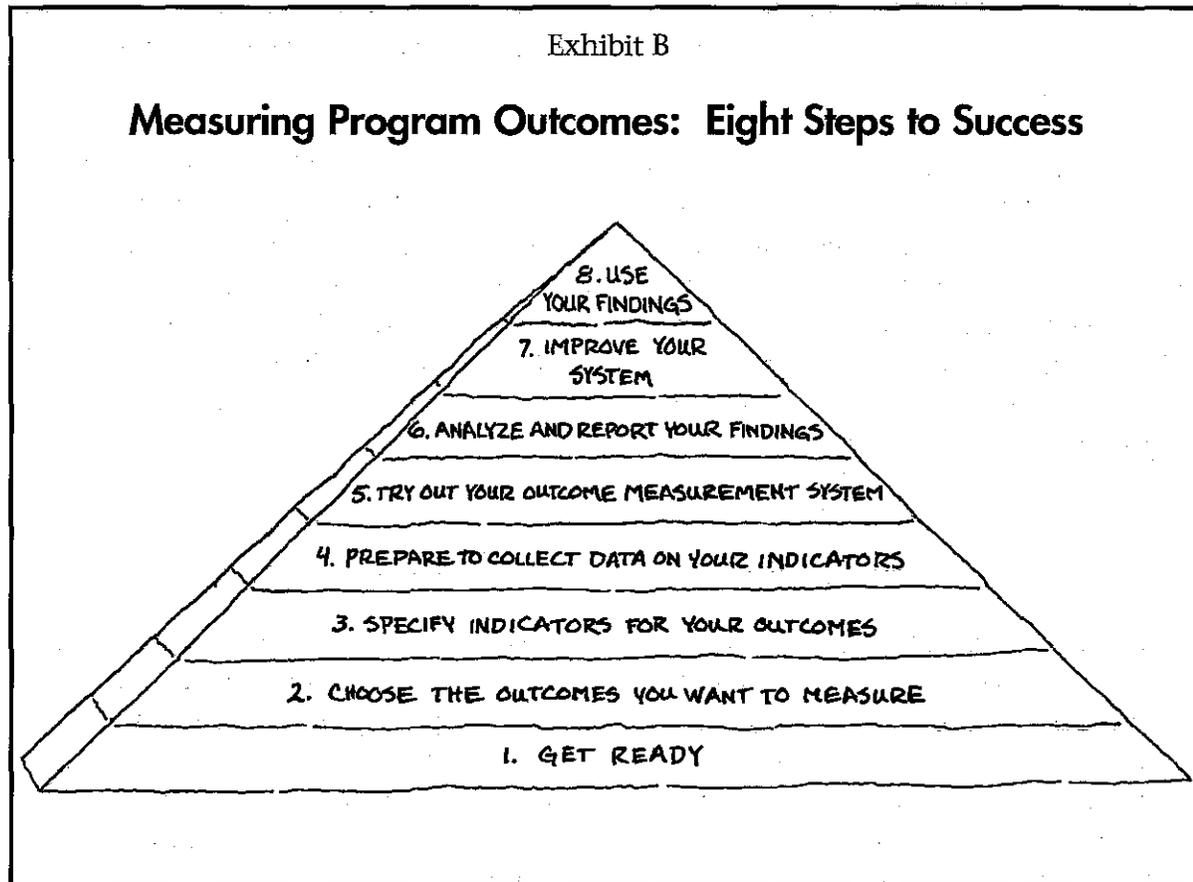
## About This Manual

*Measuring Program Outcomes: A Practical Approach* is written for executive directors and program managers in a broad range of services. These include not only intervention and remediation programs for individuals and families experiencing difficulties, but also prevention and development programs that protect and enhance well-being. The manual relates not only to direct-service providers, but also to programs engaged in advocacy, public education, capacity building, and related activities. The manual refers to this broad constellation of programs by the general term of "human services."

*Measuring Program Outcomes* provides a step-by-step approach to developing a system for measuring program outcomes and using the results. The approach, based on methods implemented successfully by agencies across the country, is presented in eight steps, shown in Exhibit B. Although the illustration suggests that the steps are sequential, this is actually a dynamic process with a good deal of interplay among stages.

To help illustrate the steps and their relationships to each other, *Measuring Program Outcomes* follows four hypothetical programs through the process of planning for outcome measurement. These programs are introduced in Step 2. The manual also provides worksheets and other tools to help you plan your program's outcome measurement system.

Experiences of human service providers who have implemented outcome measurement in their programs are spotlighted throughout the manual. They offer their encouragement, and their assurances that real benefits of this effort enjoyed by real programs can be gained by your program as well.



*Measuring Program Outcomes* is designed to apply to diverse agencies and programs. It makes no assumptions about agency size, budget, funding sources, level of computerization, setting, or participant characteristics. However, a one-size-fits-all guide to developing an outcome measurement system is not realistic. Parts of the manual may miss the mark for some readers in one or more ways. For example:

- Readers with training in program assessment may find the material in some sections to be too elementary. They can simply move on to the next section.
- Readers with no program assessment experience who need or want more detail can refer to the bibliography for more specialized references.
- The nature of some programs presents particular challenges related to identifying and measuring outcomes. Examples include advocacy programs (how do you measure system change?), prevention programs (how do you measure that something did not happen?), and programs that have only one-time contact with participants (how do you assess change in individuals you never see again?). The manual discusses special issues confronting various types of programs and offers suggestions for dealing with them.

It should be stressed that the processes and procedures described here are *not* the only valid approach to designing an outcome measurement system. Other organizations have prepared useful guides on the subject, as well as more comprehensive program assessment systems of which outcome measurement is a part. Agencies should use whatever model or combination provides guidance that meets their needs most closely. Funders are urged to respect this approach by focusing on the merits of a program's measurement process, rather than on the details of how the program went about developing the process.

## **Notes About Terminology**

Just as programs use different approaches to create outcome measurement strategies, so they apply different terms to various aspects of what this manual calls "outcome measurement." This Introduction defines some key terms as they are used in this manual. Other terms are defined as they are introduced.

When communicating with others, and particularly with funders, agencies are advised to use whatever terms will communicate most effectively with that audience. Funders, on the other hand, should recognize that different systems label outcomes as "client impacts," "program results," "effects," and other terms. The important issue must be the benefits of services for people, not the words used to describe them.

Two other notes about terminology are important for this manual's audience.

- The service providers for whom this manual is intended call themselves by varying terms, including "agency," "organization," "council," "program," "society," and "club." The manual uses the term "agency," in its meaning of a business or service that acts for others, to denote this broad spectrum of organizations dedicated to strengthening and empowering families and individuals.
- Agencies refer to those who participate in their programs by many terms, including "client," "customer," "recipient," "enrollee," and "member." This manual uses the term "participant" to represent the individuals and families whom the agency serves. As used in this manual, "participant" does not refer to program staff or volunteers.

## In Conclusion

"For those of us in the social service world, there's a certain trepidation in really finding out what we do. People should be excited about this possibility. There's no such thing as failure—it's feedback and shows where to make changes. It's a whole new arena now. I've been in social services for 20 years and, especially in prevention work, we've never been able to show what's effective. Now with this we're on the horizon of showing what works."

*Judy Selle,  
Executive Director,  
Neighborhood House,  
Milwaukee, Wisconsin*

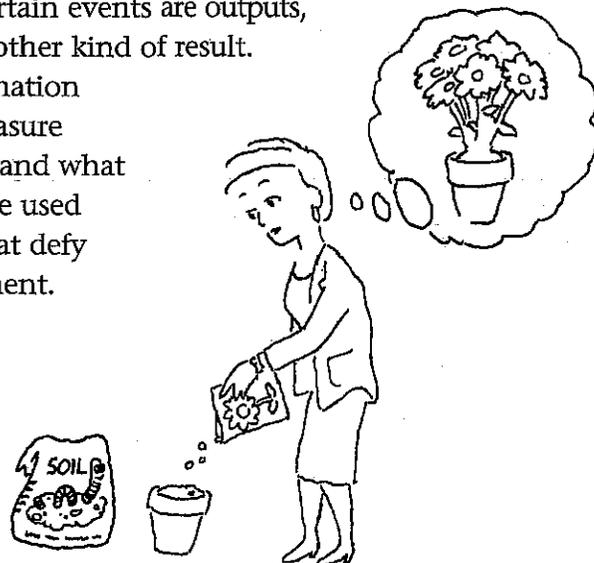
Outcome measurement is not a passing fad. In the years ahead, collecting data on benefits for program participants will be as common as collecting data on the number of program participants is today. Measurement methods and tools will have been refined, there will be greater agreement on terminology and criteria, and the subject will be part of the core curriculum in college and university courses for human service professionals.

Currently, however, the concept is relatively new and many issues are unresolved. There is debate over whether certain events are outputs, outcomes, or another kind of result.

There is consternation over how to measure some outcomes and what surrogates can be used for outcomes that defy direct measurement.

There is tension

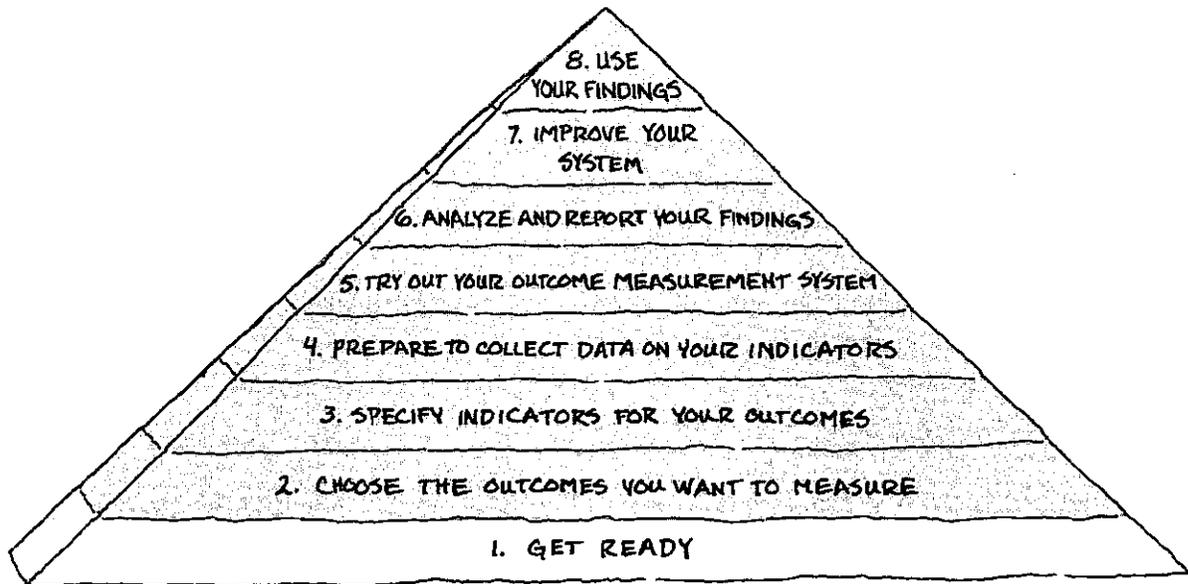
between the need for technically sound methodologies, which can be expensive and time consuming, and the staffing, funding, and workload realities that constrain nearly all service agencies.



This manual does not pretend to resolve these issues. It does, however, lay ground-work that will continue to be tested and refined. As the field progresses, the manual will be revised and updated. For now, its purpose is to provide help for those committed to improving the quality of life of people they serve.



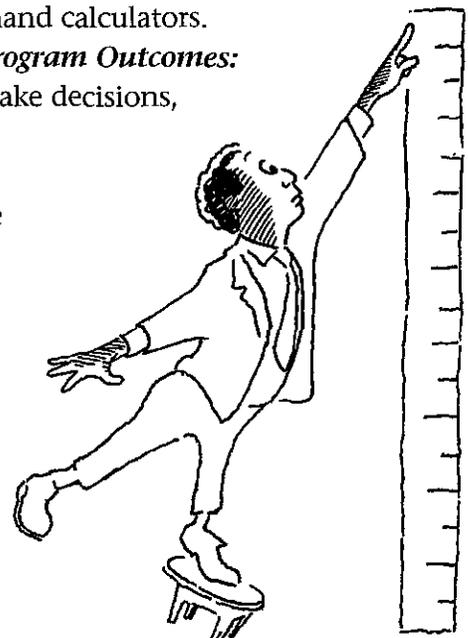
# Step 1: Get Ready



**P**lanning and implementing a system for measuring program outcomes is a challenging assignment. This is true whether your organization is large or small, has one program or many, is relatively new or has existed for several years, uses state-of-the-art computer systems or electric typewriters and hand calculators. Regardless of your agency's characteristics, *Measuring Program Outcomes: A Practical Approach* can help you examine options, make decisions, and avoid some of the pitfalls others have experienced.

The contents and approach of the manual reflect the following assumptions.

- Your agency has decided that it is going to implement outcome measurement for at least one program.
- Although one person will lead the process of planning and implementing outcome measurement, a work group of several individuals will be involved over a period of weeks or months in discussing and making decisions about each step.



- The work group will seek additional input, feedback, and perhaps expertise as appropriate during the process.
- Even if the major impetus for your agency's involvement in outcome measurement is a key funder's requirements, your top priority in developing the system will be its value as a tool for program improvement.

Given these assumptions, the manual starts at the beginning of the process with this chapter on getting ready to plan and implement an outcome measurement process. Key tasks involved in this step are:

1. Assemble and orient an outcome measurement work group.
2. Decide which program(s) to start with.
3. Develop a timeline.
4. Distribute your game plan to key players.

While this is the first step you will take in developing your outcome measurement system, many of the decisions to be made during this step depend on an understanding of the job that lies ahead. Therefore, you should review all the steps in *Measuring Program Outcomes: A Practical Approach* before actually implementing these tasks.

## **Task 1: Assemble and Orient an Outcome Measurement Work Group**

There are at least two reasons for having a group assume responsibility for developing an outcome measurement system. First, it is a large burden for one individual to carry. Second, each step of the process requires important decisions that must take many factors into account. Involving several perspectives will help assure that all issues are considered.

If your agency has several programs, but all of the programs have similar participants and similar objectives, one work group probably should address outcome measurement for all of them. You may not want to implement outcome measurement for all programs at the same time, however. Task 2 of this step gives guidance on deciding which programs to start with.

If your programs have different aims and involve different population groups, it may be more effective to have separate work groups for each program. If you create several work groups, also establish a mechanism to provide cross-group coordination of timelines, data collection plans, data processing schedules, reporting, and other logistics. One approach to this is for all work group leaders to meet periodically with the agency director.

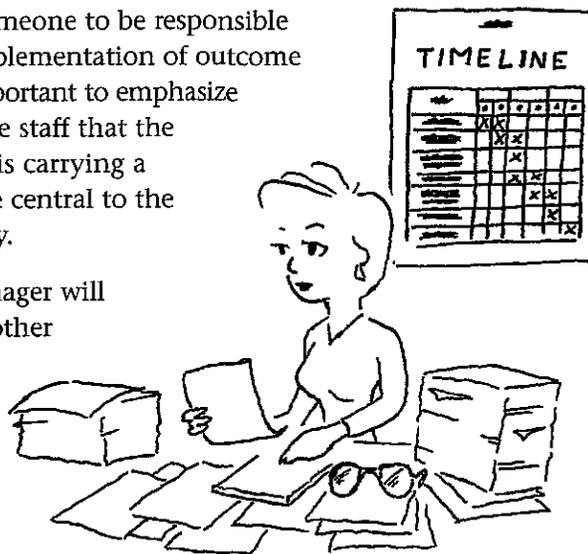
---

**To the Agency Director:** Your active support of your agency's outcome measurement process is vital to its success. You must demonstrate in words and actions that you are committed to this effort and expect the rest of the organization to follow your lead. This demonstration includes:

- Allocating resources to plan and implement the process successfully.
- Conveying the importance of this effort to your agency's volunteer structure and to its many publics.
- Keeping board members informed of progress and involved in key decisions.
- Setting an example by participating fully in whatever parts of the process relate to your responsibilities.

You may choose to designate someone to be responsible for day-to-day planning and implementation of outcome measurement. If you do, it is important to emphasize to that person and the rest of the staff that the manager has your support and is carrying a responsibility you consider to be central to the long-term success of your agency.

The outcome measurement manager will need adequate time freed from other responsibilities to lead this effort. The time required varies from agency to agency; five to 10 hours per week for approximately 12 months probably is a good average.



You or your designee may need to explore options for meeting or reducing the costs associated with outcome measurement. Options for this include:

- Cost-sharing collaborations with other organizations.
  - In-kind assistance from board members or other volunteers, local businesses, universities, and similar sources.
  - Development or capacity-building grants from a funder or local foundation.
- 

## Work Group Responsibilities

Members of the work group will have two types of responsibilities. They will, of course, have task-oriented responsibilities for identifying alternatives, making decisions, and allocating and carrying out assignments related to each planning and implementation step. These responsibilities are outlined in Exhibit 1-A. In the course of carrying out these tasks, the work group will:

- Make decisions that are appropriate for your agency, given its size, type, resources, and needs related to outcome measurement.
- Determine what resources are needed for each step, including staff training, technical assistance, or other expertise.
- Identify in-house resources that can be used in the effort, such as program monitoring and data collection systems already in place, staff and volunteer expertise related to various issues, and existing contracts or relationships that can be tapped creatively.
- Estimate costs and suggest ways to meet or reduce them, either by pursuing alternative implementation options or by seeking outside assistance.

To help assure successful implementation, the work group needs to communicate with and involve staff and volunteers at each step. As others have learned, it is vital to get widespread input and buy-in from the outset on such issues as:

- Why the agency is going to measure program outcomes.
- What outcomes it will measure.
- What data it will use to measure outcome.
- How it will collect and process data.
- How staff and volunteers will use the findings.

"We found it was important to get staff consensus up front—especially staff from different departments. We explained why the process was important and urged everyone to offer suggestions. We were able to talk as peers, not as management and employees. It also made the staff want to know what happened. Everyone wants to know the end of a story."

*Sue Malkin,  
Executive Director,  
Deaf, Hearing and  
Speech Center,  
Detroit, Michigan*

## Work Group Membership

The work group should be small. Four or five carefully chosen members—perhaps one or two more for large programs—can provide an adequate range of perspectives for your process. A group larger than seven or so may be too unwieldy for active discussions and timely decisions.

Keep in mind that the group will seek input from many sources at several steps in the process. At certain points, it may involve someone with expertise in areas such as questionnaire development and data analysis. Individuals can be involved at various times on an ad hoc basis while still keeping the core group small.

Following are some guidelines for assembling your work group. (Note that one individual may fit more than one of these categories.)

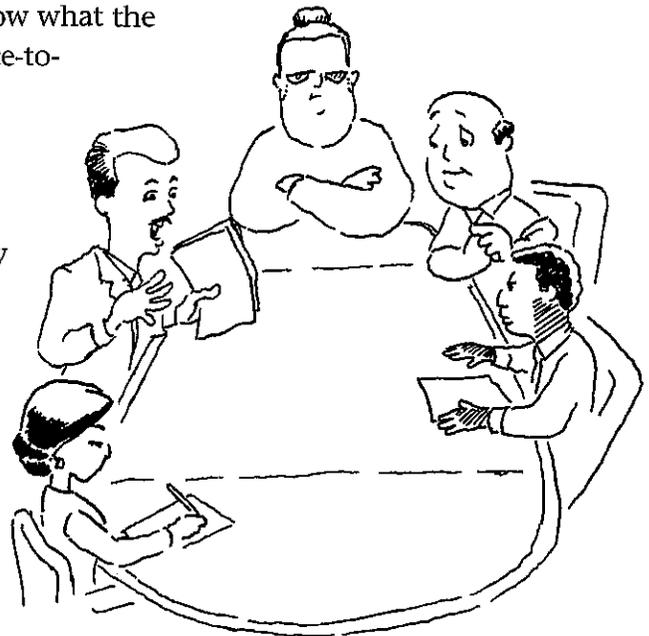
## Exhibit 1-A

### Responsibilities of Outcome Measurement Work Group

1. Decide which program(s) to start with. *(see Step 1: Get Ready)*
2. Develop a timeline. *(see Step 1)*
3. Identify outcomes to measure. *(see Step 2: Choose the Outcomes You Want to Measure)*
4. Construct a program logic model. *(see Step 2)*
5. Identify outcome indicators. *(see Step 3: Specify Indicators for Your Outcomes)*
6. See that data collection instruments are prepared, data collection plans are developed, and all instruments and plans are pretested. *(see Step 4: Prepare to Collect Data on Your Indicators)*
7. Plan for and monitor a pilot process. *(see Step 5: Try Out Your Outcome Measurement System)*
8. Monitor data analysis and report preparation. *(see Step 6: Analyze and Report Your Findings)*
9. Evaluate the results of the pilot and make necessary changes. *(see Step 7: Improve Your Outcome Measurement System)*

- It is critical that some group members know what the agency actually does at the day-to-day, face-to-face-with-participants level of operation. Thus, the group needs to include direct-service and/or first-level supervisory staff.

- Some members should represent other key functions, such as marketing or resource development; quality monitoring (e.g., licensing or accreditation compliance, program assessment); data processing; administrative support; and volunteer (e.g., member of board or of program evaluation or strategic planning committee).



---

**Note:** The work group should seek input from program participants at various points as it develops the outcome measurement system. Participants should play a role in identifying outcomes, reviewing survey questions, and testing data collection methods. For most programs, however, participants will not be members of the work group. The time required and the focus on administrative details of the program generally are not a match for participants' schedules and interests.

---

## Initial Discussions

Before the work group begins planning its approach to preparing for outcome measurement, it should discuss several issues to reach common understandings and expectations. Some topics for discussion are outlined below.

## Definition of Terms

It is imperative that group members share a common understanding of the concept of outcome measurement and that they are able to explain it to others. To facilitate discussions of key terms, you may want to provide group members with copies of:

- Exhibit A from this manual's **Introduction to Outcome Measurement** (page 3).
- Exhibits 1-B and 1-C, which capture the definitions of input, activity, output, and outcome presented in the **Introduction**.
- Exhibit 1-D, which discusses program components that sometimes are thought to be outcomes but usually are not.
- Exhibit 1-E, which gives examples of outputs and related outcomes.

## Value of Outcome Measurement

Measuring outcomes is not simple. It will help work group members and others in your agency pursue the tasks ahead if they focus on the value of outcome measurement and the benefits it can have. Exhibit 1-F shows hoped-for benefits generated by representatives of several agencies in Pierce County, Washington before they embarked on outcome measurement. Your work group may want to generate its own list and share it with other staff and volunteers.

"I view this as a benefit, as another tool to use rather than another chore. Some staff were not enthusiastic initially, but once they understood the benefit they became invested. They began to see the value to us as an agency and to our clientele."

*Marilyn Lange,  
Executive Director,  
Village Adult Day Center, Inc.,  
Milwaukee, Wisconsin*

Exhibit 1-B

## The Traditional Service Program Model

Most of us are familiar with the basic program model shown below. It has three major components: program inputs (resources), activities (services), and outputs (products).



- **Inputs** include resources dedicated to or consumed by the program. Examples are money, staff and staff time, volunteers and volunteer time, facilities, equipment, and supplies. For example, inputs for a parent education class include the hours of staff time spent designing and delivering the program.

Inputs also include constraints on the program, such as laws, regulations, and requirements for receipt of funding.

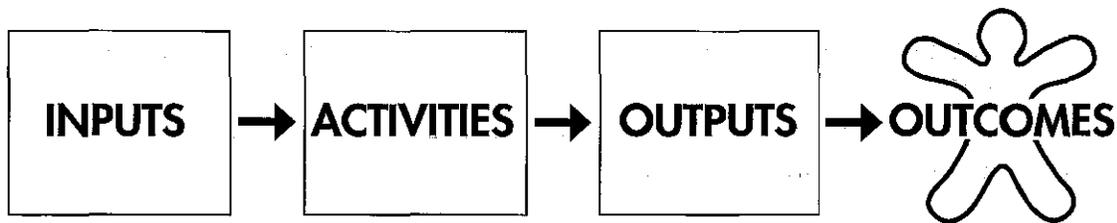
- **Activities** are what the program does with the inputs to fulfill its mission. Activities include the strategies, techniques, and types of treatment that comprise the program's service methodology. For instance, sheltering and feeding homeless families are program activities, as are training and counseling homeless adults to help them prepare for and find jobs.
- **Outputs** are the direct products of program activities and usually are measured in terms of the volume of work accomplished—for example, the numbers of classes taught, counseling sessions conducted, educational materials distributed, and participants served. Outputs have little inherent value in themselves. They are important because they are intended to lead to a desired benefit or change for participants or target populations.

If given enough resources, managers can control output levels. In a parent education class, for example, the number of classes held and the number of parents served are outputs. With enough staff and supplies, the program could double its output of classes and participants.

## The Program Outcome Model

The traditional service model in Exhibit 1-B lacks one important focus: the benefits participants derive from the program. No amount of input, activity, and output analysis can answer the question, "Are participants or target groups better off after receiving the service than they were before?"

The question of whether, and how much, a program's participants have changed, how their status has improved, how they have benefitted, is answered by measuring the program's *outcomes*. Outcomes are the arrow beyond the outputs, as shown here.



- **Outcomes** are benefits or changes for individuals or populations during or after participating in program activities. They are influenced by a program's outputs. Outcomes may relate to behavior, skills, knowledge, attitudes, values, condition, status, or other attributes. They are what participants know, think, or can do; or how they behave; or what their condition is, that is different following the program.

For example, in a program to counsel families on financial management, outputs—what the service produces—include the number of financial planning sessions and the number of families seen. The desired outcomes—the changes sought in participants' behavior or status—can include their developing and living within a budget, making monthly additions to a savings account, and having increased financial stability.

In another example, outputs of a neighborhood clean-up campaign can be the number of organizing meetings held and the number of weekends dedicated to the clean-up effort. Outcomes—benefits to the target population—might include reduced exposure to safety hazards and increased feelings of neighborhood pride.

## Exhibit 1-D

### Hints as to What Are and Are Not Outcomes

Here are some program components that agencies sometimes have trouble classifying as activities, outputs, or outcomes. These hints provide general guidelines and point out when exceptions may be appropriate.

- **Recruiting and training staff and volunteers, purchasing or upgrading equipment, and various support and maintenance activities.** These are internal program operations intended to improve the quality of program inputs. The number of staff recruited, number of volunteers trained, amount of equipment purchased, etc., indicate the volume of these internal operations. However, the operations do not represent benefits or changes in participants, and thus *are not outcomes*.
- **Number of participants served.** This information relates to the volume of work accomplished. In most cases, volume of service is an output. It tells nothing about whether participants benefitted from the service and therefore *is not usually an outcome*.

In public education programs where the program aims to encourage citizens to seek a service, such as cancer screening, the fact that citizens become aware of the importance of the service and seek it out reflects a change in knowledge or attitudes and behavior resulting from the program. Thus, the number of citizens who are motivated to seek a service by a public education program is an outcome of that program.

- **Participant satisfaction.** Most often, whether a participant is satisfied or not with various aspects of a program (e.g., courteousness of staff, timeliness of follow-up) does not indicate whether the participant's condition improved as a result of the service. Thus, participant satisfaction *generally is not an outcome*.

In rare instances, participant satisfaction may be part of the series of changes a participant experiences in achieving a successful outcome. For example, if an individual's willingness to continue with long-term counseling is critical to the program's success and satisfaction is a key determinant of continuation, then satisfaction may be a necessary, although not sufficient, outcome. In programs whose purpose is to meet participant's basic needs, such as food kitchens and homeless shelters, it may be nearly impossible to track participants far enough beyond the immediate delivery of service to identify outcomes beyond being fed and sheltered. In these cases, the program may have to settle for participant satisfaction as the closest approximation of an outcome it can measure.

Exhibit 1-E

**Examples of Diverse Programs and Possible Outcomes<sup>1</sup>**

Program	Possible Outcomes
Comprehensive child care	<ul style="list-style-type: none"> <li>■ Children exhibit age-appropriate physical, mental, and verbal skills.</li> <li>■ Children are school-ready for kindergarten.</li> </ul>
General Equivalency Diploma (GED) preparation	<ul style="list-style-type: none"> <li>■ Participants obtain their GED certificate.</li> <li>■ Within 6 months after obtaining their GED, participants are employed full-time.</li> </ul>
Outpatient treatment for adolescent substance abusers	<ul style="list-style-type: none"> <li>■ Adolescents increase knowledge about the effects of substance abuse and about substance abuse addiction.</li> <li>■ Adolescents change attitude towards substance abuse.</li> <li>■ Graduates remain free of substance abuse 6 months after program completion.</li> </ul>
Emergency shelter beds on winter nights	<ul style="list-style-type: none"> <li>■ Homeless persons agree to come off the street and use the shelter.</li> <li>■ Those sheltered do not suffer from frostbite or die from exposure to cold.</li> </ul>
Homework guidance by volunteer tutors to children enrolled in after-school program	<ul style="list-style-type: none"> <li>■ Youths' attitude towards school work improves.</li> <li>■ Youths complete homework assignments.</li> <li>■ Youths perform at or above grade level.</li> </ul>
Full-day therapeutic child care for homeless preschoolers	<ul style="list-style-type: none"> <li>■ Children get respite from family stress.</li> <li>■ Children engage in age-appropriate play.</li> <li>■ Children exhibit fewer symptoms of stress-related regression.</li> <li>■ Parents receive respite from child care.</li> </ul>
Overnight camping for 8- to 12-year-old inner-city boys	<ul style="list-style-type: none"> <li>■ Boys learn outdoor survival skills.</li> <li>■ Boys develop enhanced sense of competence.</li> <li>■ Boys develop and maintain positive peer relationships.</li> </ul>
Congregate meals for senior citizens	<ul style="list-style-type: none"> <li>■ Participants have social interaction with peers.</li> <li>■ Participants are not home-bound.</li> <li>■ Participants eat nutritious and varied diet.</li> <li>■ Seniors experience decrease in social and health problems.</li> </ul>

<sup>1</sup> These are illustrative examples only. Programs should identify their own outcomes, matched to and based on their own experiences and missions and input of their staff, volunteers, participants, and others.

Exhibit 1-E, cont'd

Program	Possible Outcomes
Interpreter services for non-English speaking patients at a health clinic	<ul style="list-style-type: none"><li>■ Patients access needed health care.</li><li>■ Patients understand medical diagnosis and need for preventive measures or treatment.</li><li>■ Patients comply with medical recommendations.</li><li>■ Patients have decrease in health conditions that are preventable.</li><li>■ Patients recuperate from conditions that can be treated.</li></ul>
Personal safety training for residents of subsidized apartment building	<ul style="list-style-type: none"><li>■ Residents initiate Neighborhood Watch Program.</li><li>■ Residents develop and implement action plan to make building more secure.</li><li>■ Residents feel safer.</li><li>■ Personal and property attacks decline.</li></ul>

The group should go beyond creating a lengthy list to agree on the two or three most important benefits it wants your agency to gain from measuring its outcomes. These become your objectives, or mission, for this effort. These objectives should be part of what the work group distributes throughout your organization at the end of Step 1.

### Limitations and Potential Problems of Outcome Measurement

Outcome measurement offers much value to programs. However, it has limitations that are important to recognize. The work group should discuss these so that everyone is clear about what outcome measurement will *not* do. Three key limitations are discussed below.

- Although outcome findings may show that program participants are not experiencing the intended benefits, they do not show where the problem lies or what is needed to fix it. To pinpoint and correct problems, program managers need the kind of data on program inputs, processes, and outputs they probably have been collecting for some time. Thus, outcome measurement is an addition to existing data collection efforts, not an alternative.
- Outcome measurement provides an ongoing means for managers to track how many of their participants achieve the outcomes that are desired, but it does not prove that the program, and the program alone, caused the outcomes. This is a key difference between outcome measurement and program impact research, which uses random

Exhibit 1-F

## **Examples of Potential Benefits of Outcome Measurement**

(summary of brainstorming by agencies in Pierce County, Washington)

- Provide a communication tool to let people know what's being done and the difference that it makes
- Reaffirm that we are on the right track
- Get information to use for program development
- Focus on programs that really make a difference for clients
- Make programs tangible by describing expected outcomes
- Benefit the agency and its long-range planning efforts
- Benefit families that use the services
- Let clients have a say in services
- Do a better job for our clients
- Have data to show quality
- Help focus on primary tasks
- Use for future planning
- A reality check
- Demonstrate "intangibles"
- Collect information as a routine part of what we are doing
- Get information that will be useful for decision-making
- Develop a model that will stop us from having to do ongoing testing
- Keep us from being penalized for things we cannot measure
- Communicate to funders a balance of outcomes and outputs
- Help justify existence
- Help the Board
- Be accountable and cost-effective
- Take pride in accomplishment and quality

assignment of participants to either the program or to control groups and sophisticated statistical methods. This is the only way to separate the program's influence from other factors that might affect participants. While the lack of such methodology makes outcome measurement less expensive and more feasible than impact research, it also limits the conclusions that can be drawn from the findings.

- The findings of outcome measurement don't reveal whether the outcomes being measured are the right ones for a particular program—the ones that best reflect meaningful change in the status of participants. Assuring that the process is measuring the appropriate outcomes is up to those who design the process.

Not only are there limits to the power of outcome measurement, there also are problems in its implementation.

- The thorniest problem probably is that, given today's state of the art, there are not established, readily available indicators and measurement methods for the outcomes of some programs. Steps 3 and 4 of this manual offer ideas on grappling with measurement issues. However, identifying measurable-yet-meaningful approximations of hard-to-measure outcomes is an ongoing challenge for many programs.

"Anybody who develops a program should build in research and evaluation funds. We're starting to put part of our total budget for evaluation and research as part of every grant, and it's a very good thing we've done to standardize funding. Now we can have enough money to pay someone else to do that part and benefit from their objectivity."

*Jim Clark, Vice President,  
Daniel Memorial,  
Jacksonville, Florida*

- Another significant problem is cost in time, effort, and money. Although not as expensive as program impact research, outcome measurement requires resources. The biggest single cost is staff time, but reference materials, paper, duplication, postage, mileage, phone calls, and outside consultation are some of the added costs likely to be incurred for outcome measurement.

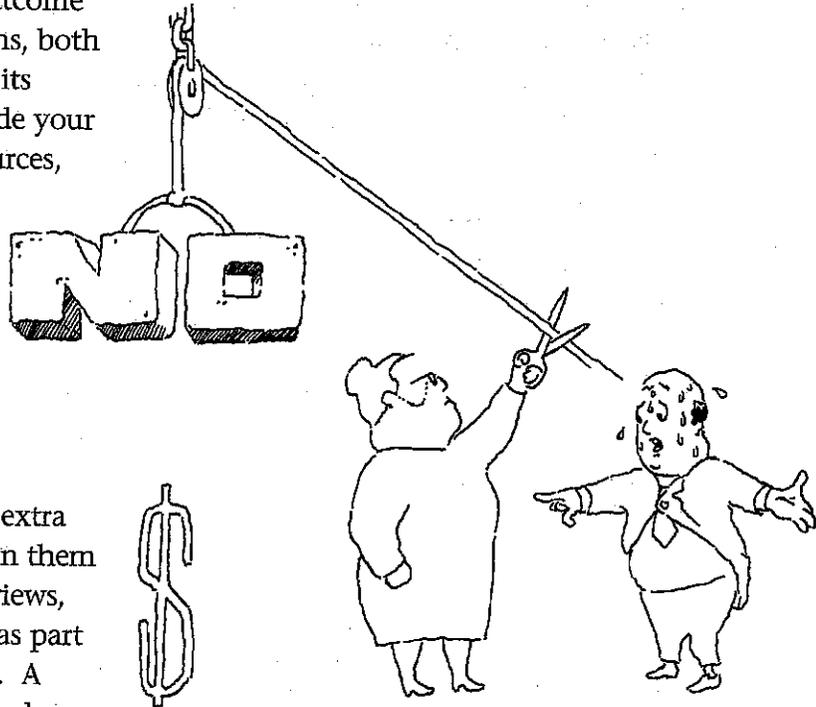
*Measuring Program Outcomes: A Practical Approach* outlines some of the trade-offs between cost and other considerations in its discussion of data collection methods in the appendix, which also provides suggestions for minimizing costs of mail surveys. Your work group will consider costs as it makes decisions at several points. Added costs will remain, however. Agencies that have incorporated outcome measurement as an ongoing part

of their operations include these in their core costs. Funders should recognize these as legitimate costs of a quality program.

- Expertise may be a problem. While some organizations have staff members with training in evaluation or research, many lack such in-house expertise. They are challenged to leverage resources creatively, perhaps arranging in-kind assistance from colleges or universities or the corporate community, seeking a capacity-building grant from funders committed to outcome measurement, or sharing the cost of specialized staff training or a consultant with other agencies. The two tasks where outside help is needed most often are designing data collection methods and analyzing collected data.
- There is concern that a focus on producing good outcome numbers may encourage some agencies to focus on participants or target groups most likely to show positive change. If followed, this practice of "creaming" (selecting certain participants because they are more prone to success) could leave the most challenging participants, who may be those most in need of help, without service. Program staff and boards should

discuss ways to guard against this problem, such as reviewing outreach practices and participant characteristics periodically and comparing findings to those of earlier periods.

- The inappropriate use of outcome findings in funding decisions, both within your agency among its various programs and outside your agency by your funding sources, can have unwarranted repercussions on programs. Exhibit 1-G provides guidelines for relating outcome data to funding decisions appropriately.



- A potential implication for program participants is the extra burden that will be placed on them if participant surveys, interviews, and/or testing are required as part of the data collection effort. A related concern is privacy and confidentiality. The more sensitive the information requested and the more individuals involved in collection and analysis, the more danger there is of serious breaches of confidentiality. The problem is compounded if the program intends to use personal vignettes or case studies in reporting or fund-raising materials. Whenever possible, complete confidentiality should be maintained in reporting outcome data. Step 4 provides guidelines for assuring confidentiality.

Many of these limitations and problems have no immediate answer. They are challenges to be addressed by agency staff and volunteers, as well as by funders and community policy makers.

### **Expectations of Agency Director, Board, Other Key Players**

The work group needs to know what expectations key individuals or bodies have for the process, timetable, costs, products, or other aspects of its work. Given the newness of this venture, these expectations may be vague. Nevertheless, the agency director, board, and others involved should try to identify what they have in mind regarding the group's work and share these thoughts when the group first meets.

## Available Resource Materials

The work group should review and discuss materials that are relevant to their task. These materials may include:

- Documents that describe the agency's purpose, mission, and programs.
- Recent program reports.
- Information on outcome measurement initiatives or program monitoring and assessment efforts in other parts of the agency.
- Descriptions of and reports from outcome measurement activities of other agencies.
- Any relevant resource materials from national organizations of which the agency or its staff are members.

**M**arc Center, Mesa, Arizona placed special emphasis on obtaining staff buy-in. The agency believes that direct-care providers and middle-management staff need to understand the complete picture because they collect all the data for outcome measurement. All staff learn how the information is generated and how it ultimately will be used.

Initially, representatives of all staff levels participated in a series of meetings to discuss their perception of outcome measurement and air their concerns. Then, through a series of focus groups, staff across all levels helped develop the system. Staff members were integrally involved in determining which outcomes are reasonable to track and how to track them.

Now, all staff levels, from direct care to middle and upper management, are represented on the Program Evaluation Outcome Measurement Committee. They analyze and interpret the data and then, peer to peer, share the results and proposed use of the information with their colleagues.

One of the measures Marc tracks is what it costs to have a client obtain gainful employment. Through outcome measurement, Marc determined that it can place a client for an average of \$1,450 and that the person will return the cost in taxes within one year of employment.

## Task 2: Decide Which Program(s) to Start With

Although some agencies that have multiple programs involve all of the programs in outcome measurement from the beginning, many start with just one or two programs. Before making a decision about which programs to start with, focus first on the question of what constitutes a program for the purpose of outcome measurement. Agencies assign the term "program" based on various criteria, including funding source and service delivery method. In outcome measurement, a program is a set of related activities and outputs directed at common or closely related purposes that a meaningful portion of the agency's resources is dedicated to achieve.

## Exhibit 1-G

### Guidelines for Incorporating Outcome Data in Funding Decisions

- Provide sufficient time (a year or more) to develop and test outcome measurement systems before launching full-scale implementation.
- Implement a “hold-harmless” period during which outcome findings have no bearing on funding decisions to allow managers time to make program adjustments based on their first outcome findings.
- Focus on whether managers are making good-faith efforts to measure their outcomes, rather than on the measurement results themselves, when initially including outcome issues in funding considerations.
- If outcome *findings* eventually are considered in funding decisions, do not consider outcome data in isolation as a measure of a program’s worth. Other factors such as number of participants, participant characteristics, staffing resources, service costs, and level of service combine with outcome data to explain the value of a particular program.
- Do not make assumptions about what level of performance constitutes a “good” outcome for a program. Without one or more rounds of experience with outcome measurement, there generally is not a sound basis for making this judgement. Even comparisons with similar programs are problematic unless the agency is sure that the compared programs serve comparable populations.
- In evaluating outcome data, focus on whether findings show program improvement from cycle to cycle rather than on how outcomes of one program compare to those of another, or on whether outcomes meet some ungrounded, intuitive target.

Your work group should examine your agency’s activities in light of this definition and group the activities into programs for outcome measurement. Then the group can decide, if the agency has more than one program, which one(s) should be the first for outcome measurement. The choice should maximize the likelihood of successful outcome measurement of an important part of the agency’s work. Programs should be considered for the initial outcome measurement effort if:

- They have a recognizable and reasonably defined mission and clientele.
- They represent a substantial portion of the agency’s activity.

“It’s best to start with small, easily measured programs where you are more likely to be successful at the end. Take something that is easy to track in order to get rid of the pitfalls. Choose a program that you feel is going to be successful or that you’ve always wondered about to help deflect the anxiety.”

Royce Van Evera,  
Executive Director,  
Law, Order & Justice Center,  
Schenectady, New York

- Funders or others have been asking about the results being produced by the program.
- Program supervisors and staff are likely to be supportive of the outcome measurement effort.

### Task 3: Develop a Timeline

"We had been working on a mission evaluation tool for over a year when the program committee of the board suggested a more comprehensive tool. At about the same time, we received United Way's tool. After some additions to the United Way document, it took us about six months to get a program evaluation tool up and running. And that was for 60 programs."

*Arthea J. Larson, Associated  
Catholic Charities,  
New Orleans, Louisiana*

Once the work group knows which program(s) will be the first for outcome measurement, it should create a timeline for the eight steps of the planning and implementing process. A sample timeline is provided in Exhibit 1-H; Worksheet 1 is a blank version for your use. Allow yourself plenty of time, including leeway for deadlines to slide. If your timeline is too tight, you may end up rushing through the process or overlooking important issues.

Considerations for creating your schedule include:

- Existing organizational deadlines or events that may affect scheduling of key steps and milestones. For example, the agency's funding cycle and fiscal year may determine when certain steps should begin or conclude. Annual board meetings or conferences may have other scheduling implications. Programs that are linked to the school year need to consider vacations in their calendar.
- Length of a typical course of service (e.g., a one-time counseling session, a six-week training program, a program of weekly activities throughout the school year).
- How long after the start or completion of service one would expect to see at least initial outcomes. Answers to this and the previous issue may affect the length of time needed for the trial run (Step 5). They also can influence the duration of program reporting cycles once full-scale implementation has started. Annual reporting cycles do not match the realities of every program.
- Availability of external resources you may need. For example, if you think you may use university faculty as consultants or student interns as aides, avoid scheduling major data collection or analysis tasks during their exams or breaks.

"It took us about two years to have our first round of outcome reports ready. The first year was spent getting the system in place and then six months to design the data tracking and another six months to do the management reporting."

*Randall L. Gray,  
Executive Director,  
Marc Center, Mesa, Arizona*

Exhibit 1-H

**Sample Timeline for Planning and Implementing Outcome Measurement in a Program**

Step	Initial Preparation							Trial Run				Implementation	
	Month							Month				Month	
	1	2	3	4	5	6	7	8-?	?+1	?+2	?+3	?+4	?+5
1: Get Ready	X	X											
2: Choose the Outcomes You Want to Measure			X	X									
3: Specify Indicators for Your Outcomes					X								
4: Prepare to Collect Data on Your Indicators						X	X						
5: Try Out Your Outcome Measurement System								XXXXX					
6: Analyze and Report Your Findings									X	X	X		
7: Improve Your Outcome Measurement System												X	
<i>Launch full-scale implementation</i>													X→
8: Use Your Findings									X	X	X	X	X→

This sample allows 7 months for initial preparation. The amount of time for Step 5, which starts in Month 8, is left open because it can vary so widely. After Step 5, the sample shows 3 months for Step 6 and 1 month for Step 7. Full-scale implementation starts 5 months after the data collection trial ends.

Worksheet 1

## Timeline for Planning and Implementing Outcome Measurement

Program: \_\_\_\_\_

	Month											
Step												
1: Get Ready												
2: Choose the Outcomes We Want to Measure												
3: Specify Indicators for Our Outcomes												
4: Prepare to Collect Data on Our Indicators												
5: Try Out Our Outcome Measurement System												
6: Analyze and Report Our Findings												
7: Improve Our Outcome Measurement System												
<i>Launch full-scale implementation</i>												
8: Use Our Findings												

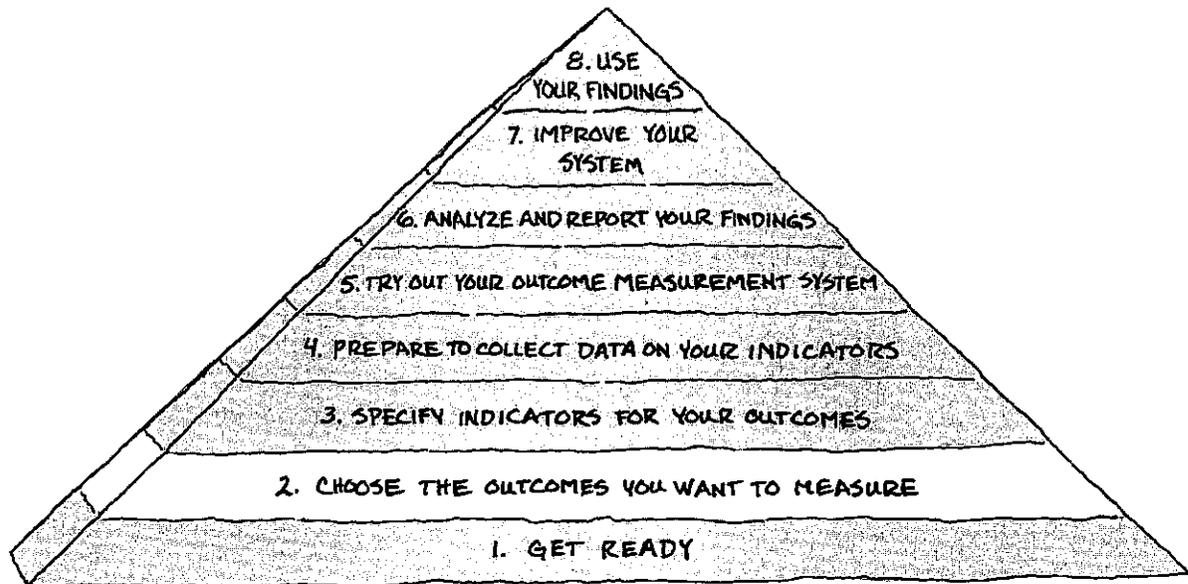
The work group should consider its initial timeline to be tentative until it has gained agreement about outcomes to be measured (Step 2) and data collection procedures to be used (Step 4). Decisions made during these steps may prompt the group to amend the schedule for the trial run. Later, the program's experience with either the data collection trial or the initial data analysis and reporting (Steps 5 and 6) may cause further changes to the schedule. It is far more important to address whatever issues arise effectively than to meet deadlines set before the challenges of the task were clearly defined.

#### **Task 4: Distribute Your Game Plan to Key Players**

Once your work group has defined the purposes of your agency's outcome measurement effort, reached a common understanding of terms and expectations, and developed a preliminary timeline for developing an outcome measurement system, be sure that everyone who will be involved knows what is planned. This includes all levels of staff, volunteers involved in service delivery, board members, and members of committees that will have a role in or be affected by the plan (e.g., strategic planning, public and community relations, resource development). If funders have provided important impetus for your outcome measurement effort, you may want to inform them as well.

Even at this early stage, amendments to your game plan may be needed. Be prepared to take extra time, if necessary, to resolve issues and gain agreement. As soon as your plan has approval, you can begin the next step of deciding which program outcomes you want to measure.

## Step 2: Choose the Outcomes You Want to Measure



**T**he next step for your organization's outcome measurement work group is to identify the program outcomes that should be candidates for regular measurement. The aim is to define a set of outcomes that track the benefits participants experience during or after encounters with the program.

Do not be concerned at this point with whether the outcomes being considered are measurable or how the program would go about measuring them. Also, do not attempt to set targets for outcomes you identify. These tasks are addressed in later steps. For now, your challenge is to develop a list of useful program outcomes.

This effort requires careful thought. The outcomes selected in this step are the foundation for all subsequent planning and implementation activities. If the outcomes are not well conceived, the value of outcome measurement is diminished. Tasks that will help the work group complete this step effectively are:

1. Gather ideas for what your program's outcomes are from a variety of sources.
2. Construct a logic model for your program.
3. Select the outcomes that are important to measure.
4. Get feedback on your logic model and the outcomes selected for measurement.

## Levels of Outcomes



To this point, *Measuring Program Outcomes: A Practical Approach* has not distinguished among different levels of outcomes. This distinction is important, however. In many cases, there is not just one desired outcome for participants, but a series, each contributing to another benefit or change that leads closer to the ultimate outcome the program hopes to achieve for participants.

Various sources recognize different numbers of outcome levels and call them by different terms. The real issue is not the terms, of course, but the concept of a hierarchy of logically related changes or benefits. This manual identifies three levels of outcomes and uses the following terms to denote them:

- **Initial outcomes** are the first benefits or changes participants experience, and are the ones most closely related to and influenced by the program's outputs. Often, initial outcomes are changes in participants' knowledge, attitudes, or skills. They are not ends in themselves, and may not be especially meaningful in terms of the quality of participants' lives. However, they are necessary steps toward the desired ends, and therefore are important as indicators of participants' progress toward those ends.
- **Intermediate outcomes** link a program's initial outcomes to the longer-term outcomes it desires for participants. They often are changes in behavior that result from participants' new knowledge, attitudes, or skills.
- **Longer-term outcomes** are the ultimate outcomes a program desires to achieve for its participants. They represent meaningful changes for participants, often in their condition or status. Although the program may hope that participants go even further in their growth and development and that similar changes will occur throughout the larger community, the program's longer-term outcomes are the most removed benefits that it can reasonably expect to influence.

Outcomes are not intrinsically initial, intermediate, or longer-term. An intermediate outcome for one program can be a longer-term outcome for another. Their designation is determined by the logical relationship among the sequence of effects for the particular program. This relationship can be thought of as an "if *a*, then *b*" chain of influences and benefits. Basic examples of this "If-Then" relationship are:

- If stop-smoking classes teach about health hazards and effective quitting techniques (output), **then** smokers acquire knowledge, change their attitudes, and gain skills to stop smoking (initial outcomes). If smokers know smoking is harmful, want to quit,

and have skills to minimize withdrawal symptoms, **then** they will quit smoking (intermediate outcome). If they quit smoking, **then** they will have fewer smoking-related illnesses (longer-term outcome).

- If employment counseling for women trying to become financially self-sufficient includes job interviewing skills (output), **then** the women will give more effective job interviews (initial outcome). If they give better interviews, **then** they are more likely to find jobs (intermediate outcome). If they find jobs, **then** they are more likely to be financially self-sufficient (longer-term outcome).
- If youth are mentored by adults who stress the importance of education (output), **then** the youth will see education as important (initial outcome). If the youth see education as important, **then** they will attend school more regularly (intermediate outcome). If youth attend school more regularly, **then** they are more likely to graduate (intermediate outcome). If they graduate, **then** they are more likely to become employed (longer-term outcome).

"In the old days we would have had a team of professionals. Now we let customers tell us how their quality of life was enhanced. We asked our clients 'what's important?' Their responses cluster into eight broad domains and are tremendously consistent. Then we went back and asked what kinds of questions we should be asking. This insures we are consumer driven, not organizationally driven. Not only did our customers help develop the instrument, they're helping design the assessment of program quality."

*Randall L. Gray,  
Executive Director, Marc Center,  
Mesa, Arizona*

The relationship between program inputs, activities, outputs, and initial, intermediate, and longer-term outcomes is depicted in Exhibit 2-A. With this understanding of levels of outcomes participants experience, your work group is ready to begin identifying your program's outcomes.

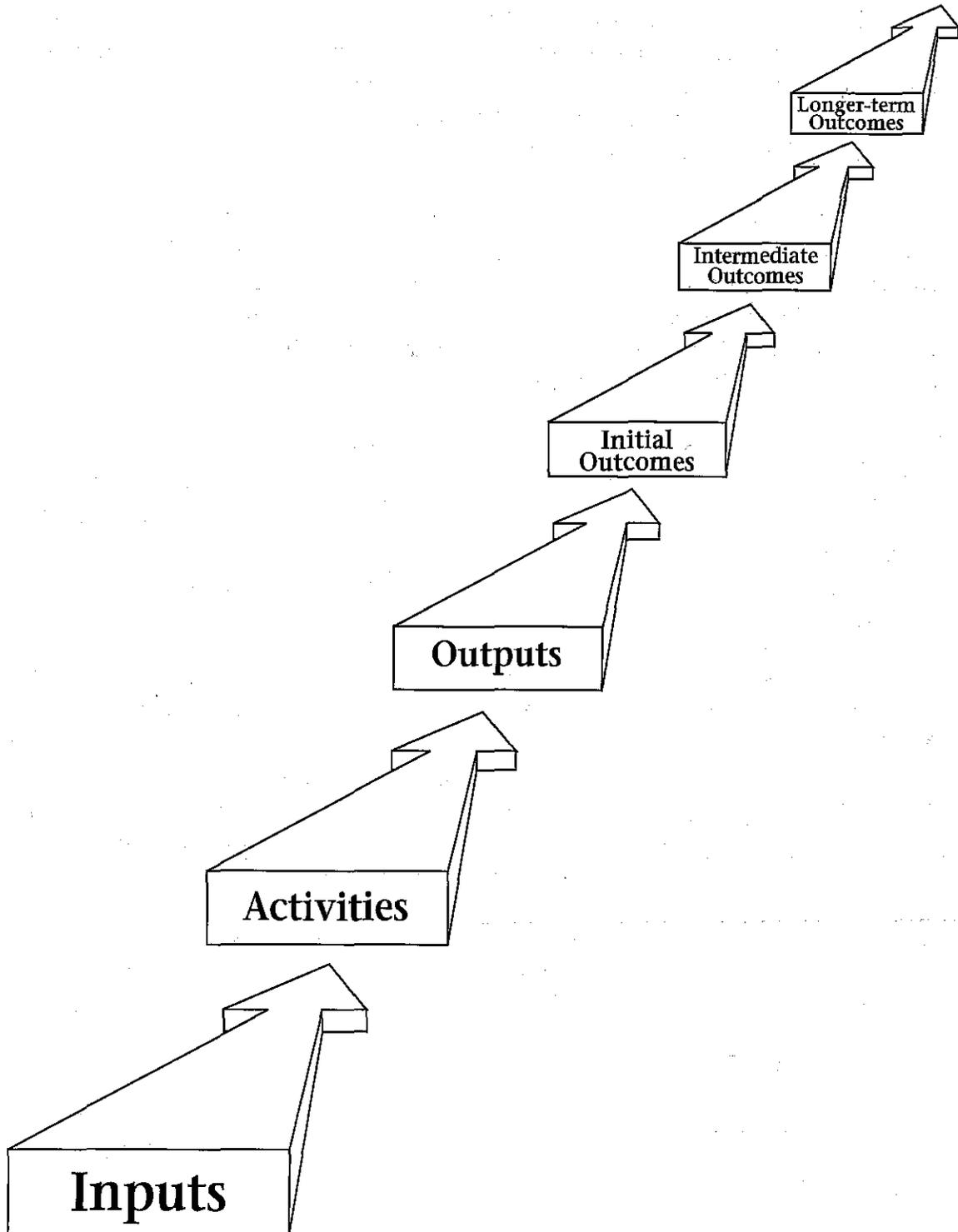
### **Task 1: Gather Ideas for What Your Program's Outcomes Are from a Variety of Sources**

The value of outcome measurement is enhanced when outcomes reflect many views of the program. Seek ideas from a variety of sources, both within and outside the program, before trying to refine a list of outcomes. Multiple perspectives will help the work group think about the program and its benefits for participants more broadly than if the group relies only on its own members. The result will be a richer set of outcomes that capture the program's most important benefits for participants.

The group also should seek ideas on unintended negative outcomes that might follow from the program.

For example, for a program that offers movie passes, fast food coupons, and video arcade tokens as rewards for improved grades, unintended negative outcomes could include

## Inputs Through Outcomes: The Conceptual Chain



students improving their grades by cheating and students being bullied to give up their prizes. The program will want to monitor these potential problems to see if such behavior is in fact occurring.

Following are some methods you might use to gain ideas regarding outcomes, including unintended negative outcomes, for participants.

- **Review agency or program materials** that might suggest what the program's results for participants or target groups are intended to be. Examples include the organization's most recent annual report of program activities, organization or program mission statements, articles of incorporation, director's statement in the annual report, statements of purpose in funding applications, and findings from past needs assessments.
- **Talk with program staff and volunteers who work directly with participants.** They can describe, for example, what the program does from their perspective, what aspects of the program they believe are most important to participants, and how they think participants benefit.
- **Meet with key volunteers from your board and relevant committees,** such as strategic planning and resource development. Volunteers' views of outcomes can provide insight into both community concerns and attitudes related to the program and how outcome measurement can yield information that communicates effectively with the public.
- **Talk with current and past participants,** either in informal conversations, or more formally through interviews, surveys, or focus groups. (Exhibit 2-B describes key steps in conducting a focus group. This manual's bibliography lists more comprehensive references on the topic.) Participants can tell you what they expected the program to do for them and what they expected the result to be. They also can tell you what they think the actual result of the program has been. (Maybe you're having outcomes you hadn't thought of!) In addition to helping identify sound outcomes, consulting with participants enhances the credibility of your outcome measurement effort.

"We include kids as part of the evaluation system. It gives the social workers a forum to talk about expectations. For example, the kids may want a safe place to hang out with friends, whereas we may want to work on conflict resolution. This gives us the opportunity to reach a common ground, to get some buy-in from them."

*Judy Selle, Executive Director,  
Neighborhood House,  
Milwaukee, Wisconsin*

## Conducting Focus Groups

1. Invite approximately 8-12 participants to participate in each focus group. Consider offering an incentive for participation.

Group members can be chosen from lists of current and former participants. The main criterion is that they have had experience with the program. If program participants are diverse in age, race/ethnicity, income, or other characteristics, the full range of your participants should be represented. Note, however, that conversations may be freer if each group is relatively homogeneous. If your program serves entire families, conduct separate focus groups with adults and young people.

2. Schedule the meeting for a maximum of two hours. Hold it in a comfortable, accessible location. Soft drinks and snacks might be provided.
3. Select an experienced focus group facilitator for the session.

The purpose and conduct of focus groups are different from those of task-oriented meetings. Choose a facilitator who knows and is experienced with the differences. Do not select as facilitator a staff person who works with any of the focus group members. This could inhibit an open sharing of participants' views.

4. After introductions and an overview of the session's purpose, the facilitator asks the participants open-ended questions to elicit their views on the benefits the program provides—what changes participants want, expect, or have experienced. Pose the question in several ways, such as:

- What did you expect to change for you as a result of being in the program?
- If the program really helps you, how will you be better off?
- Are there things you don't like about the program?

The facilitator encourages each participant to express her or his views on these questions. The facilitator's job is to establish an open, non-threatening environment and obtain input from each participant. To do this, the facilitator accepts anything participants say and does not attempt to explain the program, correct misconceptions participants may have, or gain agreement or consensus among group members.

5. Someone, such as a member of the work group, serves as recorder. Soon after the session, the reporter and facilitator summarize what was said.

Note that focus groups are purely qualitative in their method. The summary does not attempt to quantify how many persons expressed various views or seemed to share the same position.

For some programs, such as those dealing with young children and persons with diminished mental capacity, parents and other caregivers need to be included. These individuals often can point out not only desired outcomes, but unintended negative ones as well.

- **Review records of complaints** participants have made about the program. They may suggest outcomes participants expected to achieve, but did not. They also may reveal unintended negative outcomes.

**B**ig Brothers Big Sisters of Fresno County, Inc., in Fresno, California feels that doing their homework up front helped streamline the outcome measurement process. The organization invests approximately 40 hours per year to document the impact their program has on children and on the community.

First, the staff revised the mission statement to focus more specifically on the organization's current direction: juvenile delinquency prevention. Next, the executive director worked to identify variables that can contribute to juvenile delinquency. She read periodicals to understand the national thinking, interviewed people in the penal system, and then consulted the local United Way to gain a community perspective.

She discovered that the strongest predictive variables include low-income, single-parent families; behavioral problems in school; problems with authority figures; loss issues; and the existence of alcohol or drug problems combined with physical and/or sexual abuse. After pinpointing these variables, the staff had a clear idea of changes it wanted to promote in its participants. They went on to develop outcome indicators and measures enabling caseworkers to obtain more specific and useful information from the volunteers, parents, and children.

This process helped the agency track the progress of children in the Big Brothers Big Sisters program and conclude that after their involvement with a volunteer mentor, less than one percent of the children in the program become involved with the juvenile justice system.

- **Talk with representatives of agencies that might be a "next step" for former participants** of your program. For instance, if your program prepares individuals for employment, it could be valuable to talk with employment counselors or employers about attributes or characteristics that are desirable in their participants. An early childhood education program might talk with kindergarten teachers about attributes of school readiness. Obstetricians and nurse midwives could provide insights to a prenatal care program. A youth internship director could be helpful to a program that provides youth mentoring. Such sources also may relate concerns that suggest potential unintended negative outcomes.

- **Obtain materials from other organizations that deliver similar services**, both within your community and in other communities. Also check with national professional associations and trade organizations whose members represent your type of service. Outcomes, standards, or objectives such groups have identified may be applicable to your program as well.

- **Collaborate with other programs that have missions and participants similar to yours.** Pooling your differing perspectives and experiences may help you identify a more comprehensive and focused set of desired and negative outcomes.
- **Ask staff members from a different department or informed outsiders to observe your program** as it operates and infer outcomes from what is observed. Exhibit 2-C is an example of what observers knowledgeable about school-age child care saw during an afternoon of observing the recreational component of an after-school program for inner-city children and the outcomes they inferred the program was having for participants.

From a one-time observation, one simply would note behaviors that appear to be desirable. The work group then would explore whether these desirable behaviors should be considered program outcomes. Such observations are very helpful in identifying outcomes for youth development programs and social programs for the elderly where there is a general feeling that the program is doing participants a lot of good, but it is hard to pinpoint the benefits clearly.

- **Talk with funders who are promoting outcome measurement.** Note that funders should not be telling you what your outcomes are. They can, however, clarify their expectations regarding outcome statements, and they may have suggestions for your program based on experience with other programs with similar purposes serving similar participants.

## **Task 2: Construct a Logic Model for Your Program**

After you have gained input on your program's outcomes from many sources, organize and refine them by developing a logic model diagram for your program. A program logic model is a description of how the program theoretically works to achieve benefits for participants. It is the "If-Then" sequence of changes that the program intends to set in motion through its inputs, activities, and outputs. Logic models are a useful framework for examining outcomes. They help you think through the steps of participants' progress and develop a realistic picture of what your program can expect to accomplish for participants. They also help you identify the key program components that must be tracked to assess the program's effectiveness.

"The outcome focus was a wonderful way of looking at things. In the process we spent a lot of time on connotations of the word 'impact.' 'What is happening to this girl? What are the positive pieces to her involvement?' We had used the same words before, but didn't have the back-up documentation and knowledge. A part of this process really focused our leaders' efforts, giving them encouragement and ideas for program delivery."

*Mary Beth Malm,  
Executive Director,  
Girl Scouts of Milwaukee  
Area, Milwaukee, Wisconsin*

Exhibit 2-C

**Examples of Outcomes Based on Observations of a Neighborhood After-school Program**

Observations	Comments	Inferred Outcomes
<ul style="list-style-type: none"> <li>■ Several children with learning disabilities or attention deficit disorder are able to work for relatively long periods of time.</li> </ul>	<p>The persons making the observations note that the children's attention span is longer than what they would have expected from children with these learning problems. They talk to the after-school supervisor, who explains that she has been working with these children to increase their attention span. However, this outcome has never been tracked.</p>	<ul style="list-style-type: none"> <li>■ Children with learning disabilities and/or attention deficit disorder have increased attention span.</li> </ul>
<ul style="list-style-type: none"> <li>■ Children help each other with homework and games.</li> <li>■ Children praise positive achievement of peers.</li> <li>■ Children discourage name calling and other negative interactions.</li> <li>■ The children don't swear while in the program.</li> <li>■ The children are able to organize themselves when using the video games and foosball table. They wait until it is their turn.</li> </ul>	<p>The observers note that the behavior of the children in the program is markedly different from the behavior at the local school and nearby playground and streets. Staff explain that not only do the children understand that strict behavioral norms exist in the program, but that the children themselves try to maintain these norms. Again, this desired behavior has never been tracked and is therefore not reported to funders and outsiders.</p>	<ul style="list-style-type: none"> <li>■ Children provide mutual support and encourage positive behavior of peers.</li> <li>■ Children learn self-control.</li> <li>■ Children develop social skills.</li> </ul>
<ul style="list-style-type: none"> <li>■ The children seek guidance from the adults.</li> </ul>	<p>The observers note that the children respect and trust the adults and go to them for help with a wide range of problems, ranging from simple questions about a homework assignment to personal and family problems. In talking to staff, they learn that many of the children have few positive adult role models in their lives, especially male role models.</p>	<p>Children demonstrate:</p> <ul style="list-style-type: none"> <li>■ Increased trust of caregiving adults.</li> <li>■ Increased use of caregiving adults as resources.</li> <li>■ Decreased sense of helplessness and isolation.</li> </ul>
<ul style="list-style-type: none"> <li>■ Children seek out new challenges without fear of failure.</li> </ul>	<p>The observers note the children's initiative and apparent sense of confidence. They suggest that these attributes be measured and compared with that of children who are not in the program.</p>	<p>Children demonstrate:</p> <ul style="list-style-type: none"> <li>■ Increased self-confidence.</li> <li>■ Increased choice of new activities and experiences.</li> </ul>

---

**Note:** Measuring program outcomes does not mean that you will stop collecting information on inputs, activities, and outputs. If you find that some outcomes fall short of your expectations, you will need information on these other program components to locate the source of the problem and make improvements. Once you have implemented outcome measurement, however, you may find that you do not need all of the data the program has been collecting, enabling you to reduce data collection efforts in some areas.

---

Logic models usually are diagrammed as a series of boxes representing program inputs, activities, outputs, and outcomes. They can be drawn either horizontally or vertically. To provide examples, Exhibit 2-D introduces the Southside Children's Agency, a fictitious, multi-service community organization with four distinct programs. Exhibit 2-E displays horizontal logic model diagrams depicting each program's inputs, activities, outputs, and initial, intermediate, and longer-term outcomes. You can use Worksheet 2 to create a horizontal logic model diagram for your program. Exhibit 2-F shows the same logic models for the programs of the Southside Children's Agency diagrammed in a vertical format.

## Southside Children's Agency

The Southside Children's Agency is a fictitious multi-service neighborhood center focused on enhancing the well-being of children in a specific geographic section of a city. The agency's services include parenting education, child care, public awareness, and neighborhood organizing. The four specific programs that are followed throughout this manual are:

- **Teen mother parenting program.** The Southside neighborhood has a high incidence of births to teenagers. These babies are disproportionately victims of abuse or neglect. This program targets mothers in the sixth month of pregnancy. It provides a classroom-style prenatal and parenting education program in the high school for one hour, twice a week. The program continues through the child's first birthday.
- **Developmental child care program for children of teen mothers.** The agency has identified that there is a very high incidence (almost 90%) of pregnant teenagers dropping out of school following the birth of their child. One of the key reasons identified by these mothers is lack of child care. The agency also knows that there is a lack of quality, affordable child care in the area. They have therefore established a developmental child care program located on the high school grounds for children of high school-age mothers. Only children of mothers who attend school are eligible for the program.
- **Child abuse public awareness and prevention program.** This program targets community residents, and particularly members of the education and medical professions, with information about signs of child abuse and appropriate reporting procedures for suspected abuse. The purpose of the program is to prevent child abuse from recurring by identifying it in time to intervene. However, the agency itself does not provide services for abused children or their parents.
- **Neighborhood organizing program.** The agency's mission includes creating a safe and nurturing environment for the neighborhood's children. It has formed a neighborhood organizing program to organize the residents in self-managed projects to create clean, safe, play areas and develop Crime Watch and Block Parent programs.

## Exhibit 2-E

## Example Logic Models for Four Programs of Southside Children's Agency

Program: Teen Mother Parenting Education

Inputs	Activities	Outputs	Outcomes		
			Initial	Intermediate	Longer-term
Agency provides MSW program manager, part-time RN instructor, nationally certified education manuals, videos, and other teaching tools.  Agency and high school identify pregnant teens to participate in program.	Program provides parenting classes on prenatal through infant nutrition, development, safety, and caretaking delivered in high schools twice a week for one hour to teen mothers from 3 months prior to one year after delivery of a child.	Pregnant teens attend program.	Teens are knowledgeable of prenatal nutrition and health guidelines.  Teens are knowledgeable of proper care, feeding, and social interaction with infants.	Teens follow proper nutrition and health guidelines.  Teens provide proper care, feeding, and social interaction to their babies.	Teens deliver healthy babies.  Babies achieve appropriate 12-month milestones for physical, motor, verbal, and social development.

**Discussion:** This program has two "If-Then" sequences—one related to prenatal health and the other to infant care. Both outcome sequences contribute to the longer-term outcome of appropriate 12-month milestones for the babies.

Program: Developmental Child Care

Inputs	Activities	Outputs	Outcomes		
			Initial	Intermediate	Longer-term
Child care licensing is acquired and maintained.  Agency provides director with M.A. in child development, 3 full-time child care workers, facilities and equipment for child care center.	Developmental child care is provided on school grounds for children of teen mothers attending high school.	Children are enrolled in child care.	Children are present at child care regularly.  Teen mothers attend school regularly.		Children exhibit progress toward age-appropriate motor, verbal, and social skills.  Teen mothers graduate from high school.

**Discussion:** This program has no intermediate outcomes. This does not mean that the program is of short duration or that it is not sophisticated. It means that the program expects a straightforward relationship between its initial and longer-term outcomes: if children regularly attend developmental child care, then they will exhibit progress toward developmental milestones, and if teen mothers regularly attend classes, then they will graduate (see further discussion on page 49). One might ask why attendance—children's regular presence at child care and the teen mothers' regular attendance at school—is considered an outcome for the program. In most cases, program attendance is an output. In this example, however, the children's presence at child care depends on the mothers' actions and the desired outcome for the teen mothers is attending school, not the program. Using the child care service and attending school are behaviors of teen mothers that the program wants to influence in order to achieve the longer-term outcomes. This example illustrates why program context is so important to developing and understanding a logic model.

Exhibit 2-E (cont'd)

Program: Child Abuse Public Awareness and Prevention

Inputs	Activities	Outputs	Outcomes		
			Initial	Intermediate	Longer-term
Agency has one full-time staff member with an MSW and 10 years experience in child welfare, including research on indicators of child abuse and neglect. Program is supported by a volunteer advisory board representing local media, medical and education professions, and pro bono assistance from a local advertising agency.	Agency designs public service announcements (PSAs), brochures, posters, and presentation materials on signs of child abuse and neglect and tests them for clarity, interest, and acceptability.	<p>PSAs air on radio and television.</p> <p>Brochures distributed at community events, to school administrators, pediatricians' offices, child care centers.</p> <p>Posters placed in stores and buses.</p> <p>Presentations made to community civic groups, medical and school professionals' meetings by program director, agency director, and advisory board members.</p>	General public, educators, child care workers, and medical professionals hear or read information.	Target audiences are knowledgeable of signs of child abuse and neglect and of appropriate actions to take.	Target audiences identify and report suspected abuse and neglect.

**Discussion:** Note that this program does not claim to reduce abuse and neglect rates in the community. In fact, since such rates generally are based on the number of reported cases, the rate actually may increase for a short period after this program begins. Relating one program alone to community-level change is very risky. (See page 51 for further discussion.)

Program: Neighborhood Organizing

Inputs	Activities	Outputs	Outcomes				
			Initial	Intermediate		Longer-term	
Agency has 2 community-building staff, meeting space, and board representing major constituencies of the community.	Agency staff and a team of board members hold community meetings to discuss proposed response to community-identified need of a safer neighborhood for children.	Residents attend meetings.	Residents sign up to clean up vacant lots and build playgrounds.	Volunteers for each vacant lot meet and develop action plans.	Volunteers create clean, drug-free play areas.	Children play in clean, safe areas.	Residents feel neighborhood is safer place for children.
		Action plans developed for 3 priorities: <ul style="list-style-type: none"> <li>■ clean play areas</li> <li>■ crime watch</li> <li>■ block parents</li> </ul>	Residents sign up as Crime Watch organizers for their block.	Residents attend "Block Parties" for Crime Watch training.	Residents follow Crime Watch guidelines.	Crime in participating blocks decreases.	
		Residents sign up to serve as Block Parents to provide emergency support for children on their block.	Block Parents are trained.  Parents and children know of program.	Children go to Block Parents in emergencies.	Fewer children are harmed by accidents or are victims of crime.		

**Discussion:** The initial outcomes of this program involve residents signing up. Why would this be an outcome? The target for this program is the neighborhood. By increasing citizen participation and voluntarism, the agency benefits the neighborhood. This is the first behavior the program must influence through its organizing program. Therefore, it is the initial outcome.

## Worksheet 2

**Program Logic Model**

Program: \_\_\_\_\_

Inputs	Activities	Outputs	Outcomes		
			Initial	Intermediate	Longer-term

Exhibit 2-F

## Example Logic Models for Four Programs of Southside Children's Agency

Program: Teen Mother Parenting Education

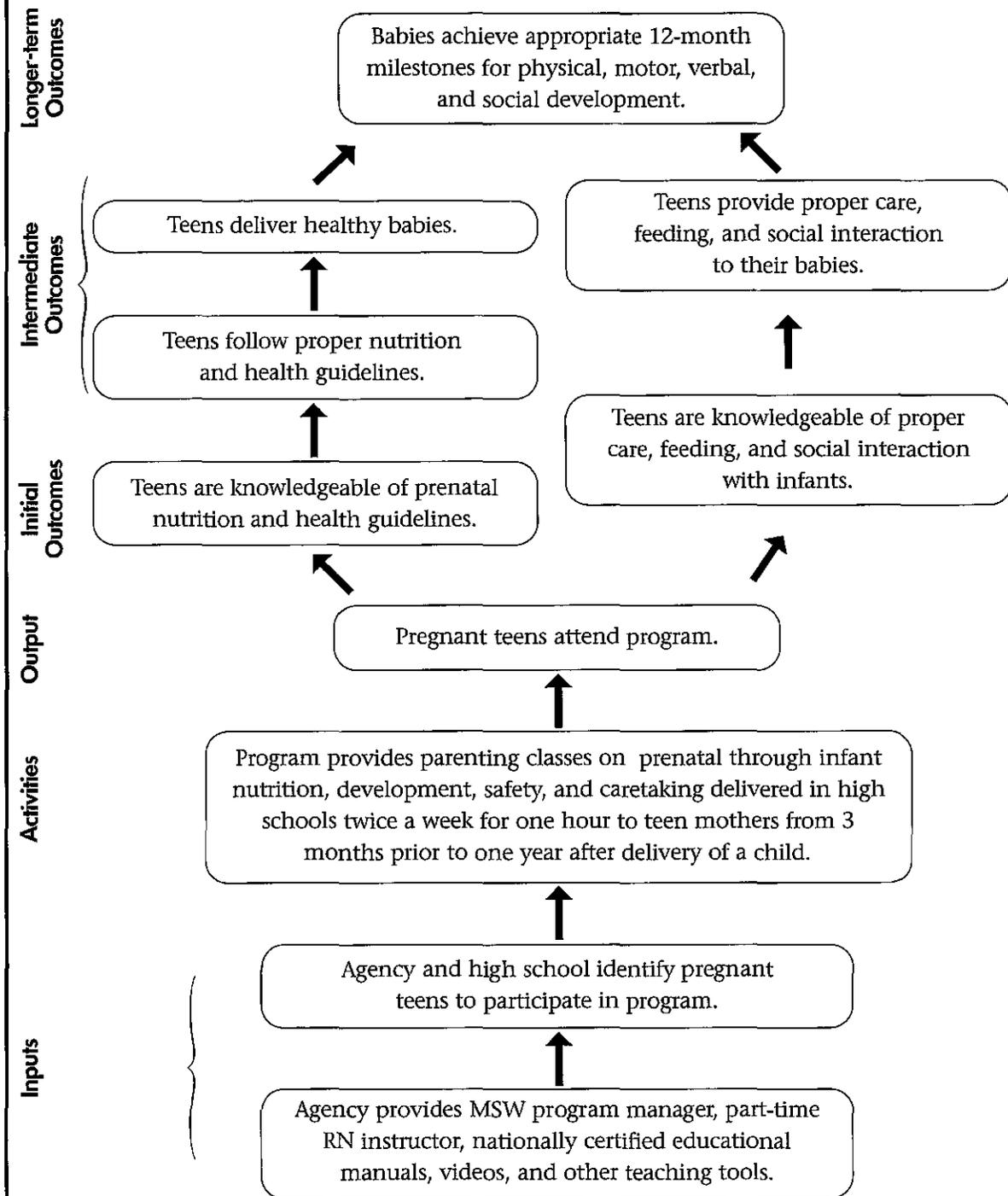


Exhibit 2-F, cont'd

Program: Developmental Child Care

Longer-term  
Outcomes

Children exhibit progress toward age-appropriate motor, verbal, and social skills.

Teen mothers graduate from high school.

Initial  
Outcomes

Children are present at child care regularly.

Teen mothers attend school regularly.

Output

Children are enrolled in child care.

Activities

Developmental child care is provided on school grounds for children of teen mothers attending high school.

Inputs

Agency provides director with M.A. in child development, 3 full-time child care workers, facilities and equipment for child care center.

Child care licensing is acquired and maintained.

Exhibit 2-F, cont'd

Program: Child Abuse Public Awareness and Prevention

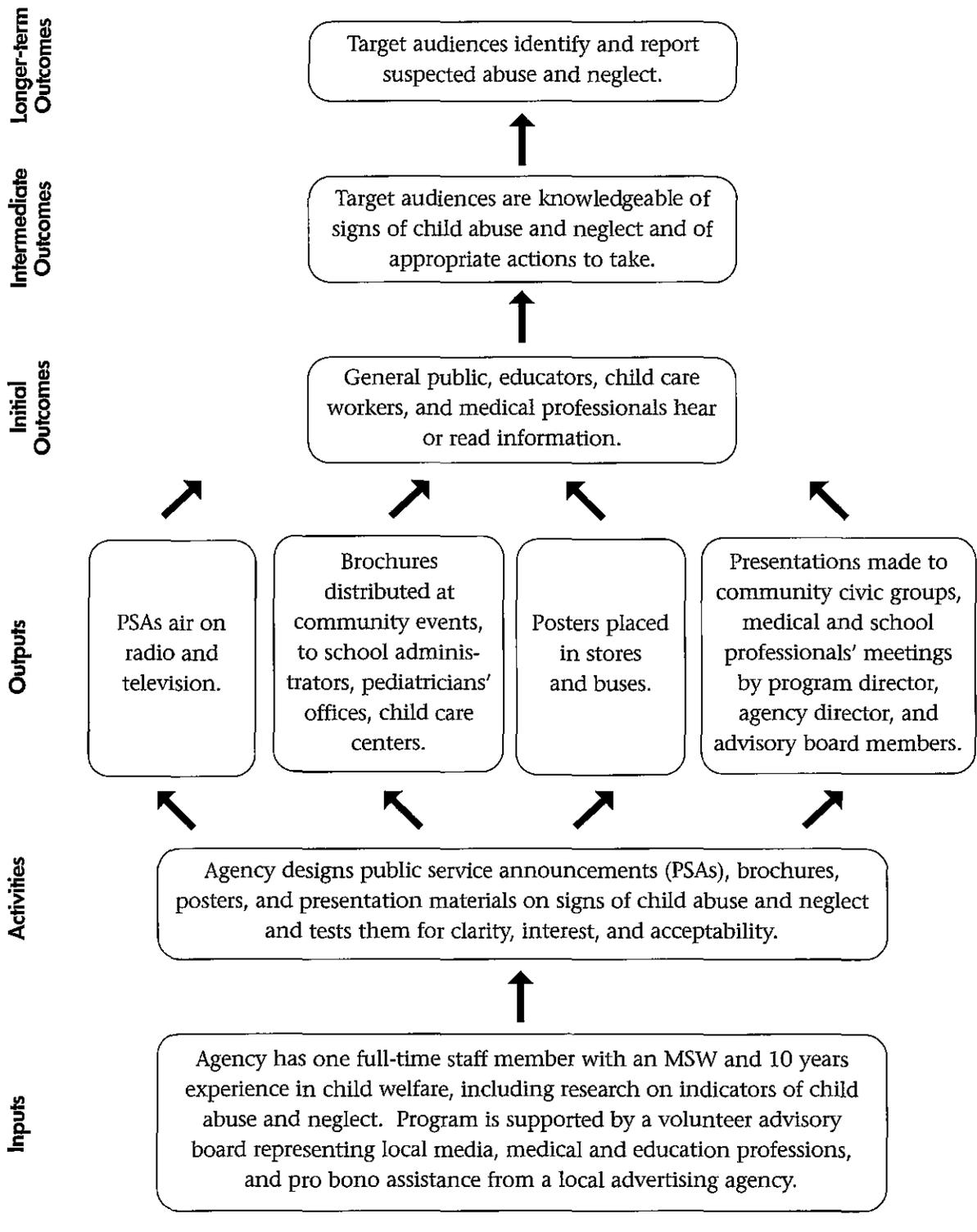
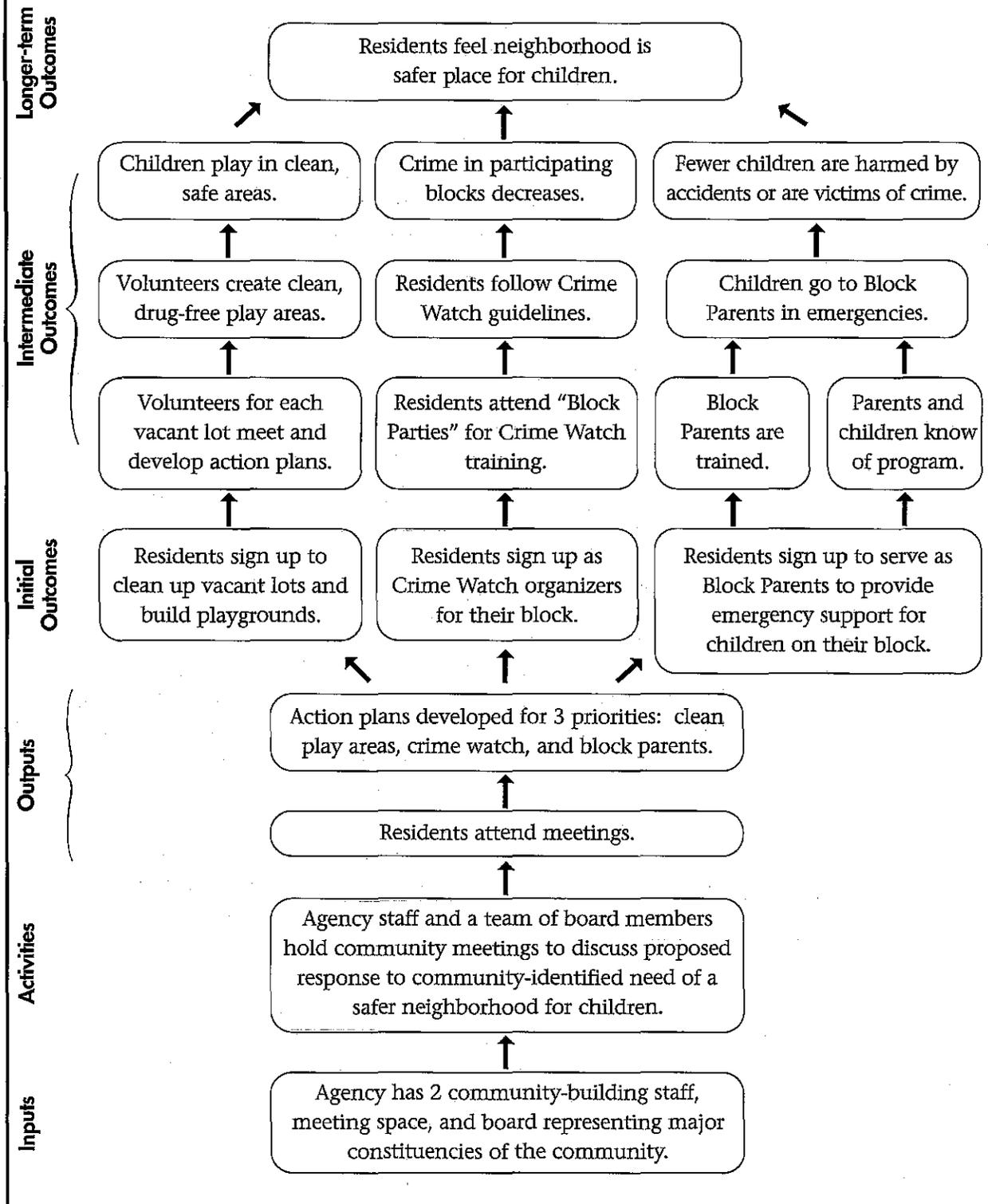


Exhibit 2-F, cont'd

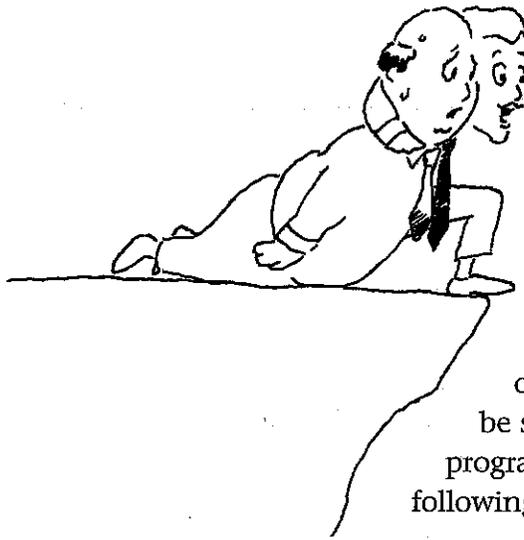
Program: Neighborhood Organizing



## Issues in Defining Outcomes

The example logic models illustrate a number of principles for identifying appropriate program outcomes. For example:

- **There is not a right number of outcomes for a program.** The Child Abuse Public Awareness and Prevention program of the Southside Children's Agency has three outcomes while the Neighborhood Organizing program has fourteen.
- **Some programs may have more than one "outcome track,"** which means that the same outputs set in motion multiple "If-Then" chains. The Developmental Child Care program has two outcome tracks—one for teen mothers related to school graduation and one for their children related to developmental milestones. The Neighborhood Organizing program has three tracks of initial and intermediate outcomes leading to one longer-term outcome of a safer neighborhood for children.
- **For some programs, initial outcomes may be arguably closer to outputs.** This is because initial outcomes rarely represent major change. In Exhibit 2-E, the discussions regarding the Developmental Child Care program and the Neighborhood Organizing program illustrate the importance of context in deciding whether program components that at first look like outputs may on occasion be initial outcomes. To help answer the question, "Is this an outcome?," refer to the descriptions of outputs and outcomes in the Introduction and to Exhibits 1-B, -C, and -D in Step 1. Recognize, however, that flexibility and interpretation often are required. Most important, keep in mind that it is not the "box" an item fits in that matters, but whether your logic model includes the appropriate outcomes.
- **The more immediate the outcome, the more influence a program generally has on its achievement.** In the example Teen Mother Parenting Education program, the initial outcomes of teens being knowledgeable of prenatal nutrition and health guidelines and of appropriate care, feeding, and social interaction with infants are largely the result of the staff's skill in teaching the information and the quality of the teaching materials and approaches the program uses.
- **Conversely, the longer term the outcome, the less direct influence a program has over its achievement and the more likely other, extraneous forces are to intervene.** In the Parenting Education program example, babies of teen mothers achieving appropriate 12-month milestones could be affected by the mothers' home environment, the knowledge level of others who care for the babies, and the availability of nutritious food for the infant—none of which are within the program's control.



A major challenge in identifying longer-term outcomes is determining “how far” to extend a program’s logic model—that is, what is the longest-term outcome for which the program should be held accountable. It should be far enough on the “If-Then” chain to capture meaningful benefit or change for participants and to reflect the full extent of the program’s influence on participants. On the other hand, it should not be so far beyond the program’s outputs that the program’s influence is washed out by other factors. The following principles may help with this determination.

- **Just because other forces may affect an outcome does not mean that it should be excluded from a program’s logic model.** For example, the purpose of the Southside Children’s Agency Teen Mother Parenting Education program is to improve the well-being of babies born to teen mothers in the first year of life. It is important for the program to measure and track this outcome, even though other factors, such as the mother’s home environment, might affect participants’ success. The program will try to track those other factors so that it can understand if less-than-desired outcomes are due to circumstances outside of the program’s control. (Step 3 further discusses factors that may influence outcomes.)

As another example, one of the longer-term outcomes of the Developmental Child Care Program is that the teen mothers graduate. The program’s “theory” regarding this outcome is that, by providing child care as a supplement to existing services, more teen mothers will graduate. However, this program deals with only one requisite for graduation: attendance. It does not deal, for instance, with grades or discipline. Other programs in the school address these issues by providing mentoring and tutoring to at-risk students. The Southside Children’s Agency program fills a gap in a larger logic chain of comprehensive services to teen mothers. As the program monitors the outcome of graduating, it also will monitor the girls’ grades and use of the mentoring and tutoring services as factors that may affect their success.

- **On the other hand, a program’s longer-term outcomes should not go beyond the program’s purpose.** For example, the Child Abuse Public Awareness and Prevention program might be tempted to specify as a longer-term outcome that families experiencing child abuse and neglect receive treatment in early stages. However, this program is designed to increase case finding; it does not provide treatment for child abuse and neglect. Because achieving this outcomes depends on action by some other organization, it is not an appropriate outcome for this program.

- **Similarly, a program’s outcomes should not go beyond the scope of its target audience.** Another longer-term outcome that might occur to the Child Abuse Public Awareness and Prevention program is that, as a result of early case finding, the overall rate of child abuse and neglect in the community declines. However, this outcome calls for community-wide change that requires action by many more individuals than just those who receive information from this program. Even if the program works and its target audiences receive, recall, and act on information it provides, if overall abuse and neglect in the community rise, perhaps related to increases in unemployment or reduction in treatment resources, Southside Children’s Agency would be holding itself accountable by setting this as a desired outcome for its program. The outcome may be part of a community logic model for a collaborative strategy to address child abuse and neglect, but it is not an appropriate outcome for this one program alone.

In the same way, the Teen Mother Parenting Education program may serve only 50 pregnant teens a year. The entire community is not its target; an outcome that the infant mortality rate in the community declines is beyond the scope of its target audience. However, the Neighborhood Organizing program does have the entire Southside neighborhood as its target. Therefore, increasing residents’ feeling of safety is an appropriate target for this program.

A basic question to help clarify whether a longer-term outcome is appropriate for your program to measure is, “Can our program alone have a significant influence on this outcome?” This does not require that your program alone has total or complete influence over the outcome, but that if the program is implemented properly, participants are fully engaged, and counteracting negative events do not intervene, the outcome should occur. If you cannot answer this question with “yes,” the outcome probably is not one for which you want to be held accountable.

One of the criticisms of using a logic model, or program objectives, as a framework for outcome measurement is that it ignores consequences of the program that may be unintended. A final principle in identifying program outcomes is:

- **It is important to consider carefully what unintended and possibly negative consequences your program may have for its participants or the community.** For example, the Developmental Child Care program may create an appealing picture of teenage parenthood by having babies in the school and allowing their mothers to visit them during classes and receive special programs. The Southside Children’s Agency will monitor both the attitudes of other girls in the school toward teen pregnancy and the pregnancy rate in the school, assessing whether it begins to rise.

Another example of negative, unintended consequences can be found in the example Neighborhood Organizing program. It is not uncommon to find that when residents effectively organize against crime and see it decline, it actually rises in adjoining neighborhoods. This happens because the criminals move to a less-challenging environment as a result of the neighborhood's organization.

## Preparing Your Logic Model

After information sources have been tapped, the work group can make a first attempt at constructing a program logic model. One approach is to give each work group member a pad of large, self-sticking notes. Before the group meeting, ask members to review all the information from the various sources they have tapped and write one idea about a program input, activity, output, or outcome on each note.

At the meeting, hang several sheets of flipchart paper on the wall. Using Worksheet 2 as a model, mark off sections for "Inputs," "Activities," "Outputs," "Initial Outcomes," "Intermediate Outcomes," and "Longer-term Outcomes." Group members can post the sticky notes they have made on the chart, move them around, and draw arrows between them to create "If-Then" chains showing the logical sequence of program effects from inputs through activities and outputs to outcomes.

"It's important for anyone who is getting into outcome measurement to sit down with the people who carry out the programs and come to a consensus on the program objectives. A lot of us assume there's a clear understanding of the objective and the desired outcome. But oftentimes we find out that's not the case—we have different expectations and directions."

*Pam Ferrell,  
Executive Director,  
YWCA of Summit City,  
Akron, Ohio*

Don't rush to finalize your model. Your work group may need to make several passes at this task. The issues discussed earlier about levels of outcomes and appropriateness of accountability for your agency take time to consider and debate. You may want to go outside the work group to get feedback on the model from your board, other staff of your organization, or funders. Agencies that have gone through the activity of identifying outcomes report that it takes substantial time and effort. Many program personnel are likely to have difficulties in shifting from a focus on internal activities and outputs to a focus on outcomes.

Don't be surprised to find a lot of variation in views of what your program's outcomes should be. One value of this step, aside from its contribution to the development of an outcome measurement system, is that it makes the variations visible, allows development of consensus, and gets the key players—staff and volunteers—focused in the same direction.

### Task 3: Select the Outcomes That Are Important to Measure

Your pursuit of input from varied sources and your work on a logic model are likely to produce a long list of outcomes for your program. The next task is to decide which of these outcomes your program should try to measure.



Do not consider whether an outcome appears to be measurable, or how much effort it might take to measure it. Focus on identifying a set of outcomes that reflect the program's intended benefit for participants and any potential negative outcomes.

Begin by weeding out outcomes that are duplicative, overlapping, or clearly unimportant. Although agency executives, board members, and those external to the program, such as funders and potential participants, are likely to be most interested in achievement of longer-term outcomes, do not be tempted to drop initial and intermediate outcomes from your measurement approach. Program managers require this information to inform them about all of the elements of the program and to pinpoint problem areas.

For example, in the Teen Mother Parenting Education program, the program may learn that babies are not achieving 12-month milestones. If it does not monitor initial and intermediate outcomes, it cannot determine if this is because the nutrition and health information provided to the teen mothers is unintelligible; because the teens know what they should eat but do not change their behavior; because the program's instructions on infant care are confusing; or because the teen mothers are not caring for their infants appropriately, even though they know how.

After you have eliminated duplications and insignificant outcomes, evaluate each outcome still on your list against the following criteria. Worksheet 3 may help with this task.

- Is it reasonable to believe the program can influence the outcome in a non-trivial way, even though it can't control it? (Be careful about implying that your program will influence community-wide change when in fact it aims to influence only program participants, not the entire community.)

## Assessing Possible Outcomes

Program: \_\_\_\_\_

Outcome	Does it meet test...					
	1?		2?		3?	
	Yes	No	Yes	No	Yes	No
	Yes	No	Yes	No	Yes	No
	Yes	No	Yes	No	Yes	No
	Yes	No	Yes	No	Yes	No
	Yes	No	Yes	No	Yes	No
	Yes	No	Yes	No	Yes	No
	Yes	No	Yes	No	Yes	No

Test 1: Is it reasonable to believe the program can influence the outcome in a non-trivial way, even though it can't control it?

Test 2: Would measurement of the outcome help identify program successes and help pinpoint and address problems or shortcomings?

Test 3: Will the program's various "publics" accept this as a valid outcome of the program?

- Would measurement of the outcome help identify program successes and help pinpoint and address problems or shortcomings?
- Will the program’s various “publics”—staff, volunteers, participants, collaborating organizations, funders, and the general community—accept this as a valid outcome of the program?

After looking at each outcome individually, examine the set of outcomes you are considering.

- Do program outputs and initial, intermediate, and longer-term outcomes relate to each other logically? (To check this, walk through the implied “If-Then” relationships as shown for the three examples on pages 32 and 33.)
- Do these relationships reflect the logic of the program—the sequence of influences and changes that program inputs, activities, and outputs are intended to set in motion?
- Do the longer-term outcomes represent meaningful benefits or changes in participants’ status, condition, or quality of life?
- Have you identified potential negative outcomes of the program?

**S**taff of the **Adult Literacy Program of Literacy Volunteers of America-Schenectady, Inc., of Schenectady, New York** struggled with the best way to evaluate its literacy program. Although standardized tests typical of the educational system would be easy to use, the staff rejected them because they are designed to test reading levels for children. Staff knew that adults tend not to learn sequentially, and that standardized tests would not measure the nuances of adult learning.

To provide a more honest portrayal of their clients’ achievements, they decided to quantify practical applications of literacy. For example, English-as-a-second-language students might strive to write a letter, borrow a book from a library, or speak English to a doctor. They also decided that literacy outcomes need not be sequential and should not be limited to a list of predetermined achievements. Instead, they should be established by the clients themselves. The program’s desired outcome is for participants to demonstrate at least two new applications of literacy skills.

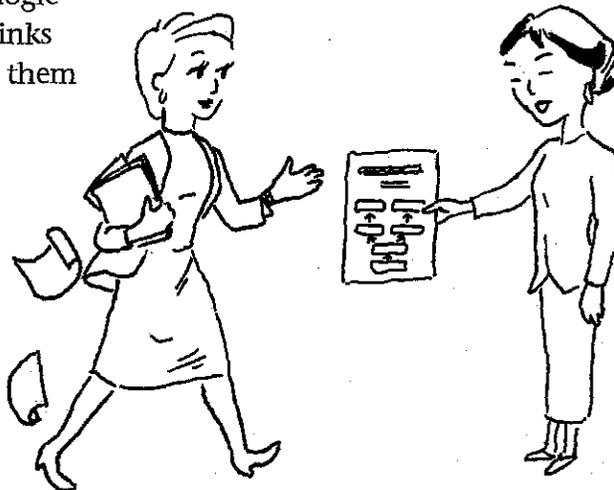
The agency recalls one example illuminating the wisdom of this strategy: A Chinese man was struggling to learn English. For him, the first milestone was successfully ordering a pepperoni pizza over the telephone.

Do not exclude outcomes because the program already does well in those areas. Take credit for them. For example, a health or safety program may have reduced the number of incidents of an important health or safety problem almost to “zero.” Unless you are certain that the problem has been wiped out completely, continue to track and report the outcome.

## Task 4: Get Feedback on Your Logic Model and the Outcomes Selected for Measurement

After your work group has drafted a logic model and identified the outcomes it thinks should be measured, distribute copies of them and seek reaction from:

- Individuals who are in positions to comment on the proposed model and outcomes insightfully.
- Individuals whose support and participation will be important for later planning and implementation.



This includes agency and program managers, service delivery staff, volunteer leaders, representatives of national service or professional associations, and funders that are asking programs to specify outcomes. Exhibit 2-G lists questions for reviewers' consideration.

Feedback from reviewers may result in further work group discussions of the program logic model and the outcomes that are important to measure. Time taken to fine-tune ideas and reach consensus now will pay off later as the group works from a stronger and more broadly supported conceptual foundation.

**T**he Samaritan Counseling Center of the Capital Region in Scotia, New York struggled with establishing outcome targets for its Project Recovery program, which counsels adults who were sexually abused as children. To adjust for the different stages of client entry and progress, the agency focused on measures of improvement, rather than of completion.

For example, one desired outcome is that clients show reduction in the intensity of nightmares, improvement in performing daily tasks, and reduction in suicidal ideation after participating in the program for six months. With this outcome, treatment remains individualized, but there is consistency in evaluation across the program.

Once there is agreement on the program's logic model and the program outcomes to be measured, the next step is to determine what information will help you measure your program's achievement on the outcomes. This is the subject of Step 3.

Exhibit 2-G

## Questions Regarding Proposed Program Logic Model and Outcomes

Does the logic model:

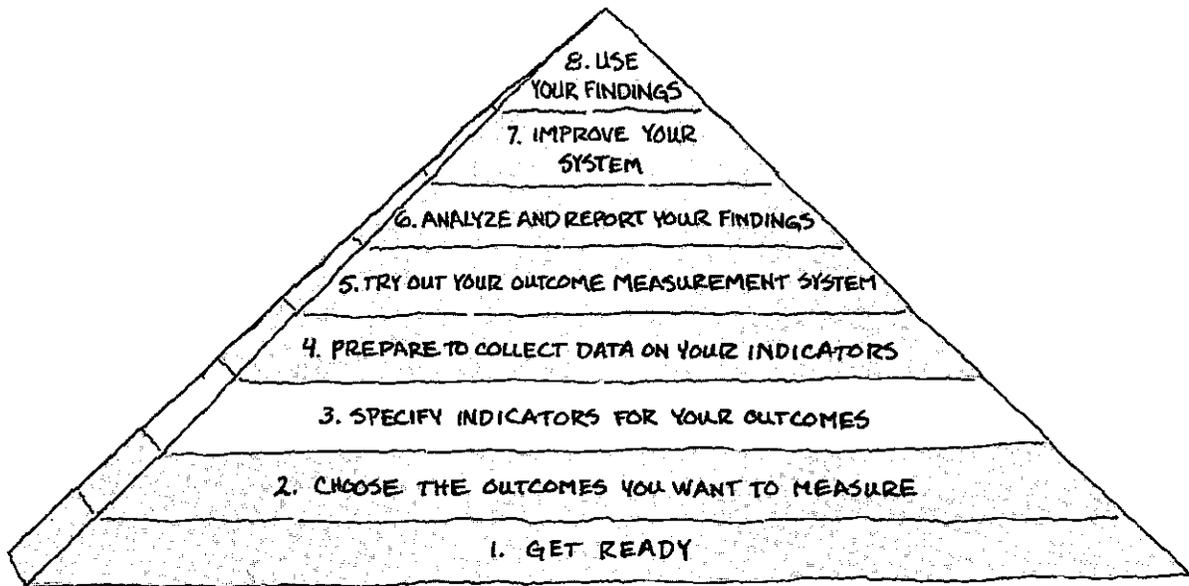
- Include all the activities and outcomes that are important?
- Make the appropriate connections between inputs, activities, outputs, and outcomes?

Are the outcomes identified as important to measure:

- Relevant to the mission/objectives of the program?
- Outcomes for which the program should be held accountable?
  - Are they important to achieve if the program is to fulfill its mission or objectives?
  - Do they represent meaningful benefits or changes for participants?
  - Is it reasonable to believe the program can influence them in a non-trivial way, even though it can't control them?
- Clear in defining the intended scope of the program's influence?
- Useful to program managers in efforts to identify both points of success and problems the program can correct?
- Likely to be effective in communicating the benefits of what you do to your program's various publics?



## Step 3: Specify Indicators for Your Outcomes



**A**fter completing Step 2, you have a list of outcomes that describe your program's desired benefits for participants. Steps 3 and 4 guide you through the process of creating a system to measure them. Step 3 focuses on the challenge of deciding just what data you want to collect. Step 4 helps you identify data sources and data collection methods.

As you address Steps 3 and 4, you will complete Worksheet 4 (page 60) one column at a time. For now, fill in the first column by listing the outcomes you want to measure. Then you are ready to begin Step 3. The tasks involved are:

1. Specify one or more indicators for each outcome.
2. Decide what factors could influence participant outcomes.

### Task 1: Specify One or More Indicators for Each Outcome

A critical challenge in outcome measurement is deciding specifically what information will indicate how well the program is doing regarding an outcome. This information

## Outcome Measurement Framework

Program: \_\_\_\_\_

Outcome	Indicator(s) (may be more than one per outcome)	Data Source	Data Collection Method

is called the outcome *indicator*. The purpose of an indicator is to help you know whether an outcome has been achieved. Specifying an outcome indicator requires deciding:

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

Some outcomes are fairly easy to observe and measure, such as graduating from high school or getting a job. Many outcomes, however, are not concrete enough to measure directly. For example, the desired outcome of a prenatal counseling program might be that participants' infants are born healthy. This certainly is a desirable outcome, but how would the program determine whether a newborn is "healthy?" If a program defines a healthy infant as one who weighs at least 5.5 pounds, has an Apgar score of at least 7, and has no observable physical abnormalities, the outcome becomes measurable. The measurable outcome indicator then becomes: Babies that at birth weigh at least 5.5 pounds, have an Apgar score of at least 7, and have no observable physical abnormalities.

An indicator may not capture all aspects of an outcome, and some outcomes may require more than one indicator. In the "born healthy" example in the previous paragraph, the indicator does not exhaust all aspects of being born healthy. It does not, for instance, include being free of problems whose detection requires time and testing beyond what typically is done in the delivery room, such as cerebral palsy, HIV, or fetal alcohol syndrome. A second indicator of a healthy birth could be the absence of these or other problems of concern.

Outcome indicators must be observable and measurable. If a condition is not observable and measurable, it may relate to the outcome, but it is not useful as an indicator. Indicators also must be unambiguous. Terms such as "substantial," "acceptable," and "adequate" (e.g., "Participants demonstrate substantial improvement") that are subject to interpretation are not sufficiently specific.

Numerical targets (e.g., "Participants demonstrate at least a four-point improvement") meet the criterion of being unambiguous. However, during its first experience with outcome measurement, a program may have little basis for picking numerical targets. Rather than set a numerical target that may turn out to be unrealistic, it generally is wiser to aim simply for "improvement" or "gain" (e.g., "Participants demonstrate improvement"). With indicators stated in this fashion, *any* increase or advancement constitutes achievement of the outcome. The program can establish numerical targets later, after a round or two of outcome measurement reveals what numbers are reasonable.

In many cases, it is clear from the outcome statement that the outcome is expected to occur, and therefore will be measured, directly after service delivery. If another timeframe is appropriate, the outcome indicator should state this, specifying whether time is measured from the point of initiation of service or the point of service completion. Examples are, "Youth demonstrate more positive interaction with peers after two years involvement with an adult mentor," and "Youth placed with their families following residential treatment will remain there for at least 12 months."

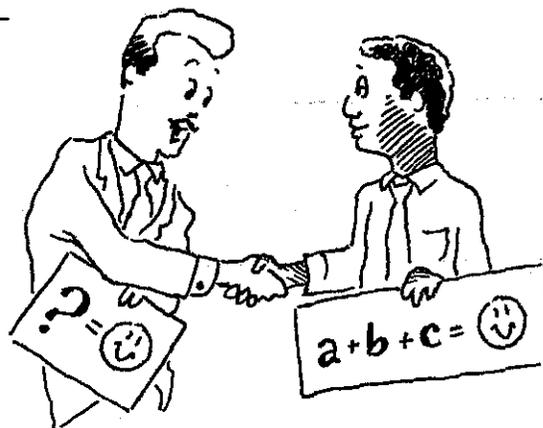
"We sat down with all the youth-serving groups and discussed, 'What does a healthy child look like?' We determined that he or she should have a sense of independence, of hope, of contribution to the community, of self, of worth, of belonging, and of closeness in relationships. Then we discussed how to measure these qualities in young people we work with. It really forced us to focus on what we are doing."

Mary Beth Malm,  
Executive Director, Girl  
Scouts of Milwaukee Area,  
Milwaukee, Wisconsin

Your work group should begin specifying indicators for your program outcomes by identifying the specific, observable accomplishments(s) or change(s) that will tell you whether each outcome has been achieved. This can be very challenging and may require much thought and discussion. Ask yourselves how you can tell if the outcome has been achieved. What does the outcome *look like* when it occurs? How do you *know* it has happened? What do you *see*? Keep refining the indicators by asking, "Can we observe and measure this? Does it tell us if the outcome has been reached?"

Do not exclude an indicator because it seems too simple. Some indicators are indeed quite straightforward, once the outcome has been defined. Note that, while development of indicators starts in Step 3, you are likely to amend some indicators during Step 4 as you make decisions about data sources and collection methods. You may make further refinements as a result of the pilot test described in Step 5.

The data you *collect* will be individual data—that is, data on individual participants. However, you will not *report* individual data. Instead, you will combine the data for individual participants into program-level information. For example, while a desired outcome for each participant may be improvement in reading skills, the program's achievement on this outcome is the aggregate accomplishment of all participants.



As you define measurable indicators of your program's outcomes, consider what statistics will show this overall accomplishment. Most often, the specific statistic programs use to convey this program-level information are the total *number* of participants achieving the outcome and the *percent* of participants achieving the outcome. In the reading-skills example in the previous paragraph, for instance, the program's success in achieving the outcome is the number and percent of participants whose reading skills improve.

Generally, if you will report the percent of participants who achieved an outcome, you also should report the number. This tells the reader whether your 60 percent, for example, refers to three participants out of five, or to 600 participants out of 1,000. In some instances, only a number can be reported, such as the number of action plans a team of participants develops or the average number of job interviews participants arrange. Occasionally, other statistics such as ratios or rates may be appropriate.

Outcome indicators developed for the four programs of the Southside Children's Agency are shown in Exhibit 3-A. Exhibit 3-B gives other examples of indicators. As your work group identifies measurable indicators for your program outcomes, you can complete the second column of the Outcome Measurement Framework worksheet (Worksheet 4).

Now check your list of outcome indicators for the following:

- Is there at least one indicator for each outcome?
- Does each indicator measure some important aspect of the outcome that no other indicator measures? (A few good indicators are better than many indicators that lack focus.)
- Is the wording of each indicator sufficiently specific? Does it tell you what characteristic or change you will count?
- Does each indicator identify the statistic(s) that will summarize the program's performance on the outcome? Will the statistic(s) convey your level of achievement effectively?

"I have found that the program staff are already evaluating the effectiveness of the programs. They tend to observe changes in the participants but need to document their observations. It doesn't have to be scientific evaluation to be useful or credible."

*Pam Ferrell, Executive Director,  
YWCA of Summit City,  
Akron, Ohio*

Exhibit 3-A

**Sample Outcome Indicators: Southside Children's Agency**

Program: Teen Mother Parenting Education

Outcome	Indicator(s)
Teens are knowledgeable of prenatal nutrition and health guidelines.	<ul style="list-style-type: none"> <li>■ Number and percent of program participants able to identify food items that are good sources of major dietary requirements</li> </ul>
Teens follow proper nutrition and health guidelines.	<ul style="list-style-type: none"> <li>■ Number and percent of participants within proper ranges for prenatal weight gain</li> <li>■ Number and percent of participants that do not smoke</li> <li>■ Number and percent of participants that eat at least 4 calcium servings and one of each of the other nutritional groups each day</li> <li>■ Number and percent of participants that take a prenatal vitamin each day</li> </ul>
Teens deliver healthy babies.	<ul style="list-style-type: none"> <li>■ Number and percent of newborns weighing at least 5.5 pounds and scoring 7 or above on Apgar scale</li> </ul>

Program: Developmental Child Care (outcomes for teen mothers only)

Outcome	Indicator(s)
Teen mothers attend school regularly.	<ul style="list-style-type: none"> <li>■ Of teen mothers enrolled in high school whose children are enrolled in child care program, number and percent who miss less than 3 days per month during school year</li> </ul>
Teen mothers graduate from high school.	<ul style="list-style-type: none"> <li>■ Of teen mothers whose children are enrolled in child care program, number and percent who graduate from high school over a four-year period</li> </ul>

Exhibit 3-A, cont'd

Program: Child Abuse Public Awareness and Prevention

Outcome	Indicator(s)
General public, educators, child care workers, and medical professionals hear or read information.	■ Percent of target audiences that recall content of PSAs, brochures, posters, or presentations
Target audiences are knowledgeable of signs of child abuse and neglect and of appropriate actions to take.	■ Percent of target audiences that can name 3 signs of abuse and neglect and can identify police or child protective services as appropriate contact
Target audiences identify and report suspected abuse and neglect.	■ Percent of target audiences that say they have made such a report since hearing or seeing information

Program: Neighborhood Organizing (safe play area outcomes only)

Outcome	Indicator(s)
Residents sign up to clean up vacant lots and build playgrounds.	■ Number of residents on sign-up sheets
Volunteers for each vacant lot meet and develop action plans.	■ Number of action plans developed
Volunteers create clean, drug-free play areas.	■ Number of vacant lots that are free of litter, have grass or other appropriate ground cover, have play equipment, and are free of drug sales and/or use
Children play in clean, safe areas.	■ Number of children observed in play areas on one random, sunny day per week for 3 months after cleanup
Residents feel neighborhood is safer place for children.	■ Number and percent of neighborhood residents that say neighborhood is safer for children today than it was one year ago

Exhibit 3-B

**Example Outcomes and Outcome Indicators  
for Various Programs<sup>1</sup>**

Type of Program	Outcome	Indicator(s)
Smoking cessation class	Participants stop smoking.	<ul style="list-style-type: none"> <li>■ Number and percent of participants who report that they have quit smoking by the end of the course</li> <li>■ Number and percent of participants who have not relapsed six months after program completion</li> </ul>
Information and referral program	Callers access services to which they are referred or about which they are given information.	<ul style="list-style-type: none"> <li>■ Number and percent of community agencies that report an increase in new participants who came to their agency as a result of a call to the information and referral hotline</li> <li>■ Number and percent of community agencies that indicate these referrals are appropriate</li> </ul>
Tutorial program for 6th grade students	Students' academic performance improves.	<ul style="list-style-type: none"> <li>■ Number and percent of participants who earn better grades in the grading period following completion of the program than in the grading period immediately preceding enrollment in the program</li> </ul>
English-as-a-second-language instruction	Participants become proficient in English.	<ul style="list-style-type: none"> <li>■ Number and percent of participants who demonstrate increase in ability to read, write, and speak English by the end of the course</li> </ul>
Counseling for parents identified as at risk for child abuse or neglect	Risk factors decrease. No confirmed incidents of child abuse or neglect.	<ul style="list-style-type: none"> <li>■ Number and percent of participating families for whom Child Protective Service records report no confirmed child abuse or neglect during 12 months following program completion</li> </ul>

<sup>1</sup> These are illustrative examples only. Programs need to identify their own outcomes and indicators, matched to and based on their own experiences and missions and the input of their staff, volunteers, participants, and others.

Exhibit 3-B, cont'd

Type of Program	Outcome	Indicator(s)
Employee assistance program	Employees with drug and/or alcohol problems are rehabilitated and do not lose their jobs.	<ul style="list-style-type: none"> <li>■ Number and percent of program participants who are gainfully employed at same company 6 months after intake</li> </ul>
Homemaking services	<p>The home environment is healthy, clean, and safe.</p> <p>Participants stay in their own home and are not referred to a nursing home.</p>	<ul style="list-style-type: none"> <li>■ Number and percent of participants whose home environment is rated clean and safe by a trained observer</li> <li>■ Number of local nursing homes who report that applications from younger and healthier citizens are declining (indicating that persons who in the past would have been referred to a nursing home now stay at home longer)</li> </ul>
Prenatal care program	Pregnant women follow the advice of the nutritionist.	<ul style="list-style-type: none"> <li>■ Number and percent of women who take recommended vitamin supplements and consume recommended amounts of calcium</li> </ul>
Shelter and counseling for runaway youth	Family is reunified whenever possible; otherwise, youths are in stable alternative housing.	<ul style="list-style-type: none"> <li>■ Number and percent of youth who return home</li> <li>■ Number and percent of youth placed in alternative living arrangements who are in that arrangement 6 months later unless they have been reunified or emancipated</li> </ul>
Camping	Children expand skills in areas of interest to them.	<ul style="list-style-type: none"> <li>■ Number and percent of campers that identify two or more skills they have learned at camp</li> </ul>
Family planning for teen mothers	Teen mothers have no second pregnancies until they have completed high school and have the personal, family, and financial resources to support a second child.	<ul style="list-style-type: none"> <li>■ Number and percent of teen mothers who comply with family planning visits</li> <li>■ Number and percent of teen mothers using a recommended form of birth control</li> <li>■ Number and percent of teen mothers who do not have repeat pregnancies prior to graduation</li> <li>■ Number and percent of teen mothers who, at the time of next pregnancy, are high school graduates, are married, and do not need public assistance to provide for their children</li> </ul>

## Task 2: Decide What Factors Could Influence Participant Outcomes

The level of achievement on specific outcomes will, of course, be different for different participants. You want to know under what circumstances successful, and unsuccessful, outcomes have occurred. Your outcome measurement system should help you determine if these differences are related to particular participant or program characteristics. This, in turn, can raise questions that may help you offer the appropriate services to participants with particular characteristics, or improve services by replicating effective program characteristics in all service units.

To be able to link level of achievement on outcomes to participant and program characteristics, you need to identify those characteristics that may make a difference and include them in your data collection efforts. Then, you will be able to break the data out based on these characteristics and look for these relationships.

Exhibit 3-C illustrates the kind of information that can be derived from such breakouts. The first table in the exhibit shows outcome results for two program units. Based on these findings, it appears that Unit 1 achieved substantially better outcomes for participants than did Unit 2.

The second table tells a very different story. Here, the data have been broken out based on the level of difficulty of participants' problems. This table shows that when level of difficulty is taken into account, Unit 2 achieved better outcomes for both difficult and non-difficult cases than did Unit 1.

As this brief example shows, including data on influential factors, such as participant's age or level of difficulty when entering the program, will make outcome information more useful to program personnel. Types of information programs may collect to enable useful comparisons include:

- Demographic characteristics of participants
- Level of difficulty presented by the participant
- Geographic location of participant's residence
- Organizational unit (if there is, for example, more than one service office)
- Type or amount of procedure used to deliver the service

Exhibit 3-D discusses typical factors that may influence outcomes. Exhibit 3-E gives examples of influencing factors appropriate for some of the outcomes in the Southside Children's Agency example. In listing these influencing factors, it is important to consider all of the program's outcomes, as some things may influence some outcomes but not others.

Exhibit 3-C

**Example of Value of Outcome Data Broken Out by Influencing Factors**

Table 1: Outcome Findings by Program Unit

	Unit	
	1	2
<b>All Participants</b>		
Total	500	500
Number Helped	300	235
Percent Helped	60%	47%

Table 2: Outcome Findings by Program Unit and Level of Difficulty of Participants' Problems

	Unit	
	1	2
<b>All Participants</b>		
Total	500	500
Number Helped	300	235
Percent Helped	60%	47%
<b>Non-Difficult Cases</b>		
Total	400	200
Number Helped	300	160
Percent Helped	75%	80%
<b>Difficult Cases</b>		
Total	100	300
Number Helped	0	75
Percent Helped	0%	25%

## Participant and Program Characteristics That May Influence Outcomes

### 1. Demographic characteristics of participants

This information helps identify which categories of participants are achieving the desired outcomes and which are not. Participant characteristics that may be relevant to your program include the following:

- Age group
- Sex
- Race/ethnicity
- Educational level
- Household income group
- Household composition (such as size and number of children)
- Disability status
- Level of functioning

### 2. Level of difficulty presented by the participant

In programs designed to help participants solve problems, participants generally vary considerably in the extent of the problem and in their strengths and resources. More difficult situations mean that the program has a harder time achieving desired outcomes. As the example in Exhibit 3-C shows, breaking out outcome data by level of difficulty helps show whether this factor is related to successful outcomes.

Reporting outcomes by level of difficulty also can help reduce the temptation to make the program's performance look good by attracting and serving easier-to-help participants while discouraging service to more difficult and more-expensive-to-help participants. (This is not meant to indicate that programs should not attempt to screen out persons they know cannot be helped by the services offered or to focus on participants who are most likely to benefit from the program. One of the results of outcome measurement may be more fine-tuned screening of potential participants and the development of different, more effective programs for different participants.)

Program staff need to develop specifications for each level of difficulty. For example, the program might decide to break difficulty into three categories: very difficult to help, moderate difficulty, and minor difficulty. Staff will need to establish procedures for identifying level of difficulty: what specific information available in participants' records or gained through interviews or observations will be used and how it will be evaluated. Staff then should be trained on using this procedure and recording the findings regarding level of difficulty in the participant's record.

## Exhibit 3-D, cont'd

### **Geographic location of participant's residence**

Geographic breakouts might be by:

- Neighborhood
- Political boundaries (e.g., school district, elected officials' districts)
- Zip code
- Census tract
- City or county

Knowing the outcome of services in each geographical area will provide information to users on where outcomes are not going well and where they are.

#### **4. Organizational unit**

This is useful for programs that have more than one service delivery facility. Managers and supervisors should each have outcome information that pertains to their own area of responsibility. If managers and supervisors are only provided outcome data that aggregates their service responsibilities with those of others, the data will be less meaningful to them. On the other hand, managers may fear that this information will be used to rate them. You may not want to tabulate this type of breakout until managers are comfortable with its usefulness.

#### **5. Type or amount of procedure used to deliver the service**

Some programs might want to examine outcomes by the amount (e.g., eight hours versus 12 hours) or type (e.g., group sessions versus one-to-one counseling) of service provided to participants. Other programs might want to examine outcomes by type of referral, such as court referrals, referrals from other agencies, and self-referrals. This type of breakout is not often used. However, it has considerable potential to be useful to creative, innovative program personnel.

Outcome information broken out by types of procedure used can help programs undertake internal comparisons of various processes and procedures. The program can try out new procedures on some of its incoming participants, while continuing to use existing practices for others. The outcomes can be compared to see whether the new practice seems to be superior and should replace the existing practices.

## Exhibit 3-E

**Factors That May Influence Participant Outcomes: Southside Children's Agency**Program: Teen Mother Parenting Education

Influencing Factor	Data Source	Data Collection Method
Mother's age		
Mother's household income		
Number of parenting classes mother attended		

Program: Developmental Child Care (factors for mothers only)

Influencing Factor	Data Source	Data Collection Method
Mother's age		
Mother's grade point average		
Mother's participation in school-run mentoring and tutoring services		
Mother's living situation (living alone, with parents or guardian, with partner)		

Exhibit 3-E, cont'd

Program: Child Abuse Public Awareness and Prevention

Influencing Factor	Data Source	Data Collection Method
Target audience (general public, medical professionals, child care workers, educators)		
Type of information received (PSAs, brochures, posters, presentations)		
Number of times information received (1 time, 2 to 4 times, 5 or more times)		

Program: Neighborhood Organizing (clean play area activities only)

Influencing Factor	Data Source	Data Collection Method
Number of volunteers for each vacant lot		
Race and sex of volunteers, children, and community residents		

"People are sometimes afraid of outcome measurement. They are afraid that what they're doing might not come out right. They often think it's because they didn't test it right and they blame the instruments. But programs based on research and theory show up with good outcomes. People should draw upon the huge amount of excellent research being done in the social science field and funded by the government."

*Lynn McDonald,  
Consultant,  
Innovative Program  
Development,  
Family Service America,  
Madison, Wisconsin*

(For example, household income of a pregnant teen may influence her ability to eat healthy foods, but not her ability to learn which foods are healthy to eat.)

Your work group should consider which of the information in Exhibit 3-D and other participant or program characteristics are likely to affect outcomes in an important way. Avoid the costly and time-consuming temptation to collect lots of data because the opportunity is there. If you break your data out by too many categories, you may not have enough participants in each category to make meaningful comparisons. The more data on these factors you collect, the more you will need a computer to analyze the data.

Use the first column of Worksheet 5 to identify influential factors you think are likely to make a difference in participants' success in achieving the desired outcomes. You will complete the rest of the worksheet in Step 4.

This is a good time for your work group to update staff, volunteers, and other key players on your progress and solicit another round of feedback. You may receive some good suggestions for making the indicators more measurable, or for ways to break out the data that will reveal useful information.

### Special Problems With Hard-To-Measure Outcomes

Some characteristics of programs or their participants present special challenges in measuring outcomes. Problematic situations are discussed below.

- **Participants are anonymous, so the program cannot later follow up on the outcomes for those participants.** For programs such as hotlines and information and referral programs, it may not be feasible to measure longer-term outcomes. For these programs, the extent to which the desired longer-term outcomes have been achieved (such as preventing suicides or other

"We learned that you can't just look at the mission statement in establishing outcomes and indicators. You need to be realistic and look at resources and limitations. We looked at how the existing system works and what was already there that we can tap into. Don't create a new system for yourselves. Use what is already available."

*Michelle Gaynier,  
Director of Communications  
and Marketing, The Detroit  
Institute for Children,  
Detroit, Michigan*

Worksheet 5

**Outcome Measurement Framework: Factors That May Influence Participant Outcomes**

Program: \_\_\_\_\_

Influencing Factor	Data Source	Data Collection Method

harmful behavior by the caller) are difficult or impossible to assess because the calls usually are anonymous. Follow-up with participants or obtaining any further information about them thus is usually not possible. Surrogate, intermediate outcome indicators will likely need to be used. Consider attempting to ask a brief set of end-of-call questions, such as "Did you get the information you wanted? Can you tell me what you will do next about this problem?"



Another option is to ask callers if they have called before, and, if so, to ask them to rate their previous experience using a few basic questions, such as "Was the information you received accurate? Were you able to get the assistance you needed from the referral you were given?"

Another option is for the program to have a second, independent person listen in on a sample of calls and evaluate the assistance made by the hotline personnel. This process also would provide the program an opportunity to identify the training needs of personnel.

A referral service might arrange for agencies to which they make referrals to ask incoming participants how they learned of the agency, and then to report to the referral service the number and type of participants the referral agency had generated, and whether they were appropriate.

- **The assistance is very short-term.** Programs such as disaster assistance and emergency food, shelter, and utility assistance provide immediate and short-term assistance to participants. In effect, the mere provision of these services can be said to be an important outcome, and for some programs their sole purpose. In these cases, the program outcomes may equal the outputs: the amount of service and number of people helped. If possible, the program can seek feedback from these participants or a sample of them about the helpfulness of the service in terms of timeliness, availability, and courteousness of the staff.
- **One or more major outcomes of the program cannot be expected for many years, so that tracking and follow-up of those participants is not feasible.** The long-range effects probably can only feasibly be examined in special, ad hoc studies due to the cost and difficulty. However, intermediate outcomes usually can be expected and should be tracked. For example, a substance abuse treatment program could determine if the participant is substance free both at the end of the program and, say, one year later. The program probably cannot determine if the participant is substance free for

the rest of his or her life. Similarly, youth development programs may find it infeasible to track their participants to adulthood to determine if they are productive, contributing members of society, but they can track intermediate outcomes such as increased school performance, positive peer interaction, or participation in community and civic activities.

- **Participants may not give reliable responses because they are involved in substance abuse or are physically unable to answer for themselves.** Prior research shows that well-conducted interviews with clearly guaranteed anonymity appear to be able to obtain candid answers to questions about drug use and even criminal activity. For programs where participants are unable to provide feedback for themselves, feedback from appropriate relatives or friends might be used.
- **The outcomes sought may appear to be too intangible to measure in any systematic way.** Usually, some aspect of the outcome can be measured, although it may require using qualitative data (things that are not numeric) or creating proxies or surrogates for the outcome itself. For example, a youth development program may seek to “build character.” On the surface this may seem very intangible, but through a combination of expert judgement and qualitative observations, an indicator could be created. A panel of leaders from diverse segments of the community, for example, could come together to develop a checklist of “attributes of sound character” and to identify observable behaviors associated with each attribute. Providers then could use these materials to develop a rating system for reviewing youths’ behavior or assessing statements youth might be asked to write about how they had achieved the outcome.
- **Activities are aimed at influencing community leaders to take action on the part of a particular issue or group, such as advocacy or community action programs.** With these programs, there generally are longer-term outcomes such as passing or defeating legislation, creating or enforcing policies, or starting or stopping behavior of corporations or institutions in the community. Certainly, these agencies should track their success or lack of it on these outcomes. But the intermediate outcomes of these agencies may look very much like the activities or outputs of other agencies. Because their purpose is to influence others, the generation of activity on behalf of a cause, such as introduction of legislation, scheduling a public hearing on enforcement of a policy, letters generated to members of Congress, or the number of people attending a rally or event, are appropriately considered intermediate outcomes for such programs.
- **Activities are aimed at the whole community, rather than at a particular, limited set of participants.** These would be things such as public education programs intended to improve the quality of life of citizens, such as a smoking prevention or other wellness campaign. The longer-term outcomes would be changes in behavior by the program’s target population. Generally, these can be tracked by publicly

available data on the problem or issue being addressed, such as death from lung cancer or drunk driving, collected before and for a period after the information program began.

However, these longer-term outcomes may take quite a while to occur, data are not always available in a timely way, and intervening forces may counteract any gains the program might have made. One way to deal with concerns about lack of timely information is to identify intermediate outcomes such as increased awareness and knowledge and track them via some type of community survey (see Step 4). An even better way is to collect data on self-reported behavior changes such as reduced smoking, reduced drinking and driving, or increased breast self-exams, since there is a strong research link between those behaviors and the longer-term outcome of a lower rate of death. To deal with counteracting forces in the community, outcome data from a community survey can be broken out by those who saw the program's information and those that did not, thus attributing outcomes more directly to the program.

- **Activities provide support to other agencies rather than direct assistance to individuals.** These include management assistance programs, volunteer placement programs, and capacity-building activities. For these programs, the recipient agencies' view of the usefulness and helpfulness of the support may be an initial outcome. The intermediate outcomes would be related to a change in the knowledge and behavior of the agencies assisted. A longer-term outcome could be improved effectiveness of the recipient agency. It will probably be impossible to create a numerical outcome indicator for improved effectiveness, but a qualitative measure, such as the recipient agency's ability to state at least one way the service made their organization more effective within a year of service, could be developed.
- **Programs are trying to prevent a negative event from ever occurring.** Programs that are trying to prevent something from happening, such as child abuse, teen pregnancy, or teen smoking, face special challenges. For programs that provide prevention services directly to participants (as opposed to the general public, which was discussed previously), the program can conduct follow-up surveys of at least a sample of participants several months or more after services have ended to determine the number that did and did not experience the negative event the program was trying to prevent.

The methods discussed in this manual will tell you the proportion of your participants that did or did not have a certain negative event (such as teen pregnancy) occur. It will not tell whether it was less likely to occur than it would have been without your program. In other words, you will not know whether your program alone prevented something. In-depth impact research studies, using control groups, are the only way to determine definitively that a program prevented an event.

In implementing the **Girls Incorporated Preventing Adolescent Pregnancy™ Program, Girls Incorporated of Memphis, Tennessee** measures outcomes each year. The program measures knowledge and skill changes among participants with written tests, demonstration exams, and role-plays at the beginning and end of each course. It also informally tracks pregnancies as an indicator of behavior. For three years in a row, the staff have observed no pregnancies or childbirths among the 300 to 400 girls aged 6 to 18 while they were in the program.

Every five years, the agency measures girls' long-term progress. During 1991 and 1992, the staff located girls who had been 12 or older in 1989 who had participated in Girls Incorporated of Memphis programming. Through questionnaires and personal visits, staff measured indicators such as employment, high school graduation, pregnancy, and involvement with the juvenile justice system. They then compared those data to teen pregnancy and high school completion rates for Memphis neighborhood schools. The longer-term study has shown the lasting positive effects of participation in Girls Incorporated of Memphis programming.

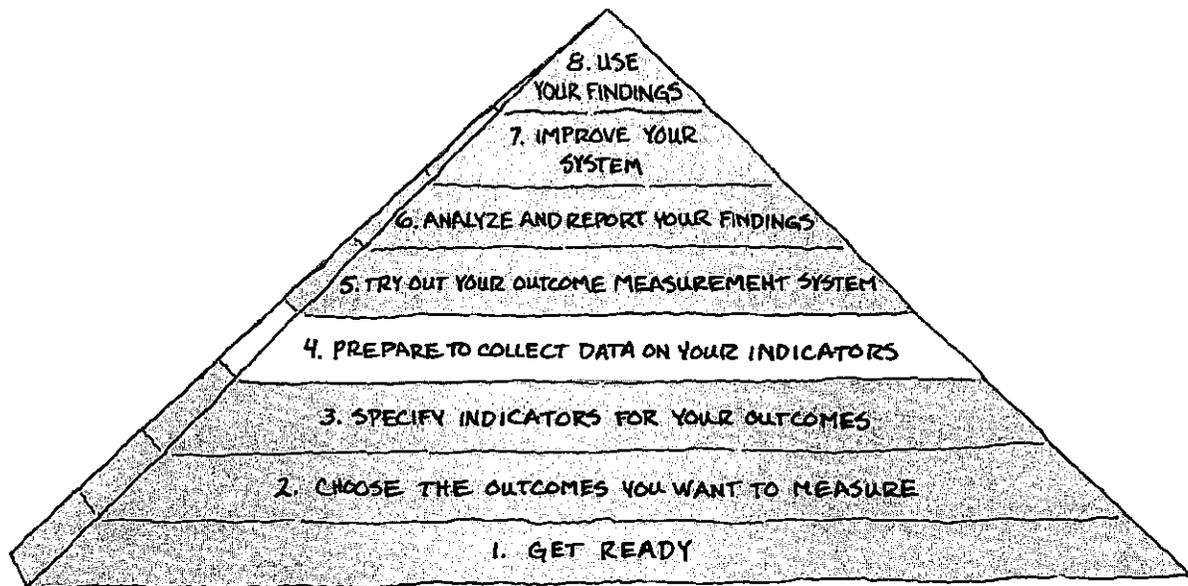
However, in the case of prevention programs, comparisons can be made between the participant group and a comparable group from the general public to determine if the participants' outcomes are better than would otherwise be expected. Step 8 discusses sources of comparative data.

Programs whose outcomes are not expected to occur until years after delivery of the service face special additional challenges addressed earlier in this section. In this case, because the program may

not be able to track the longer-term outcome, it may be necessary to use intermediate outcome measures as surrogates of the longer-term outcome. Particularly in the case of prevention programs, it is important to have strong theoretical links between the initial and intermediate outcomes and the longer-term outcome, which is the absence of whatever the program is trying to prevent. These theoretical links may already have been proven by impact research studies (e.g., we know that receipt of a polio vaccine prevents polio); they may emerge in exploratory research (girls who do not become pregnant as teens are involved in significantly more supervised activities after school than girls who do get pregnant); or they may be so intuitively strong that most would accept them on face value (youth who increase school attendance can be assumed to be less likely to drop out of school).



## Step 4: Prepare to Collect Data on Your Indicators



**B**y now, your work group has identified the outcomes you want to measure and has decided what specific information will indicate the program's level of achievement on each outcome. You also have decided what factors may influence outcomes, such as participant or program characteristics.

Now you are ready to complete the framework for your outcome measurement plan and develop the procedures that will carry the plan into action. The three tasks involved in this step are:

1. Identify data sources for your indicators.
2. Design data collection methods.
3. Pretest your data collection instruments and procedures.

### Task 1: Identify Data Sources for Your Indicators

The first task in this step is deciding where you will get the data for each indicator, and for the factors you think may influence participants' outcomes. The data sources most applicable to human service programs are:

- **Records**, from either your program or other agencies or offices (e.g., schools, hospitals, referral agencies, courts, employers, law enforcement agencies, health departments).
- **Specific individuals**, such as program participants; participants' parents, guardians, or other relatives; employers, teachers, counselors, or others who interact with program participants; service-providing staff and volunteers of your program; and staff of other programs, agencies, or organizations.
- **General public**, such as neighborhood or community residents, or population groups such as teens, parents, or homeless families.
- **Trained observers** who rate behavior, facilities, or environments.
- **Mechanical tests and measurements**, such as scales, rulers, blood analyzers, genetic tests, and strength tests.



Following is a more thorough discussion of each source.

**T**he **Detroit Institute for Children in Detroit, Michigan** was initially intimidated by the task of measuring outcomes of thousands of patients with very dissimilar needs. Because the Institute treats children with a wide range of chronic physical, developmental, and neurological disabilities, one child's goal may be to lift her head while another's may be to walk unassisted.

While it was relatively simple to measure outputs ("we serve more than 4,000 patients each year"), designing a unified system to measure extremely individualized client cases required a lot of creativity. An added complication was that data must be collected manually and over different time frames.

After considering a variety of strategies, the Institute developed a general desired outcome of making progress toward treatment plan goals. Indicators of progress for an individual child thus are based on that child's treatment plan. To measure the outcome, therapists evaluate each patient's chart and determine whether or not the child is making progress toward his or her goals. Data on this outcome show staff the proportion of its patients that are making progress.

- **Your program's records.** Before considering new sources of data for your indicators, review the information you currently collect. What data do you already have on individual participants that you can use to calculate the number and percent of participants who achieve various levels of improvement on program outcomes? Examples of outcome data that may be available from your own records are:
  - Number of job-training program participants who achieved satisfactory ratings on mock job interviews.

- Number of participants achieving the required score on a standardized test.
  - Number of former participants who re-entered the program after it appeared their service goals had been met (i.e., recidivists).
  - Qualitative data such as written summaries of participant progress.
- **Records of other programs or agencies.** Other agencies' records can provide information on participants' experiences outside your program that relate to your outcome indicators. Examples are:
- School grades and test scores of participants in an after-school tutoring program.
  - Number of youths who earned Red Cross swimming certification after taking swimming classes at summer camp.
  - Verification from the library that adult literacy program participants are using their library card.

Special considerations of confidentiality may be involved in obtaining information from other agencies. These are noted in the appendix.

- **Specific individuals.** Program participants, their parents and teachers, staff, and other key individuals can provide information on outcome indicators that is otherwise unavailable. They can give first-hand reports, for example, of:
- Participants' or former participants' demonstrations of knowledge, attitudes, or skills during and following the program.
  - Changes in participants' behavior.
  - Improvements in participants' status or condition.

Where individuals are a possible data source, consider how likely it is that their self-reports will be biased.

- For program participants, recognize that there may be some desire to provide the "right" answer—the answer that makes the participant look good or the program staff feel good. This is especially true if the participant is still in the program.
- For participants and their families, be especially concerned about unreliable self-report data if the information sought is of a very personal nature or

"We use a variety of methods to collect data for our school-based programs. We use report cards and the Piers-Harris Self-Esteem Inventory, which we administer pre- and post-match. We ask referring teachers to complete an objective evaluation form. A more subjective evaluation form is given to each parent, child, and volunteer to complete. And we use a case plan for each child which lists specific goals and objectives. With that, the case worker can check with the volunteer, teacher, and child on a monthly basis to determine progress made toward goal achievement."

*Susan Gelman,  
Executive Director,  
Big Brothers Big Sisters  
of Salem County Inc.,  
Salem, New Jersey*

is related to criminal behavior such as drug dealing or child abuse; or if participants may believe the information, such as undocumented worker status or unreported income, can jeopardize social welfare benefits.

- For staff of your and other programs, consider how much pressure there may be to provide information that makes the program, or the staff person, appear effective.

Despite potential problems, self-reports are a valuable way to collect many kinds of outcome data. You can take steps to lessen biased reporting. For example, you can stress to respondents the confidentiality of responses and the program's desire for honest answers. You can assure respondents that no one who works with participants will know what they said. You can reassure staff that the information will be used as a management tool to improve program effectiveness, not to grade individual performance.

If, despite these steps, you anticipate biased self-reports, look for either an alternative source or an additional source of data on the indicator, such as a parent, employer, or other agency's records.

---

**Note:** Focus groups or similar group meetings with clients or program personnel generally are not an appropriate source of outcome data. Focus groups typically involve only a small, non-representative segment of a program's participants and are free-flowing discussions. These sessions do not yield the kinds of individual-level information that can be tabulated and reported as outcome data. However, individual questionnaires completed during or after a group meeting may be a source of outcome data, provided that the meeting itself is not likely to have biased participants' responses.

As discussed in Step 2, focus groups are an excellent way to obtain input on what outcomes a program should measure. Focus groups also can help the program interpret outcome findings after data have been collected and tabulated.

---

- **General public.** Some programs are intended to benefit the general public or segments of it, rather than specific individuals. Examples are neighborhood improvement programs and public education and awareness campaigns. These programs may need information from representative, rather than specific, individuals on topics such as:
  - Residents' changes in behavior based on perception of increased neighborhood safety following initiation of a neighborhood watch program.
  - Neighbors' perception of changes in children's after-school play habits following renovation of local playgrounds.
  - Parents' actions to have children immunized following a public education campaign on the topic.

The general public in a neighborhood or community also is a potential data source for programs such as drop-in recreation centers or information and referral services that do not have names, addresses, or telephone numbers for their participants. For such programs, neighborhood residents or members of the at-large public can provide information on, for example, their knowledge and use of the program and the benefit of the service provided.

■ **Trained observers.** Trained observers can provide information on indicators related to behavior, facilities, environments, and other factors that (1) can be measured by physical observation and (2) can be rated on a scale that identifies variations in condition. Trained observers can yield data on, for instance:

- Use of direct eye contact during job interview role-plays by participants in an employment counseling program.
- Use of verbal rather than physical means to resolve conflicts by youth development program participants.
- Ability of rehabilitation program participants to undertake activities of daily living, such as preparing nutritious meals.
- The condition of a long-term care facility, including odors, temperature, visual privacy, condition of beds and other furniture, safety elements, lighting, presence of rats and roaches, and appearance and taste of food, after an advocacy group presses for better conditions for facility residents.
- The condition of parks and playgrounds, including safety conditions, amount of litter, and presence of broken glass, before and after a neighborhood improvement committee organizes a local clean-up campaign.

"Outcome measurement doesn't have to be a burden. It is empowering for staff and for program participants as each document their success. With a skill-based approach there is no failure. By definition we are not documenting lack of skills or incompetency: we're simply documenting how people have progressed."

*David L. Hoffman,  
President,  
Family Service of Milwaukee,  
Milwaukee, Wisconsin*

■ **Mechanical measurements.** Scales, yardsticks, and other devices can provide data, for example, on:

- Amount of weight gained by pregnant mothers.
- Growth rates of young children.
- Presence of controlled substances in the urine of drug abuse program participants.
- Increased lung capacity among former participants in a stop-smoking program.

Each data source has advantages and disadvantages as a source of information about outcomes. These are summarized in Exhibit 4-A.

## Advantages and Disadvantages of Data Sources for Outcome Indicators

Advantages	Disadvantages
<b>Your Program's Records</b>	
<ul style="list-style-type: none"> <li>■ Available</li> <li>■ Accessible</li> <li>■ You know how data were collected</li> <li>■ You can amend procedures to collect needed information on future participants</li> </ul>	<ul style="list-style-type: none"> <li>■ Value of data depends on how carefully it was recorded</li> <li>■ Existing records seldom contain all data needed</li> <li>■ Generally do not provide post-service information on participants' experiences or conditions</li> </ul>
<b>Other Agencies' Records</b>	
<ul style="list-style-type: none"> <li>■ Offers a perspective on participants' experiences different from yours</li> <li>■ May provide information on outcomes achieved after participant left your program</li> </ul>	<ul style="list-style-type: none"> <li>■ Value of data depends on how carefully it was recorded</li> <li>■ Existing records may not contain all data needed</li> <li>■ Confidentiality and other issues may make data unavailable to you</li> <li>■ Their time frames may not match yours</li> <li>■ Identification of participants may be a problem (e.g., is their "John Smith" your "John Smith"?)</li> </ul>
<b>Specific Individuals</b>	
<ul style="list-style-type: none"> <li>■ Can provide first-hand view of participants' experiences and outcomes during and after program involvement</li> </ul>	<ul style="list-style-type: none"> <li>■ Information can be biased by memory, interpretation, perceived pressure, fears</li> </ul>
<b>General Public</b>	
<ul style="list-style-type: none"> <li>■ Can provide information when specific individuals cannot be identified</li> <li>■ Can provide information on programs that serve geographic areas or population segments rather than specific individuals</li> </ul>	<ul style="list-style-type: none"> <li>■ Often, only a small proportion of the at-large group has experience with the program</li> </ul>
<b>Trained Observers</b>	
<ul style="list-style-type: none"> <li>■ Good for providing information on behavioral skills and practices</li> <li>■ An alternative to or supplement for participants' self-report information</li> </ul>	<ul style="list-style-type: none"> <li>■ Applies only to indicators based on physical observation</li> <li>■ Value of data depends on training and skill of observer, specificity of rating scale</li> <li>■ Problem of inconsistent ratings if more than one observer is involved</li> </ul>
<b>Mechanical Measurements</b>	
<ul style="list-style-type: none"> <li>■ Relatively objective, quantified, standardized</li> </ul>	<ul style="list-style-type: none"> <li>■ Findings are affected by accuracy of testing devices, training and skill of administrator</li> <li>■ Limited applicability to outcome measurement</li> </ul>

**C**lients of **Self Help for the Elderly**, a San Francisco-based agency, are predominantly Asian. Most do not speak English. The program thus deals with many language and cultural challenges in measuring outcomes of its services for seniors.

To begin, of course, data collection efforts with clients must be conducted in Chinese. The agency often uses focus groups with a moderator fluent in the language and familiar with the culture. This technique is effective in soliciting issues and concerns of elderly clients, including seniors who are simultaneously providing care for grandchildren and their own very elderly parents.

Elder abuse is one area of agency concern. However, the Chinese community's silence on this issue because of shame and cultural belief complicates monitoring efforts. To supplement first-hand reports, the agency monitors the number of cases reported by the police and doctors. Even these data are unreliable as indicators of incidence. To focus on a more measurable outcome, the agency has modified its goal for this program from decreasing incidence to encouraging more reporting through intervention and counseling.

Availability of comparative data poses another challenge. Comparisons to either American seniors of European ancestry or to Chinese elderly still living in China present complications related to lifestyle, diet, environment, and other issues. However, the Chinese-American population is relatively small and data on this group often are not available.

The center uses census information on this country's older population as a source of comparative data about health implications of its meals program, but recognizes the questionable comparability of the data. To develop baseline data on the incidence of diseases such as heart disease, diabetes, and cancer among Asians, the agency has incorporated health questions on its own intake form.

volunteer assistance from local universities and colleges, businesses, or retired professionals. If not, it is likely to be worth the expense to hire an appropriate consultant. A poorly designed data collection instrument or procedure can cancel out careful efforts to identify sound outcomes and indicators.

The choice of a data collection method may represent a trade-off between cost, response rate, time required to obtain the data, and other factors. Some characteristics of

Based on the information just presented, your work group can discuss sources of data for the outcome indicators and influencing factors you identified in Step 3. As you consider a potential data source, ask yourselves if it is reasonable to believe that the data source will provide useful, reliable information related to the outcome. Your work group can record its decisions about data sources on your program's Outcome Measurement Framework (Worksheets 4 and 5), which you began in Step 3.

## **Task 2: Design Data Collection Methods**

Once you have decided where you want to get the data for each of your indicators and influencing factors, you are ready to decide how you will obtain the data from those sources, find or create data collection instruments, and develop procedures for how the instruments will be used. Unless agencies have personnel experienced in designing instruments and procedures, they probably should seek professional help. Agencies may be able to obtain

four common data collection methods are summarized in Exhibit 4-B. As you consider these and other approaches to obtaining data, it is helpful to keep the following questions in mind:

- Is the data collection method feasible and not overly expensive? Is there a less time-intensive or less expensive way to collect this information?
- Will the resulting data be useful to program managers for program improvement?
- Will the resulting data be credible to those outside the program who are likely to look at the information?

Do not drop an indicator too quickly because it appears hard to measure. Examine the data source and data collection options fully. A little creativity may present a solution. If it does not, then seek another indicator for the outcome.

"Any university is equipped to do research and evaluation and it's an easy, cheap place to go. You can always hire a management consultant, but universities are cheaper. To find the right department, look into Policy Making, Public Administration or some department where they are accustomed to doing program evaluation work. Most universities have statistics courses and you can trace those to an appropriate department in the university."

*Jim Clark,  
Vice President,  
Daniel Memorial,  
Jacksonville, Florida*

Exhibit 4-B

### Comparison of Major Data Collection Methods

Characteristic	Data Collection Method			Rating by Trained Observer
	Review of Program Records	Self-administered Questionnaire	Interview	
Cost	Low	Moderate	Moderate to high, depending on how administered	Depends on availability of low-cost observers
Amount of training required for data collectors	Some	None to some, depending on how distributed	Moderate to high, depending on collectors' previous experience and survey complexity	Moderate to high, depending on complexity, subtlety of observations
Completion time	Short, depending on amount of data needed	Moderate to long, depending on how distributed	Long	Short to moderate
Response rate	High, if records contain needed data	Depends on how distributed	Generally moderate to good	High

A valuable source of feedback at this point is someone with experience in the various data collection methods you have identified. This may be a member of your staff, one of your volunteers, faculty from a local college or university, someone in a national organization with which you are affiliated, or someone else known to your organization. Their input on potential feasibility and cost can be quite useful.

Your work group's choices of data collection methods can be recorded in the fourth column of your Outcome Measurement Framework (Worksheets 4 and 5). The completed frameworks, including data sources and collection methods, for the four programs of the Southside Children's Agency are provided in Exhibit 4-C as examples.

## Preparing Data Collection Instruments

Developing sound questionnaires, interview guides, tests, observer rating scales, and other data collection instruments can be a time-consuming task. Before starting to construct your own materials from scratch, locate instruments that others have developed and tested to see what you can use or adapt. Many pre-tested instruments provide scales for measuring a variety of human conditions and attitudes, such as self-esteem, family health, parenting skills, and mental outlook. Review of these might save your program considerable development time.



References listed in the bibliography may help you locate applicable questionnaires, rating scales, observation guides, and other instruments. Other potential sources include national professional organizations, accrediting bodies, college

and university faculties and libraries, Internet searches, local foundations, your local United Way, and other service agencies in your community.

While you are researching instruments, review information on data collection procedures others have used. Their approaches to timing of collection, sample selection, training of collectors, challenges they have faced, and other topics may help guide some of the decisions you will be making.

## Sample Outcome Measurement Framework: Southside Children's Agency

Program: Teen Mother Parenting Education

Outcome	Indicator(s)	Data Source	Data Collection Method
Teens are knowledgeable of prenatal nutrition and health guidelines.	Number and percent of program participants able to identify food items that are good sources of major dietary requirements	Participants	Self-administered survey after second week in the program
Teens follow proper nutrition and health guidelines.	Number and percent of participants within proper ranges for prenatal weight gain	School scales	Weekly weigh-ins
	Number and percent of participants that do not smoke	Participants	Self-report on daily checklist
		Teachers	Observation reported on weekly record
	Number and percent of participants that eat at least 4 calcium servings and one of each other nutritional group each day	Participants	"Healthy Baby" checklist for recording daily food intake
	Number and percent of participants that take a prenatal vitamin each day	Participants	"Healthy Baby" checklist for recording daily food intake
Teens deliver healthy babies.	Number and percent of newborns weighing 5.5 pounds or above and scoring 7 or above on Apgar scale	Hospital records	Contact hospital for birth records

Influencing Factor	Data Source	Data Collection Method
Mother's age	Program records	Record review
Mother's household income	Program records (based on mother self-report)	Record review
Number of parenting classes mother attended	Program records	Record review

Exhibit 4-C, cont'd

Program: Developmental Child Care

Outcome	Indicator(s)	Data Source	Data Collection Method
Teen mothers attend school regularly.	Of teen mothers enrolled in high school whose children are enrolled in child care program, number and percent who miss less than 3 days per month during school year	Child care records, school records	Monthly attendance report from school MIS
Teen mothers graduate from high school.	Of teen mothers whose children are enrolled in child care program, number and percent who graduate from high school over a four-year period	Child care records, school record	Track all teen mothers of children ever enrolled in child care program until they drop out or graduate, using school MIS

Influencing Factor	Data Source	Data Collection Method
Mother's age	Program records	Record review
Mother's grade point average	Program records	Record review
Mother's living situation (living alone, with parents or guardian, with partner)	Program records	Record review
Mother's participation in school-run mentoring and tutoring services	Program records	Agreement with school to identify program participants who are receiving these school-run services

## Exhibit 4-C, cont'd

Program: Child Abuse Public Awareness and Prevention

Outcome	Indicator(s)	Data Source	Data Collection Method
General public, educators, child care workers, and medical professionals hear or read information.	Number and percent of target audiences that recall content of PSAs, brochures, posters, or presentations	General public	Community telephone survey
		Child care workers, medical professionals, educators	Telephone survey to 20% of those who received materials
		Presentation attendees	Follow-up mail survey
Target audiences are knowledgeable of signs of child abuse and neglect and of appropriate actions to take.	Number and percent of target audiences that can name 3 signs of abuse and neglect and can identify police or child protective services as appropriate contact	Same as above	Same as above
Target audiences identify and report suspected abuse and neglect.	Number and percent of target audiences that say they have made such a report since hearing or seeing information	Same as above	Same as above

Influencing Factor	Data Source	Data Collection Method
Target audience (general public, medical professionals, child care workers, educators)	Self-report	Survey used for outcome data
Type of information received (PSAs, brochures, posters, presentations)	Self-report	Survey used for outcome data
Number of times information received (1 time, 2 to 4 times, 5 or more times)	Self-report	Survey used for outcome data

Exhibit 4-C, cont'd

Program: Neighborhood Organizing

Outcome	Indicator(s)	Data Source	Data Collection Method
Residents sign up to clean up vacant lots and build playgrounds.	Number of residents on sign-up sheets	Sign-up sheets	Distribute forms to all meeting attendees asking them to indicate interest in volunteering
Volunteers for each vacant lot meet and develop action plans.	Number of action plans developed	Staff report	Staff and board member teams organize and observe planning meetings and document action plans
Volunteers create clean, drug-free play areas.	Number of vacant lots that are free of litter, have grass or other appropriate ground cover, have play equipment, and are free of drug sales and/or use	Staff/volunteer observation	Apply observation checklist
Children play in clean, safe areas.	Number of children observed in play areas on one random, sunny day per week for 3 months after clean-up	Staff/volunteer observation	Apply observation checklist
Residents feel neighborhood is safer place for children.	Percent of neighborhood residents that say neighborhood is safer for children today than it was one year ago	Neighborhood residents	Mail surveys to every household in the neighborhood

Influencing Factor	Data Source	Data Collection Method
Number of volunteers for each vacant lot	Staff	Observation
Race and sex of volunteers, children, and community residents	Staff (for volunteers and children)	Observation
	Self-report (for community residents)	Mail survey (same survey used for outcome data)

If appropriate data collection instruments are unavailable, you will need to develop them. The appendix offers specific suggestions for developing record extraction forms, questionnaires, and trained observer rating scales. Exhibit 4-D provides guidance in considering cultural issues in instrument development.

In developing data collection instruments, it is tempting to include a lot of questions, categories, and topics because the information seems "interesting" or would be "nice to know." It also is easy to overlook information essential to measuring the outcomes you have selected.

#### Exhibit 4-D

### Considering Cultural Issues in Instrument Development

- In developing questionnaires that will be administered to persons from different cultures, consider carefully the meaning of the words you use. Consider also that "culture" is not "race." Even if all your clients belong to the same racial group, the concept of "physical discipline," for example, can mean many different things.
- Consider the possibility of cultural response sets. Most people, parents especially, want to look good. Parents of northern European extraction, for example, are likely to rate highly their child's sense of independence and initiative, while parents from more collectivist cultures may want to emphasize how well their child adheres to group norms and practices.
- Be aware that many standardized instruments were developed for white middle-class populations. A youth worker in New Mexico, for example, will tell you that an instrument designed to measure the functioning, well-being, and self-esteem of white middle-class adolescents does not address the issues of cultural identity, biculturalism, and cultural assimilation faced by Native American and Hispanic youth. Check with the author of any instrument you plan to use to determine whether it has been used or adapted for the population your program serves. Better yet, search for instruments developed specifically for your population.
- If you plan to measure changes between pre- and post-tests, then consider the content of the pretest carefully. For some cultural groups, it may be appropriate to ask questions at follow-up after trust has been established between the agency and the client, but it may not be appropriate to do so at the time of intake. Some cultural groups expect and accept in-depth diagnostic interviews at time of intake. Other groups consider this to be prying. For these groups, an intrusive pretest questionnaire is inappropriate.
- If in doubt about any aspect of your questionnaire, ask a representative of the cultural community you are surveying to review not only the instrument, but also the data collection procedures and the interpretation of results.

As you finish drafting an instrument, two checks will help ensure that you are collecting all the information you need. First, compare the data elements the instrument will collect against your outcome measurement framework. Does the instrument in fact request all the information assigned to it in the framework? Second, sketch out the tables, charts, and types of findings you want to be able to present with the data you collect. Can you identify specifically which question in which instrument will provide each data item? If the answer to either question is “no,” then additional information needs to be collected.

As a check against collecting unnecessary data, look at the questions in each instrument that are unrelated to the outcomes you want to measure. Do you have a plan for using the information from each question for other purposes? If not, do you really need to collect the data? If not, don't.

Be sure the individual responsible for coding and tabulating your data has ample opportunity to review and make suggestions on your draft instruments. Many design features can facilitate these tasks, saving time, money, and frustration.

“I've been in the business a long time and we used to have records two feet thick that didn't say anything. Now those days are gone, hopefully.”

*John Steele,  
Executive Director,  
Catholic Family and  
Community Services,  
Schenectady, New York*

## Developing Data Collection Procedures

Along with any data collection instruments—record review form, survey and interview questionnaire, observation guide, and so forth—you need procedures for using the instrument to obtain the information you need. Some procedural issues, such as timing of data collection, who is considered a participant for data collection purposes, sampling, who collects the data, and confidentiality protection, are common to all data collection efforts. Guidance in reaching decisions on these issues is provided below. The appendix discusses procedural issues specific to various types of instruments.

Although this manual touches on many procedural issues, it provides only an overview of these topics. It does not address technical details of issues such as sample size statistics and selection methods; frequency of administration; administration procedures; obtaining accurate, reliable, data; and obtaining adequate response rates. The bibliography lists sources for more detailed information on these topics.

## At What Milestones Should Data Be Collected?

Is there information that should be obtained when a participant enters the program? Is information needed at various points while a participant is receiving service? If so, what are those points: weekly? three months after entry? one year? What information is needed after service completion and at what intervals after completion: one month? six months? one year?

Timing depends in part on what the outcome indicator says. For example, if the indicator calls for the program to track the number of participants who *increase* their knowledge of a certain topic or who show *more* ability to deal with a certain situation, to be able to measure change the program may need to collect data both before and after the service is received.

Timing also depends on the characteristics of the program's services. For example, a few services are one-time only or very short term and are not necessarily intended to have significant longer-term effects on participants. Examples of these services include emergency food and shelter programs and some recreational services. For such immediate outcomes, collection of outcome information may take place at the time that the service is delivered. Participants can report adequately on the helpfulness of the service at that time.

Most human service programs intend their services to have longer-term effects that cannot be determined immediately after receipt of service. In these cases, outcome data need to be collected at one or more follow-up points. These points in time should be tied to some recognizable, administratively determined event related to the individual participant's affiliation with the program. The two most common events are the *date of entry* into the program and the *date of completion* of the service. This means that follow-up data collection will occur on different dates for participants who started and/or completed the program at different times.

Collecting follow-up data for all participants at once can result in a hodgepodge of time intervals since the date participants entered or completed service, with outcomes for some participants measured one week after they started with the program, for others at one month after completion, for still others at five months, and so on. When data for all these participants are combined, the varying time intervals may obscure findings that data collected on an individualized schedule would reveal.

As a rule of thumb, programs should collect follow-up outcome data at approximately six to twelve months after a participant either enters or completes service. In principle, the best outcomes to track are the longer-term ones. The dilemma is that tracking these effects generally is difficult. First, it usually is more costly and frequently is impossible to

locate participants who left a program many years ago. Second, many other influences will have occurred in participants' lives in the interim, making it difficult for the participants to respond usefully to questions. Third, the information obtained may pertain to past program procedures, policies, and personnel which may no longer be relevant. Examination of outcomes that follow a service by two or more years is more likely to be appropriate by undertaking special, ad hoc studies and evaluations, rather than by regular outcome measurement.

### **Who Is Considered a Participant?**

Should your data collection include individuals who started the program but left before completing the service? What about those who sought service but were ineligible or otherwise were refused service? For guidance, review your outcome indicators. Do they call for you to compare outcomes of persons who left the program before completing the service with outcomes of persons who completed it? Did you plan to compare those who completed the program with those who were refused service for some reason? If your indicators do not call for this type of comparison, should they? If this information would be valuable to the program, amend your indicators.

### **Include All Participants or Only a Sample?**

If outcome measurement is to be used as an ongoing management tool, the preferred option is to measure outcomes for all program participants. Programs often can build data collection into the program's normal operations without adding significantly to the burden on staff. If it is not feasible to collect outcome data on all participants, such as in a public awareness program that serves the entire community, an educational prevention program that holds workplace seminars for tens of thousands of participants, or a youth development program that covers a large population area, it is necessary to collect data on only a portion, or sample, of individuals served.

Community-wide surveys may be too costly for a large-enough sample to detect the outcomes of public education programs. If possible, target the population of interest more precisely and then survey the target population.

Guidance on sample sizes and selection methods is beyond the scope of this manual. References in the bibliography provide information on these and related issues. If the number of participants in your program is so large that sampling is essential, you may be well served by obtaining professional advice for your first round of data collection.

## Who Will Collect the Data?

Regardless of the data collection method used, the value of the data ultimately depends on the care and skills of the data collector. Data collectors must follow procedures carefully and record information accurately. They may be members of your staff or volunteer structure, staff of other agencies, program or research experts recruited from professional organizations or universities, students, or community members.

To assure that procedures are followed uniformly, data collectors must be well trained in the particular data collection assignment they will undertake for you. To minimize bias of information, program staff members and volunteers should not collect data from or about participants with whom they have worked.

Select one or two individuals to train the data collectors. Training should include explanations of the purpose of the work and how it will be used. The issue of confidentiality, discussed below, should be stressed. Data collectors should practice their task by conducting mock interviews, reviewing sample records, or performing sample observations.

Depending on your program and common characteristics of participants, the trainer may need to prepare data collectors for special situations. Interviewers, for example, may need a list of agencies to which they can make referrals related to child rearing, employment training, domestic violence, or other problems that are revealed during an interview. They also may need information and instruction on your state's requirements regarding reporting of suspected child abuse.

## How Will Confidentiality Be Protected?

In all data collection efforts, you need procedures that protect the confidentiality of participant information. Protecting confidentiality does not mean simply avoiding use of an individual's name. It means that data collectors and others involved in outcome measurement refrain from discussing situations, locations, affiliations, and all other information about participants with anyone other than the data collection monitor or other authorized persons. Some ways to protect confidentiality are:

- As mentioned above, stress the importance of this issue in training data collectors. Discuss the ethical implications and the legal consequences for both the program and the data collector. Especially if data collectors are not staff members, you may want to require that they sign a pledge of confidentiality such as the sample in Exhibit 4-E before they have contact with participants or records.

## Sample Pledge of Confidentiality for Data Collectors

**I understand that:**

- I may be collecting information of a personal and sensitive nature.
- Individuals participating in this study have been assured that their names will not be disclosed and that all information will be kept confidential.
- The responsibility of fulfilling this assurance of confidentiality begins with me.

**In recognition of this responsibility, I hereby give my personal pledge to:**

1. Keep confidential the names of all respondents, all information and opinions collected during the data collection process, and any information learned incidentally while collecting the data.
2. Refrain from discussing or disclosing, except privately with my data collection supervisor, information that might in any way identify or be linked to a particular individual.
3. Terminate data collection immediately if I encounter a respondent or begin reviewing a record for an individual whom I know personally, and contact my supervisor for further instructions.
4. Take precautions to prevent access by others to data in my possession.
5. Take all other actions within my power to safeguard the privacy of respondents and protect the confidentiality of information I collect.
6. Devote my best efforts to ensure that there is compliance with the required procedures by persons whom I supervise.

Signed: \_\_\_\_\_

Date: \_\_\_\_\_

- Include a cover sheet on each data collection form that gives the participant's name and an identification number. On all other pages of the form, use only the identification number. As soon as the data collector submits the completed form, detach the cover sheet from the rest of the form. File cover sheets and completed forms in separate files.
- Keep cover sheets and completed forms in locked files.
- Take steps to ensure that persons who are not working with the data do not have access to it.

"The Community Service Center is working with adult ex-offenders. Our goal is to help individuals attain self-sufficiency, thereby reducing recidivism. In order to determine effectiveness we implemented several methods which would assist with client tracking. CSC utilizes both telephone calls and post card mailings to follow up or track clients' progress. The post cards give the clients a feeling of a personal touch as well as enable staff to obtain information from family or friends because it serves as an introduction tool. The agency is experiencing a 70 to 80 percent response to the combined telephone call/post card effort."

*Octavia Edinburg,  
Executive Director, Community  
Service Center, Inc.,  
New Orleans, Louisiana*

### **How Will Participants Be Informed About the Data Collection Process?**

Participant rights need to be protected. Such rights include the right to informed consent or refusal, and the right to privacy and confidentiality. Often information about your agency's outcome measurement process can be incorporated into the registration or intake process for the program. For most adult participants, written consent is not needed. If in doubt, check with your agency's legal counsel. However, if participants are children or youth, parental consent generally will be required if other program data will be extracted, or if follow-up surveys will be used. Exhibit 4-F is a sample consent form.

## Sample Consent Form

To learn more about how well the Teen Center's *Teen Mothers Program* is working, we routinely interview participants after they leave the program to ask how they are doing.

If you agree to a telephone interview, everything you say will be confidential. The information you provide will be combined with the information from all other teen mothers we interview. No one will be able to tell which answers are yours.

We also would like to obtain information about how well participants are doing in school. If you give us permission, we will contact your school counselor and ask for a copy of your attendance records and grades. This information also will be confidential.

Participation in either of the follow-up studies is completely voluntary. Whether you participate or not will not affect your eligibility for services from any of the Teen Center's programs.

If you agree to participate in these follow-up studies, and we hope you will, please read (or have read to you) both agreements below and then sign them.

*I have read this form (or this form has been read to me), and I agree to participate in the Teen Mother's Program follow-up telephone survey. I understand that my participation is totally voluntary, that I can refuse to answer any question that is asked, and that I can stop the interview at any time.*

\_\_\_\_\_  
Participant's Signature

\_\_\_\_\_  
Parent/Guardian's Signature

\_\_\_\_\_  
Printed Name

\_\_\_\_\_  
Printed Name

*I have read this form (or this form has been read to me), and I agree to have my school counselor provide the Teen Mother's Program with a copy of my grades and attendance records. I understand that my agreeing to this arrangement is totally voluntary, and that I can stop the arrangement at any time by contacting either the Teen Mother's Program or my school counselor.*

\_\_\_\_\_  
Participant's Signature

\_\_\_\_\_  
Parent/Guardian's Signature

\_\_\_\_\_  
Printed Name

\_\_\_\_\_  
Printed Name

\_\_\_\_\_  
Date

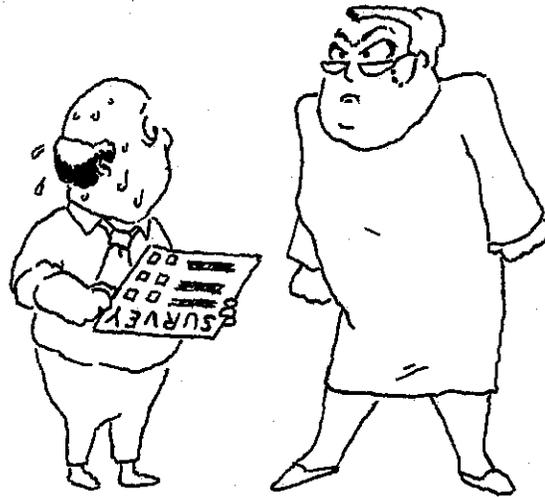
\_\_\_\_\_  
Date

Thank you very much!

### Task 3: Pretest Your Data Collection Instruments and Procedures

No matter how carefully data collection instruments and procedures are developed, some problems will be discovered only when the instruments and procedures are used. The purpose of a pretest is to try them out on a very limited basis and uncover problems while things are still in a draft stage. A pretest is well worth the relatively modest effort involved.

Generally, it is quite acceptable to pretest each instrument and its related procedures as they become ready, rather than waiting to pretest everything at the same time. In fact, you are likely to learn some things in your first pretest that can help in developing other materials. If your data collection procedures involve sampling, include the sampling methodology in your pretest.



**For record extraction,** collect data from five to 10 records at each program involved. If the extraction is being done manually, involve at least two or three reviewers. If the records are complex, or if reviewers need to make a lot of judgements about what the information means, as with qualitative information, have multiple reviewers extract data from the same set of cases to see if their interpretations are consistent.

**For questionnaires,** pretest each questionnaire with six to 10 persons representative of the group from whom you'll be collecting data. For example, if you will use one interview form with older persons, pretest it with older persons. For questionnaires to be administered through interviews, involve at least two or three interviewers. Try out not only the interview itself, but also the contact and consent procedures.

**For trained observer ratings,** ask three to five raters to observe and rate the same five or so behaviors, environments, or facilities. Compare ratings to see if observers apply the criteria consistently.

For all pretests, inform the participating individuals—respondents as well as data collectors—in advance that this is a pretest. After data have been collected, request their feedback on all aspects of the data collection, such as:

- **The wording of the questions.** Did some questions seem confusing or misunderstood? Did any seem biased? Was the vocabulary appropriate for the target group? It often is surprising which apparently simple terms do not communicate as expected. In one outcome measurement project, for instance, elderly participants did not understand the meaning of the word “coping” in a question that asked whether they had difficulty coping with a problem.

- **The content of the questions.** Did certain questions make respondents feel uncomfortable? Did any questions seem offensive or useless? Did respondents refuse to answer any questions? If so, why? You may be able to reword troublesome questions or add an explanation. Questions relating to alcohol and other drug use may fall into this category.

**A** part-time program evaluator proved essential to the outcome measurement process for **The Deaf, Hearing and Speech Center in Detroit, Michigan**. Initially, she was to provide staff orientation, participate in introductory meetings with the executive and assistant directors, help improve the tools and forms that were already in place, and advise the staff on analyzing and reporting the data. Her assignment expanded to include working with a computer programmer to design a data input and analysis system.

Her biggest contribution was improving the agency’s measurement tools by suggesting better wording and response categories for questions and recommending changes to make standardized tests more appropriate for the agency’s use. A re-focused questionnaire for hearing-impaired support group participants revealed that more success was noted in learning to understand and accept their hearing loss than in learning technical skills. This explained to staff why participants were so loyal to the program despite a somewhat slower than expected mastery of lip reading.

The consultant and staff met monthly while they worked on questionnaire design and implementation. Her involvement scaled back to quarterly meetings while she helped the agency analyze and report the data.

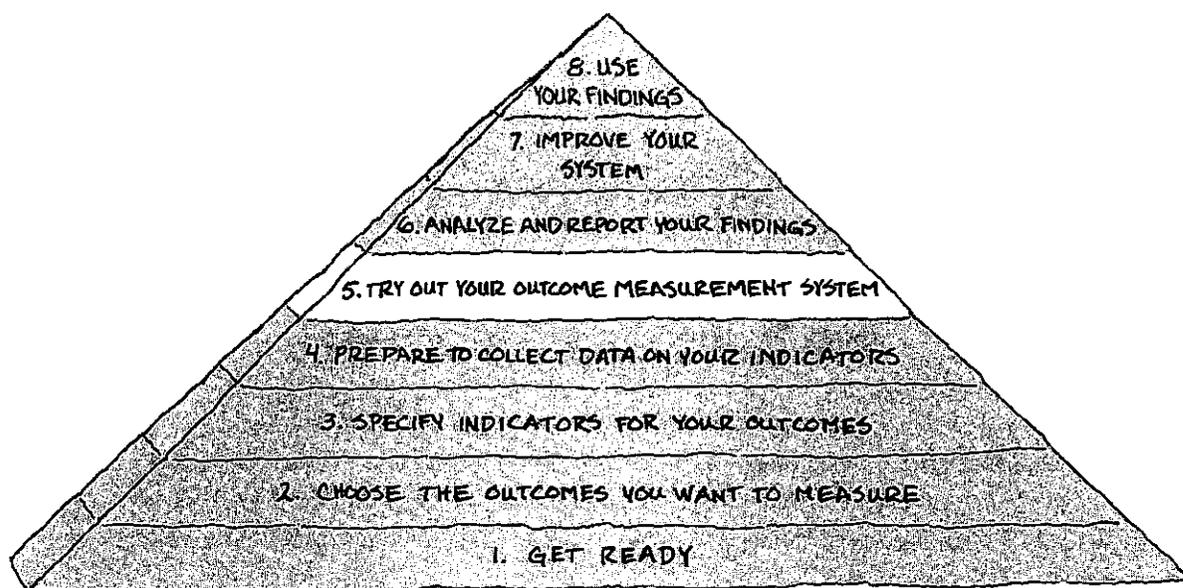
- **The adequacy of the response categories.** Were there enough categories? Too many? Did they offer the right choices? Some respondents may not like being forced into Yes or No answers. There may be a need for categories such as “sometimes” or “don’t know.”
- **The clarity of the instructions.** Where were data collectors or respondents unsure about what to do or how to provide information?
- **The layout and format of the instrument.** Did it look inviting or intimidating? Was it easy to read, easy to use?
- **The length of time required to complete the data collection.** Did it take too long to provide the information? Were respondents losing interest? In addition to learning if you need to adjust the number of questions, you can use this information to tell future data collectors and respondents about how long the data collection will take.

Look closely at completed instruments. Are people following the instructions? Maybe the instructions are not clear. Has some information been omitted? Maybe the format makes it easy to miss a question. Did people have enough space to write? Maybe blanks should be longer to provide more space. Are "catch-all" categories such as "other" used excessively? Maybe the other response categories are inadequate.

Someone who will be involved in data entry and analysis should use the completed forms to pretest this part of the process. How easy will it be to code the information and enter it into computer files or onto tally sheets?

Following the pretest, revise the instruments and the data collection procedures as appropriate. If the revisions to an instrument are substantial, retest it to be sure the problems have been addressed without adding new ones. If procedures have been changed significantly, retest them as well to be sure the revised approach is effective.

## Step 5: Try Out Your Outcome Measurement System



**T**he time has come. You are ready to try out your system in a “pilot test.” In this trial run, you will implement all parts of your system just as you plan to do in full-scale implementation, although you may not apply the system to all parts of the program. The system you are piloting includes not only collecting outcome data, which is discussed in this step, but also analyzing and reporting it, as discussed in Steps 6 and 7.

During the trial, you are likely to identify issues such as overlooked outcomes, inadequately defined indicators, cumbersome procedures, and analysis and reporting dilemmas. The tryout gives you a chance to address these problems and make the process smoother and more comprehensive while you still are at the trial stage. **Please do not skip the trial step.** It will more than repay you in time and aggravation saved later.

“In our haste to collect the information, we were at the statistical analysis stage when we realized we had forgotten very important questions about birth outcomes. We had to go back and collect low-birth-weight and prematurity information.”

*Peggy Ann Dorrian,  
Talking, Listening, Caring  
Program Supervisor,  
Catholic Family and  
Community Services,  
Schenectady, New York*

Tasks in this step are:

1. Develop a trial strategy.
2. Prepare the data collectors.
3. Track and collect outcome data.
4. Monitor the outcome measurement process.



## Task 1: Develop a Trial Strategy

If your program has few participants at any one time, you may decide to implement your trial run using all participants in your program. If your program is larger, however, you may want to keep the process more manageable by limiting the trial to a subset of participants. There are various ways to accomplish this, depending on how your program is organized. For example:

- If your program operates at more than one site, you could try the system out at only some sites.
- If program staff are organized into units, choose only some units for the trial.
- If participants start and end their experience with your program in small groups, such as parenting classes, measure outcomes of participants in only some of the groups.

Your trial run must reflect the system as it eventually will be used. If you decide to measure outcomes for only a subset of participants, for example, pick sites or staff units in a way that assures a representative group of participants. If you plan to use student interns as data collectors once full-scale implementation begins, use student interns as data collectors for the trial. If you eventually will hand tally your data, then hand tally your trial-run data as well.

Examine your schedule for the trial run in the planning and implementation timeline you developed in Step 1. Now that you know what your data collection procedures will be, you may need to adjust the schedule—perhaps significantly. Be sure you have allowed enough time to prepare for the trial—train data collectors and set up tracking and monitoring procedures—as well as to conduct it.

## Task 2: Prepare the Data Collectors

During Step 4, you developed plans for recruiting and training data collectors. Now is the time to try out that part of your outcome measurement system. Recruit appropriate individuals to collect the various types of data you need and train them on the use of your instruments. Include opportunities to practice administering the instruments in mock interviews, record extractions, or observations. Provide, as necessary, information on contact and consent procedures. Be sure to stress the importance of safeguarding respondent confidentiality.

## Task 3: Track and Collect Outcome Data

When your data collectors have been trained, you are ready to begin collecting data. Activities that will help you implement this task effectively include the following.

- Assign an individual to monitor and track the data collection process.
- If necessary, develop a master list of participants identifying what data elements need to be collected for each individual. You will not always need this master list. But if your participants receive a number of services, it may be helpful to have a central list of what data needs to be collected for which participants. Exhibit 5-A is an example for a transitional housing program whose participants are former residents of a homeless women's shelter. The log shows what data collection is to be initiated for each woman, depending on her status in three areas.

Exhibit 5-A

### Sample Master List of Data Elements Needed for Each Participant

ID #	Has children (Observe parent-child interaction.)	Former drug abuser (Assess drug use via urine test.)	Enrolled in GED program (Obtain data from program.)
001		X	X
002		X	X
004	X		X
005	X	X	
...			

- **Develop a log for monitoring the flow of data from its source to the data collection monitor.** Generally, the information should be tracked by participant. Log elements could include: name of data collector, whether consent was obtained, number of efforts made to reach the participant, and date of data collection. It also is helpful to keep track of intermediary data handlers. If the data pass from one source (say, a teacher) to another (the school principal) to the agency, the log will help identify bottlenecks.

Exhibit 5-B is a sample data collection log for an adolescent mentoring program that collects outcome

data on former participants from two sources: telephone interviews and school records. While still participating in the program, young women sign consent forms agreeing to respond to the follow-up telephone interview and to allow the agency to collect information about school status and school and academic performance from the school counselor.

"Based on our pilot test, we recognized that we weren't including a measure on our Knowledge and Awareness Seminar. We added an additional outcome and included a measure in a pre/post-test."

*Royce Van Evera,  
Executive Director,  
Law, Order & Justice Center,  
Schenectady, New York*

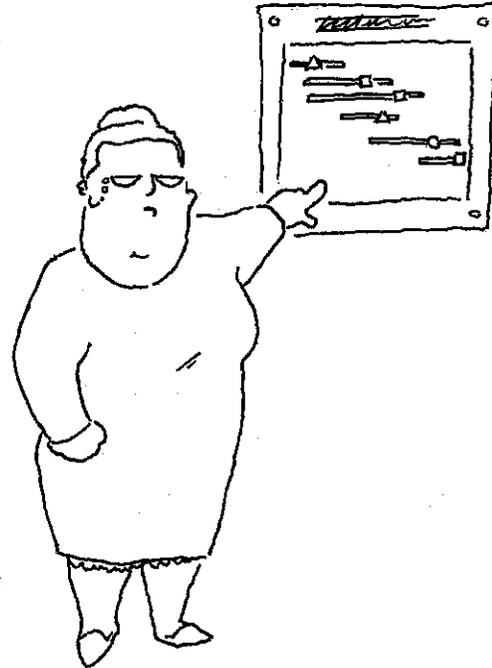
- **Develop procedures for monitoring the collection of data.** For data from participant surveys, the agency will need to track attempts to reach respondents, such as mailings, follow-up mailings, and telephone calls, and whether questionnaires have been received. For data from case records, the activity of collecting the data needs to be monitored.
- **Develop procedures to safeguard the confidentiality of the files.** Confidentiality procedures must cover not only the data collected from the participants, but also whether the participant agreed to provide outcome information. As data forms are received, confirm that the correct identification number has been recorded for each participant. Then detach the cover sheet with the participant's name and file it separately from the form. Keep data logs, data forms, cover sheets, and the master list that links names to identification numbers in locked files.
- **Develop a quality monitoring plan for reviewing the quality of the data received.** The data coordinator should spot-check incoming forms to be sure they are being completed correctly. Instruct data processors to return forms that are incomplete or illegible to the data monitor for resolution. You also may want to call a sample of persons interviewed to verify that the interviews were really conducted, were respectfully conducted, and seemed relevant.



## Task 4: Monitor the Outcome Measurement Process

During the trial run, you will want to monitor the outcome measurement system itself. Your purpose is not only to identify problems, but also to gain a clear picture of what the system is requiring in terms of time, money, and other resources. Your work group should consult with program and fiscal managers to identify issues to track. Examples include:

- How much time data collectors spend on each interview, record, or observation.
- Response rates on first, second, and third mailings for mail surveys.
- Former participants unable to be located.
- Refusal rates for in-person or telephone interviews.
- Data frequently missing in program records.
- Planned observations that could not be completed and why.
- Data collection errors (missed questions, coding discrepancies, etc.).
- Data needed for analysis but unavailable.
- Printing, postage, and other costs beyond staff time.



The work group also should devise methods for capturing information on system issues. For example:

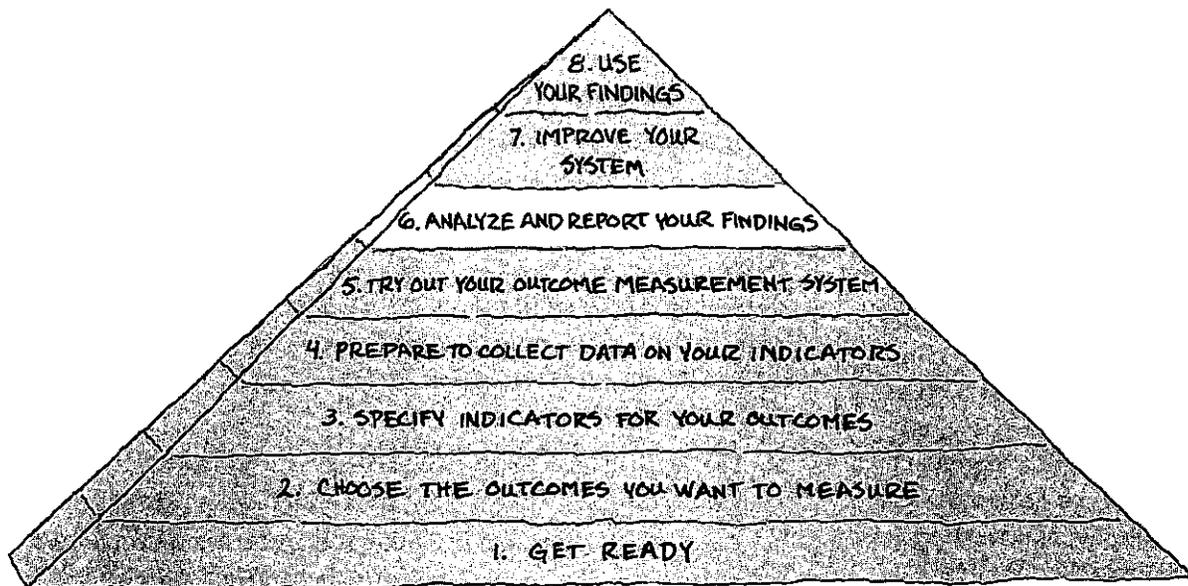
- Interview forms can include places for interviewers to record time started and time stopped for each interview.
- Data collectors can answer a short set of questions regarding every (or every fifth or every tenth) survey, record review, or observation.
- A form asking how long the survey took to complete can accompany a sample of mail surveys.
- The work group can meet semi-weekly with data collectors to identify problems and suggestions.

A common dilemma during a trial run is whether to correct a faulty data collection instrument or procedure in the middle of the test. In general, any mid-stream change weakens the strength of your findings. It may confuse the data collectors, making it hard to tell if the adjustment had the desired results. Further, if the change is significant, you will need to analyze pre-adjustment and post-adjustment data separately, running the risk that your numbers in both groups will be too small to draw any valid conclusions. These

hazards argue against mid-trial changes. On the other hand, if the problem is so substantial that you are likely to get little data, or if your data will be meaningless, then a mid-course correction probably is in order.



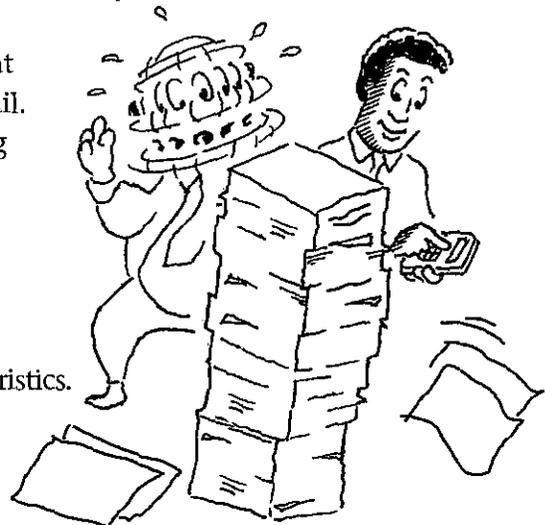
## Step 6: Analyze and Report Your Findings



**F**ollowing your trial data collection effort, you will be ready to try out the next part of your outcome measurement system: data analysis and reporting. In this process you will learn, for example, whether you collected all the data you need to measure the outcomes you selected, whether your data collection instruments make it easy to tabulate your findings, and whether you collected data for which you don't have a use.

This manual's bibliography lists resources that address data analysis and reporting issues in detail. Step 6 highlights key issues involved in analyzing your data and reporting it clearly and usefully. Tasks discussed in Step 6 are:

1. Enter the data and check for errors.
2. Tabulate the data.
3. Analyze the data broken out by key characteristics.
4. Provide explanatory information related to your findings.
5. Present your data in clear and understandable form.



Begin Step 6 by assigning someone the task of monitoring data processing activities. Depending on the size of your agency and the size of the data collection effort, this person may or may not be the same person responsible for monitoring the data collection process. The work group should coordinate with this individual to monitor the time invested and problems encountered in each data coding, analysis, and reporting task.

## Task 1: Enter the Data and Check for Errors

"Although we serve nearly 4,000 boys and girls in our school-based and club programs, we only have one full-time professional staff person and no computer. By keeping data collection simple and focusing on what's really important to know, we have a do-able outcome measurement system that has been effective and helpful."

S. "Boots" Provensal,  
Executive Director,  
Greater New  
Orleans Council,  
Camp Fire Boys and Girls,  
New Orleans, Louisiana

Processing the data means transferring the information recorded on questionnaires, observer rating forms, case record abstraction sheets, and other documents to either a computer or a new form that helps you summarize the data. Data processing does not necessarily need to involve computers, although the use of computers usually will make data analysis easier.

- **Computerized data processing.** In computerized data processing, the main activity is data entry: transferring data codes from the data collection forms directly to a computer file. A number of easy-to-use database or spreadsheet software programs are available for data entry and subsequent data analysis. You typically will not need to use more sophisticated statistical packages.
- **Manual data processing.** If you will be tabulating data by hand, transfer the data onto a simple spreadsheet such as that shown in Exhibit 6-A.

Whatever means you use for data entry, it is important to check for errors. Here are several approaches you can use.

- Transfer or enter all data twice and compare the two data files. If this appears too cumbersome, start with about 10 percent of the data. Check any discrepancies between the two data files. If errors are pervasive, recheck *all* data.
- Examine a listing of the data on a simple spreadsheet to look for obvious mistakes. For instance, if the answer to a question is either "1" for yes or "2" for no, the appearance of another number indicates an error. This simple "eye-balling method" works well for data entered electronically as well as for manually produced spreadsheets.

Exhibit 6-A

## Sample Spreadsheet for Programs Without Computers

Progress of Individuals Contacting Employment Training Program													
ID #	Started program		Completed program			Found employment within 3 months after completing program				Still employed 6 months after starting job			
	Yes	No	Yes	No	NA	Yes	No	NA	Unk	Yes	No	NA	Unk
101	X		X				X					X	
102	X		X			X					X		
103		X			X			X				X	
104	X		X			X							X
...													
...													
Totals													

- Check to see if answers to various questions make sense. (This is called a “logic check.”) For example, if a respondent indicates that he or she has no children, then all subsequent questions regarding his or her children should have a “not applicable” code. Any other response points to a data entry error.

The result of the data entry and error checks is an error-free data sheet or computerized data file you can use to analyze the data.

### Task 2: Tabulate the Data

The data obtained on each participant for each outcome indicator need to be added together to provide the overall value for that indicator for the reporting period. Most outcome indicators are expressed as the *number* and the *percent* of something, such as the number and percent of participants that achieved improvement. To calculate basic data:

- Count the total number of participants for whom you have data.
- Count the number achieving each outcome status (e.g., the number who demonstrated knowledge of a topic immediately following the program, the number who demonstrated behavior using the knowledge three months following completion of the program, the number who reported a change in condition after 6 months).

- Calculate the percentage of participants achieving each outcome status.
- Calculate other needed statistics, such as averages or medians.

To calculate the *number* of participants that achieved some particular improvement, identify the participants that met the achievement criterion and then count how many met it. To calculate the *percent*, divide this number by the total number of relevant participants. Who are considered “relevant” participants is determined by how the outcome indicator is defined. For example, if the outcome indicator is the “percent of participants who show improvement six months after completing the service,” then relevant participants are those who reached the six-month-since-completion mark during the reporting period. The percent is calculated by dividing the number of participants showing improvement by the total number of participants who reached this milestone; not by *all* participants; nor by the number of participants who *had been gone* from the program for six months but *didn’t complete* the program.

“We worked with a consultant from a nearby university. His graduate students did the tabulation and analysis, reducing our cost. Their work also enabled us to fill in gaps in our expertise and time.”

*Jane Robbins,  
Therapist, Samaritan  
Counseling Center of  
the Capital Region,  
Scotia, New York*

If your program obtained outcome data from a sample of participants using proper random sampling procedures so that the sample represents the full population of participants served by your program, then the program can estimate the total number of participants that achieved the specified outcome for that reporting period. This is done by multiplying the percent of participants in the sample that achieved the outcome by the total number of participants.

Some programs may include indicators expressed as *averages*, such as the average score on a test taken by many participants. However, averages provide no information on the *distribution* of scores and can be affected significantly by a few very high or very low scores. More information is conveyed if the indicator is expressed as a percent, such as the percent of participants whose score exceeded a certain level.

### **Task 3: Analyze the Data Broken Out by Key Characteristics**

Recall that identifying factors such as participant and program characteristics that could influence outcomes was an important task in identifying exactly what data you wanted to collect (Step 3). Now is the time to look at data broken out by these factors to see what information these analyses provide. For example:

- Data broken out by level of difficulty of incoming participants will help you assess whether changes in outcomes are related to an increase in the proportion of participants coming to a facility with more difficult-to-help conditions. They also can indicate whether differences in outcomes of two facilities are related to the level of difficulty of participants at each facility.
- Comparing program success rates for males and females, or for different racial or ethnic groups, may indicate that the program works better for some types of participants than for others.
- Data for participants at different program locations may suggest that some locations are having more success than others.

Many computer programs will do simple data breakouts through cross tabulations. Breaking out data manually is more time consuming. For instance, you have to count separately the females with “yes” responses, the females with “no” responses, the males with “yes” responses, and the males with “no” responses. The easiest way to accomplish this may be to use two spreadsheets—one for females and another for males.

Exhibit 6-B shows data on one outcome indicator for the Southside Children’s Agency Teen Mother Parenting Education program broken out by the age of the teen mother. Because the mother’s age generally is related to the health of the newborn, this is an appropriate way to break out the data for this program. The cross-tabulation table in the exhibit shows that, in fact, older teens were more likely to deliver healthy infants than were younger girls.

Exhibit 6-B

**Outcome of Delivering Healthy Babies, by Age of Mother**

Newborn weighed above 5.5 pounds, scored 7 or above on Apgar scale		Age		All Participants	
		<17	17+	Total	Percent of All
Yes	Number	8	5	13	72%
	Percent of age group	67%	83%		
No	Number	4	1	5	18%
	Percent of age group	33%	17%		
Totals		12	6	18	

## Task 4: Provide Explanatory Information Related to Your Findings

It is not enough simply to present your outcome data to your various audiences. You also should provide discussions and explanations of your findings to help readers understand what the numbers may mean. They need a context for the numbers that will deepen their appreciation for what your program does. Some programs, for example, are designed specifically to serve participants that present difficult problems. Discussions of outcome findings, especially when directed toward funders, should make this clear.

At times, your outcome findings may not meet expectations. There are many reasons for less-than-desired outcomes. Some reasons are *external* to the program. For example, an increase in unemployment in the community can result in an increase in domestic violence; a reduction in federally subsidized food programs may affect the health of women in your prenatal program.

Some reasons for disappointing outcome data are *internal* to the program. For instance, high staff turnover can result in participants progressing more slowly than expected. Or perhaps you are implementing a new curriculum that is less effective than the previous one. Conversely, in an earlier year, perhaps external factors were so favorable that the outcomes exceeded expectation. When this is the case, the following year might show what appears to be a lack of improvement.

When you present outcome data, include information that tells users the probable reasons why the outcomes look unusually high or low. For outcomes that are substantially worse than expected, also describe what you are doing or plan to do to improve them. This strategy lets you demonstrate that you recognize and have a plan to overcome the effect of events or occurrences that have affected your outcomes negatively. It also decreases the likelihood that funders and others to whom you are accountable will penalize you for what they might otherwise have viewed as “poor” outcomes.

## Task 5: Present Your Data in Clear and Understandable Form



As you prepare reports for various audiences, visual presentations in tables and charts will make the data more understandable to readers. Once you have developed presentation formats that your audiences like, using those same tables and charts each reporting period will help track changes over time. Although it is important to accompany visual displays with narrative discussions of the findings they portray, each table and chart

should be as self-explanatory as possible. Be sure to label all information clearly so readers know what is being presented.

Tables and charts you can use to record and summarize outcome indicators are described below. They can be generated by computer or drawn manually.

- **Data tables.** Data tables can summarize numerical data effectively by grouping together or breaking out relevant information. Exhibit 6-C is an example of a data table reporting data for one outcome of the Southside Children’s Agency’s Teen Mother Parenting Education program.

Exhibit 6-C

**Sample Outcome Data Table: Southside Children’s Agency  
Teen Mother Parenting Education Program**

	Total	Age			Living Situation	
		12-14	15-16	17+	Alone	w/ Parent
Number of babies born in reporting period	18	4	8	6	3	15
Newborns weighing above 5.5 pounds and scoring 7 or above on Apgar scale						
Number	13	2	6	5	3	10
Percent	72%	50%	75%	83%	100%	66%

- **Bar charts.** Bar charts display distributions of categorical data, such as year, sex, race, ethnicity, or type of referral. Each bar shows the number or percentage for whom a given outcome occurred. Exhibit 6-D is an example of a bar chart showing the percent of boys and of girls whose scores from pre- to post-test increased by six points or more.
- **Pie charts.** A pie chart illustrates the distribution of a given indicator. Each percentage is represented as a slice of the total pie; therefore, all slices must total 100 percent. Exhibit 6-E displays a set of pie charts showing contrasts between boys and girls and between children’s and parents’ responses.

Exhibit 6-D

**Percent of Participants With Pre- to Post-Test Increases  
of Six Points or More**

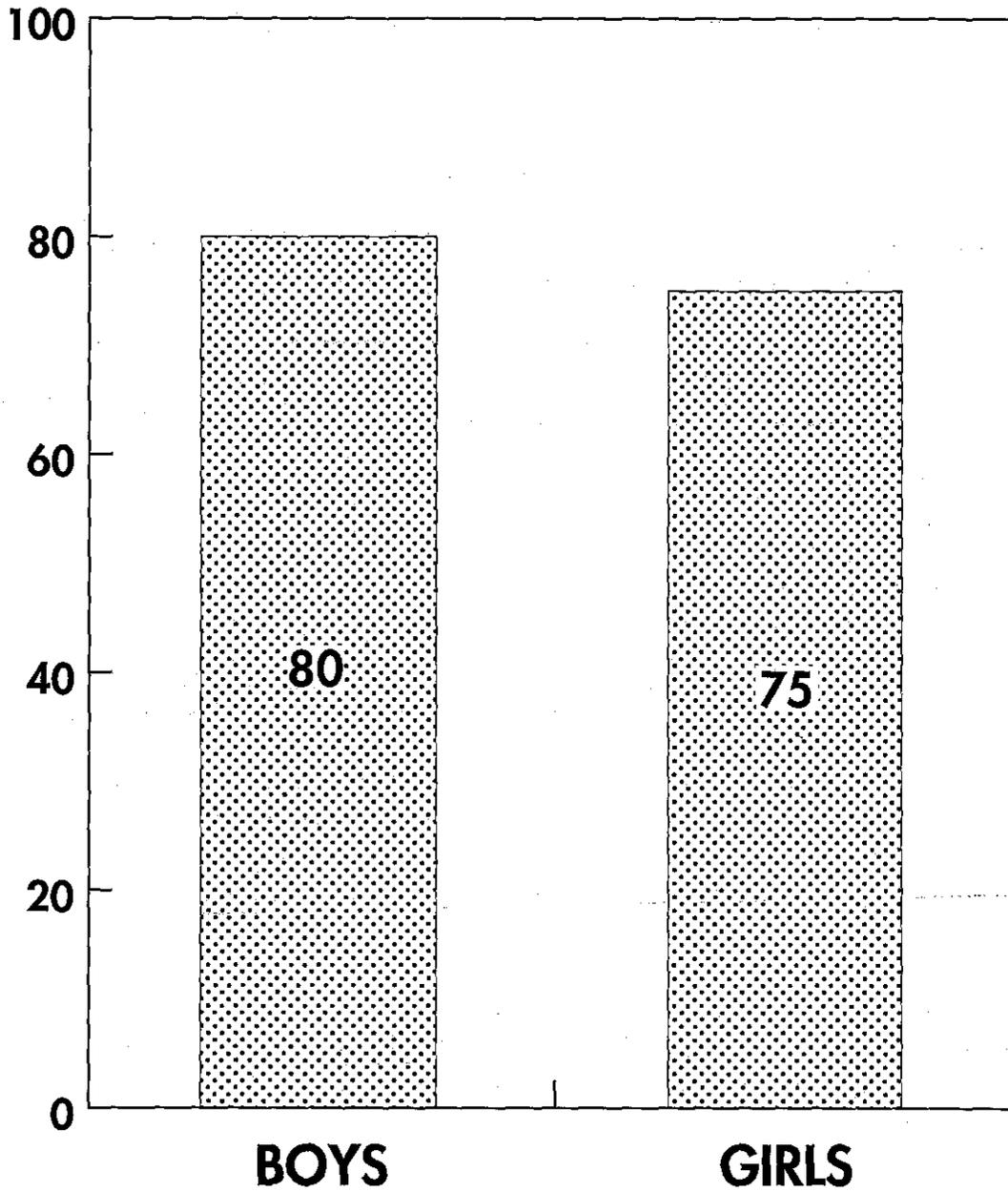
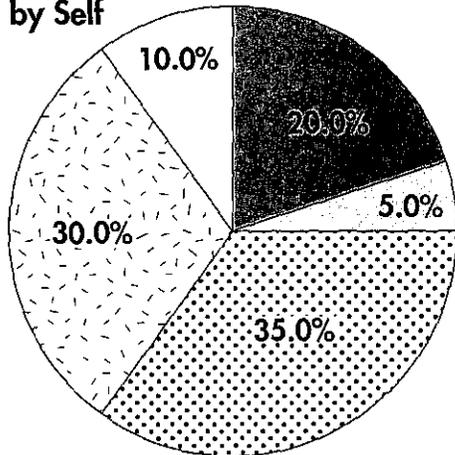


Exhibit 6-E

### Changes in Youths' Conscientiousness About Homework

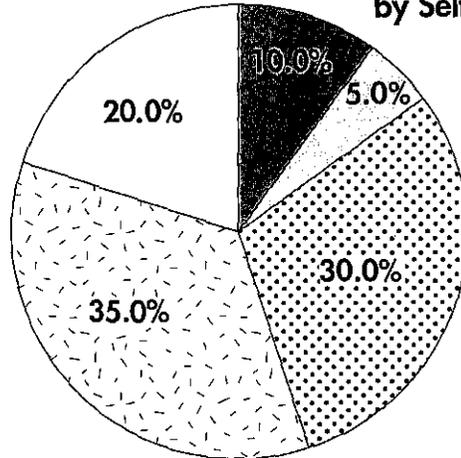
#### **GIRLS**

as reported  
by Self



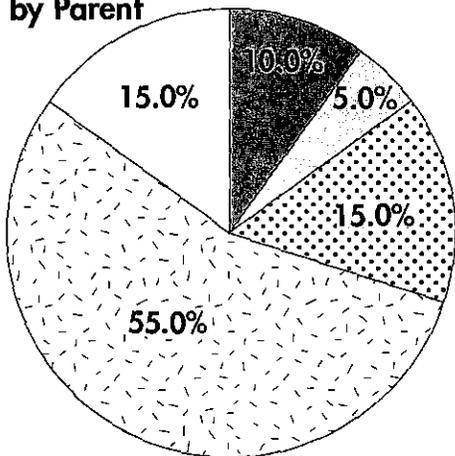
#### **BOYS**

as reported  
by Self



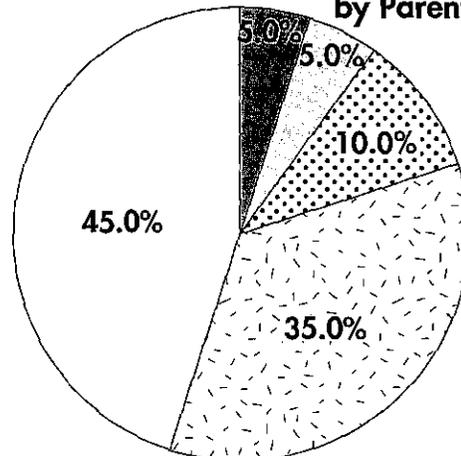
#### **GIRLS**

as reported  
by Parent



#### **BOYS**

as reported  
by Parent



Much Better  
Somewhat Better  
No Change  
Somewhat Worse  
Much Worse



- **Maps.** Mapping is an excellent way to present the geographic distribution of a problem or resource visually. Maps are particularly useful for tracking community indicators. Because different data sources often have unique reporting districts, you may need multiple maps for your outcome indicators. For example, you may need a health map, divided into health districts; an education map, divided into school districts; and a law enforcement map, divided into police jurisdictions.

After you have prepared sample reports of your trial-run outcome findings, get reaction to them from your key audiences. Ask staff, volunteers, board members, and others for their impressions and suggestions.

**T**he management at **Kingsley House and New Orleans Day Nursery Association, New Orleans, Louisiana** has established a feedback loop involving all relevant staff in the collection and analysis of its outcome measures. The loop begins when the front-line staff gather the raw data, which is compiled by individual program managers. An administrative assistant tabulates the information and presents it to agency administrators, who analyze and summarize the information.

The executives find that the most effective method of communicating the data is straightforward bar charts in which the first bar is the annual goal and the following bars represent each quarter's performance. The report is shared with program managers, who are encouraged to discuss it with their staffs. Each quarter the staff meets to interpret the newest data, discuss the reasons for the results, and identify any adjustments that need to be made.

As a result of this process, the director feels the staff have a greater feeling of pride and an expanded sense of initiative. The staff are more inclined to offer suggestions for their programs with an eye to affecting outcome measures.

- Do the findings seem reasonable?
- Are they presented clearly?
- What questions do they raise that are not answered in the report?
- Are explanations of problem areas and proposed remedies satisfactory?
- What other tables or charts would be helpful?

- Does anything seem to be missing, such as an overlooked outcome or influencing factor?

Such feedback can improve your outcome reporting significantly. It also may suggest useful modifications for your outcome measurement system.

Finally, be sure to talk with the individuals who coded, entered, tabulated, and analyzed your data. Their suggestions regarding data collection instruments and procedures can help tighten that crucial part of your process. They also may have ideas for how you can respond to feedback from key audiences on the outcome reports you prepared.

The main payoff for the **Adult Literacy Program of Literacy Volunteers of America-Schenectady, Inc.** is that the staff now have an abundance of information to help portray the benefits of the program to constituents, including funders, individual donors, and the community.

The information is presented in a two-page Program Profile. Staff believe that the new quantitative summaries of outcomes make a more compelling case for the program. Outcomes are presented as achievements enabled by literacy skills. *"Participants report 195 achievements in Personal Survival Skills, such as reading labels on a medication bottle or grocery store product, learning to write one's name and address, and learning to read U.S. currency; and 119 achievements in Social & Community Life, such as voting, applying for citizenship, being removed from public assistance, and obtaining and using a library card."*

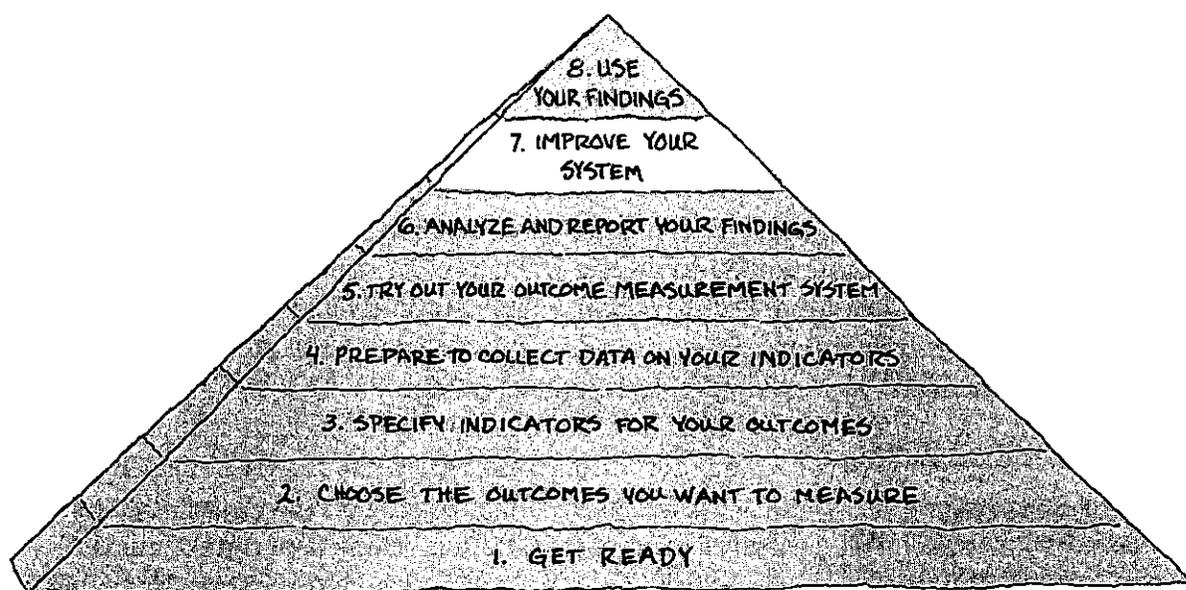
Anecdotal information makes the program real by showing what actual program participants have accomplished. *"Having gained confidence in his writing ability, J. now leaves a 'to do' list for the person on the next job shift."*

Skill development data (*51% of basic literacy students advanced at least one reading level since enrollment*) and output statistics (*286 adults served, 195 in English as a second language and 91 in basic literacy*) complete the picture.

The agency has used this format for three years and has received positive feedback from various recipients.



## Step 7: Improve Your Outcome Measurement System



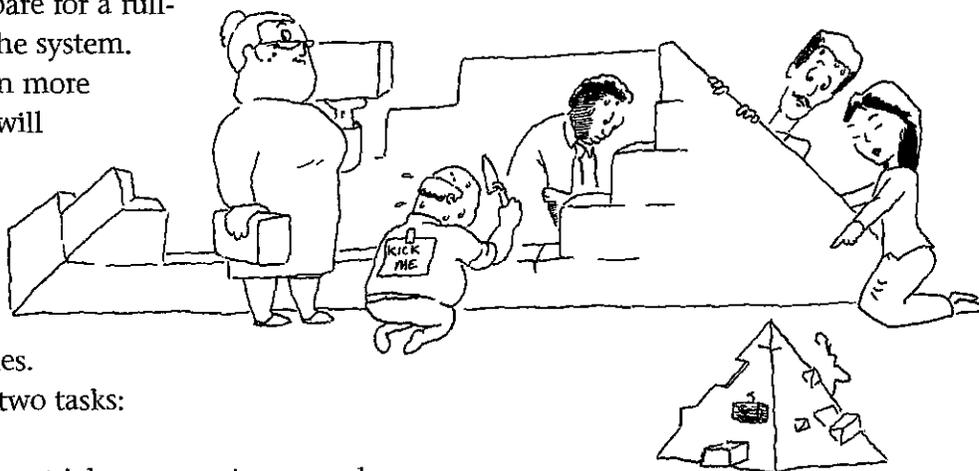
**T**he trial run of your outcome measurement system should have given you a lot of valuable information about your process and how it might be improved. Now is the time to prepare for a full-

scale launch of the system.

Then, as you gain more experience, you will continue to monitor and improve your approach to measuring program outcomes.

Step 7 addresses two tasks:

1. Review your trial-run experience, make necessary adjustments, and start full-scale implementation.
2. Monitor and review your system periodically.



## Task 1: Review Your Trial-run Experience, Make Necessary Adjustments, and Start Full-scale Implementation

During the trial run, you monitored various aspects of your outcome measurement process and gained feedback from data collectors, respondents, and data coders and analysts. You saw the collected data transformed into outcome findings and prepared sample reports for various audiences. You also got feedback on the content and presentation of the reports.

Now, your work group should review all aspects of your trial run to identify what worked well and how the system can be improved. The place to start probably is with the outcome findings themselves. Did you get all of the data you needed? Did you actually measure what you intended to measure? Does what you measured still seem to represent important outcomes for which your program should be accountable? Even if every part of your data collection process went well, a “no” to one of these questions calls for adjustments in your outcome measurement system.

“After the first year, we reviewed the process. I wish we’d have monitored it more closely on a quarterly basis so we could have adjusted one of our outcome measurements earlier.”

*Michael Moreau,  
Executive Director,  
Kingsley House and New  
Orleans Day Nursery  
Association,  
New Orleans, Louisiana*

Key uses for outcome information are to track changes in outcomes from one reporting period to the next and to identify longer-term trends. These uses cannot be met until the system has been operating for a while. Nevertheless, it should be possible to tell whether ongoing data collection will yield information useful to program planning and decision making. This examination may suggest additional information that program personnel will want to collect to give them a fuller perspective on program outcomes.

If it appears that you are on the right track regarding the outcomes you are measuring and the data you are using to measure them, you still need to assess other aspects of your system. Your work group should review:

- **Data collection instruments.** Even though you pretested your instruments in Step 4, the trial run may reveal problems. Talk with data collectors and examine findings to detect difficulties with instructions, question wording and content, response categories, layout and format, and length.
- **Training of data collectors.** Check with the data collectors to assess whether their training was adequate and what improvements are desirable. Check the results of the data collection effort to assess whether the data obtained was what was intended.

- **Data collection procedures.** Talk with data collectors to assess the adequacy and quality of the procedures. Was the data sought from agency records available as expected? If not, why? Review the data obtained to determine whether it was what was intended and if data collectors or respondents had problems providing the information requested. Examine survey response rates. If they are low, try to determine why.
- **Data entry procedures.** Did the process work smoothly? Were data entry personnel trained adequately? Were the data available to them in an understandable manner? Did equipment and computer software work properly? Were quality control methods effective?
- **Time and cost of collecting and analyzing the data.** The biggest cost is likely to be staff time from various types of personnel, including service delivery staff, clerical personnel, and staff involved in data coding and analysis. Other costs may include paper, duplication, postage, and perhaps computer software and consultant time.
- **Monitoring procedures used during the trial run.** Your monitoring procedures initially were established to give you information about your try-out period. However, you will continue to monitor your system after you have started full-scale implementation. Therefore, after reviewing the information your monitoring procedures provided, review the procedures themselves. What do you wish you had tracked, or had tracked differently? Would a new form be helpful? Do you need less detail than the current procedures obtained? Should you collect and review the information more—or less—often? Should other individuals be involved?

Based on your review, make whatever refinements are indicated to improve the effectiveness and ease of administration of your outcome measurement system. Document the revised procedures so they will be implemented consistently. This is important to produce outcome data that are comparable from year to year and, if the program has more than one facility, comparable across facilities.

Once the revisions are in place, you are ready to start full-scale implementation. If all goes well, you will not need to make other major changes to the system for a while. Unless the changes you have made are major, data collected during the trial run can be combined with data collected during the subsequent full-scale implementation.

## **Task 2: Monitor and Review Your System Periodically**

Although you hope not to need major changes for a while, over time, there will be changes. You will gain more experience with all aspects of the system; your program will

alter in various ways; your environment will change; and your priorities will be modified. Your outcome measurement system needs to keep pace with these changes.

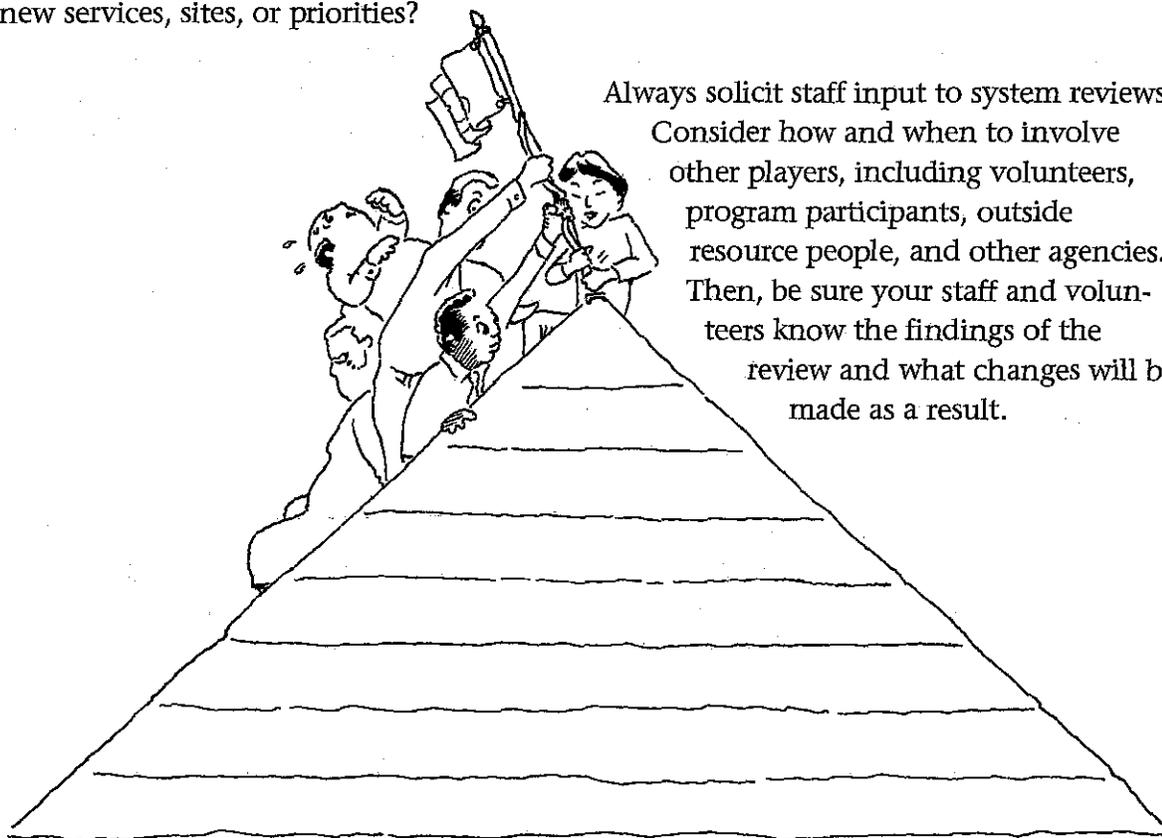
Although the trial run has ended, you will continue monitoring your outcome measurement process. Using the monitoring procedures you revised following the trial run, you will continue to track time, costs, response rates, success of various collection procedures, and other information important to your agency.

Build in a periodic formal review of your system, much like the review you completed at the end of your trial run. Examine the monitoring information to see what processes have improved, which have slipped, and what continues to be troublesome. Also consider changes that have occurred in the program and how to reflect them in the system. Finally, look ahead to program alterations still in the planning stage. What will they mean for outcome measurement? Should you start now to accommodate new services, sites, or priorities?

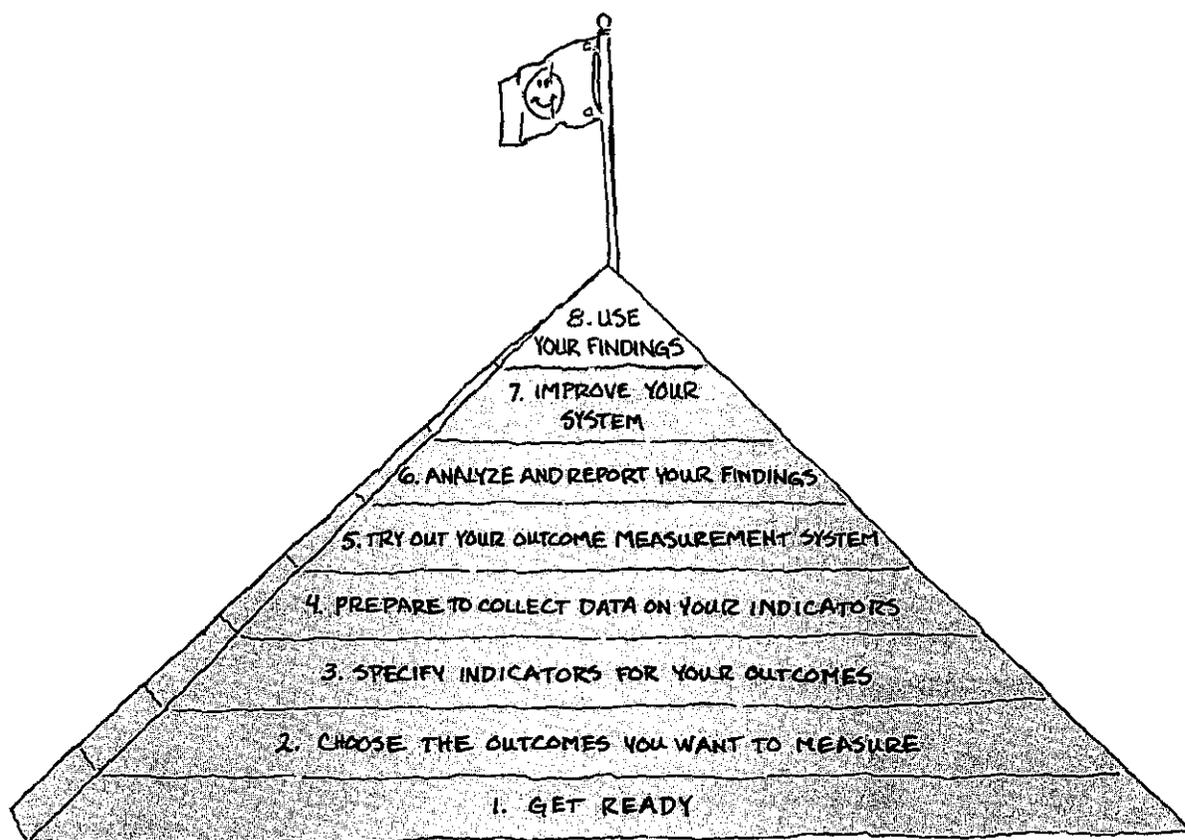
"We reassess our measures quarterly. This keeps us constantly aware of what's working and what may need changing. We're going to convert all our programs to outcome measurement. That way we know when we're really doing something and not just counting numbers."

*Royce Van Evera,  
Executive Director,  
Law, Order & Justice Center,  
Schenectady, New York*

Always solicit staff input to system reviews. Consider how and when to involve other players, including volunteers, program participants, outside resource people, and other agencies. Then, be sure your staff and volunteers know the findings of the review and what changes will be made as a result.



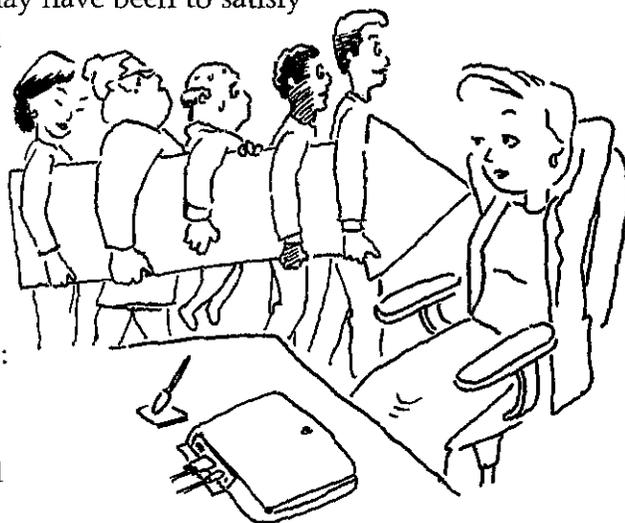
## Step 8: Use Your Findings



**Y**ou have designed and implemented a system for measuring your program outcomes and have analyzed the resulting data. Now, it is time to make maximum use of your findings to improve and promote your program. Although your initial reason for implementing outcome measurement may have been to satisfy a funding requirement, outcome information has many more, and often more important, applications in program planning and management.

Some uses of outcome data are internal to your organization. The findings represent feedback on how well the program is doing—feedback that can, for example, help:

- Provide direction for staff.
- Identify staff and volunteer training and technical assistance needs.



- Point out program improvement needs and effective strategies.
- Support annual and long-range planning.
- Guide budgets and justify resource allocations.
- Suggest outcome targets.
- Focus board members' attention on programmatic issues.

Other uses are external. Outcome findings are a demonstration of the program's value and role in your community's human service network. By casting the findings in this light, your program can use them to:

- Recruit talented staff and volunteers.
- Promote your program to potential participants and referral sources.
- Identify partners for collaborations.
- Enhance your program's public image.
- Retain and increase funding.

The more you incorporate outcome information into ongoing management and operation activities, the more the effort invested to develop and implement your outcome measurement system will pay off.

## Internal Uses of Outcome Findings

Following are some ideas for using outcome data in program management activities.

### Provide Direction for Staff

Favorable outcome findings give staff a well-deserved boost. Problematic findings focus staff attention on service aspects or participant groups needing additional attention. As soon as outcome findings are available, they can be conveyed to staff—perhaps by disseminating copies of the outcome reports or by posting the information at convenient locations.

Also useful to staff are the compiled comments participants made during interviews or on questionnaires, including explanations for why they gave a negative rating to some aspect of the program and suggestions for program improvement. (Participants' identities must not be shared with staff.) These comments can help staff pinpoint problems and devise program improvement options.

"The staff is excited because they can be in touch with the area they work in as well as the entire agency. It also helps the staff work more diligently. We wouldn't have imagined we impact lives as well as we do."

*Sue Malkin,  
 Executive Director, Deaf,  
 Hearing and Speech Center,  
 Detroit, Michigan*

Program managers can use the latest outcome report and narrative comments at “How Are We Doing?” meetings where managers and staff discuss:

- Where have we done well and why?
- Where are we not doing well and why?
- What can we do to improve results?
- Do the measurement techniques need tuning?

After improvement plans are implemented, staff can monitor later outcome reports to determine whether the results of the actions are as expected and whether more changes are needed.

## Identify Training and Technical Assistance Needs

“At a project or program level, this has helped the staff focus and be more organized in timeliness of service delivery and data collection, and it’s fueled by this outcome measurement.”

*Karen Russell,  
Public Health Nurse  
Mentor/Project Manager,  
Healthy Start Program,  
Tacoma-Pierce County  
Health Department,  
Tacoma, Washington*

Outcome findings can indicate areas where service delivery personnel—both staff and volunteers—are having special difficulties and could benefit from focused training or technical assistance. Problem areas might include specific subjects, certain service delivery methods, or particular participant groups.

Program managers also can examine outcomes for participants served by different staff members or volunteers to help identify training and technical assistance needs for individual personnel. If undertaken, this use should be strictly internal, confidential, and constructive. Consider differing levels of difficulty among participants served by various personnel in interpreting the findings.

## Identify Program Improvement Needs and Strategies

As part of a program’s continuous improvement process, information on a program’s level of success in achieving the desired outcomes for its participants is a valuable resource for organizational learning. Even trial-run data can raise questions about service delivery methods, priorities, and training needs.

Outcome findings do not, however, tell the program why it is achieving at a particular level, whether that level is



good or bad, or how to improve. It is up to the program—its managers, staff, and volunteer policymakers—to decide what the outcomes mean and what to do about them. This requires going beyond measuring outcomes to interpreting and evaluating them. There are many ways that a program can assess the meaning of its outcome data.

“Outcome measurements helped us evaluate the effectiveness of a new program area and identify an important market niche. We tiptoed into a program that matched developmentally disabled children with local high school students. The feedback from this small-scale effort was overwhelming. Parents are ecstatic that their kids are having normal interaction and companionship with other kids, and we learned that it was not difficult to make these kinds of matches. The outcome measurements have proven that we have really met a need.”

*Joani Good,  
Executive Director,  
Big Brothers Big Sisters  
of Delta County, Inc.,  
Escanaba, Michigan*

The strategies described below involve comparing outcome data for two or more groups. In examining differences between groups, it is important to distinguish between statistical significance and practical significance. Differences that are very small are not likely to be of practical importance, even if they are “statistically” significant. As a rule of thumb, differences of less than 5 percentage points seldom are meaningful for program managers or funders. Differences of 10 or more percentage points are more likely to be of practical concern; they warrant attention and further examination of their causes. Keep in mind, however, to consider the number of participants in each group. If a group has only 10 participants, then a difference of one person creates a 10 percent difference in outcome results—a difference that does not have much practical significance.

Various statistical procedures can determine the likelihood that observed differences could have occurred by chance alone. If you want this level of analysis of your data, you will need assistance from someone trained in statistics. Possibilities include a volunteer from a local college, university, or business, or a retired expert.

To learn more about where it is being effective and where it needs to improve, a program can examine its outcome findings to determine where it is having more and less success in achieving desired outcomes. It can, for example:

- **Compare outcomes for different participants.** By looking at outcome data broken out by important participant characteristics, a program can determine what characteristics seem to relate to better outcomes. This may guide program managers to target services toward specific participant groups. It also may guide program staff in finding more effective strategies for serving participants who are not experiencing the desired benefits.

- **Compare outcomes for different program units, facilities, or other organizational entities.** Breaking out outcome findings by organizational unit will show what parts of the program are achieving better outcomes. Large programs particularly are likely to experience uneven success across units.

The data table in Exhibit 8-A, which was used in Step 3 to demonstrate the value of breaking outcome data out by key factors, reflects at least two findings that could guide service-improvement efforts by this program. First, the data show that the level of success for Unit 1 is lower than that of Unit 2 for participants who came in with both difficult and less-difficult problems. This would lead the staff of Unit 1 to evaluate why its outcomes are lower, and perhaps learn from the other unit. Second, the table shows that the distribution of difficult and non-difficult cases between the two units is quite uneven, with Unit 2 serving three-quarters of the difficult cases. If this is not intentional, the program would want to examine its intake procedures to see why this pattern has occurred.

"We've adjusted our programs according to what the evaluation measures show has worked or not. For example, in our Girl Scouting in the School Day program, we used to focus more on arts and crafts until we readjusted the program to concentrate on what the parents and schools wanted. Now we focus on conflict resolution, activities to supplement the history curriculum, and living skills."

*Nancy Berg,  
Executive Director,  
San Francisco Bay Girl  
Scout Council,  
Oakland, California*

Exhibit 8-A

### Outcome Findings by Program Unit and Level of Difficulty of Participants' Problems

	Unit	
	1	2
<b>All Participants</b>		
Total	500	500
Number Helped	300	235
Percent Helped	60%	47%
<b>Non-Difficult Cases</b>		
Total	400	200
Number Helped	300	160
Percent Helped	75%	80%
<b>Difficult Cases</b>		
Total	100	300
Number Helped	0	75
Percent Helped	0%	25%

"We made several program changes as a result of the outcome measurements. For one, our pregnancy rate was higher than we wanted, so we worked with another agency to add a sex education component to the program. In the aftercare area, the data showed that kids excelled in the GED program, but they had not gone through enough life skills classes. As a result, they had not changed the behavior they were exhibiting. We decided to slow down the academic aspect of the program and infuse more concentration on life skills."

*Nina Waters, Executive Director,  
Practical and Cultural Education  
(PACE) Center for Girls,  
Jacksonville, Florida*

- **Compare outcomes for different service delivery strategies.** Program policymakers can use outcome findings to compare the relative success of different approaches to service delivery, including variations in the type or intensity of service. This can help assess both existing variations in service and new approaches that are introduced. Exhibit 8-B describes how a program can use outcome findings to compare alternative service approaches.

A simpler strategy, but one that does not provide as strong evidence, is to introduce a new service approach across the board and compare outcome data for the new approach to data from the approach that it replaced. In either strategy, program personnel will be able to distinguish under which conditions outcomes have been higher and lower. This may lead them to place more emphasis on certain approaches or to seek further information on why the differences occurred.

Exhibit 8-B

### **Comparative Findings for Two Youth Services Approaches: Team vs. Individual Activity Emphasis**

(Data are for a one-year test; about 150 youth are included in each procedure.)

Approach Used	Percent of Youth Showing Increased Verbal Skills	Percent of Youth Showing Increased Interpersonal Skills
Team Activities	88%	46%
Individual Activities	92%	28%

#### **Steps for Comparing Alternative Program Approaches**

1. Identify the two practices or approaches to be compared. Perhaps one is an existing practice and the other is a new approach.
2. Choose a method for deciding which incoming youth will be served using which approach. The method should select a representative sample of the participants for each

### Exhibit 8-B, cont'd

practice with approximately the same proportion of difficult participants in each of the comparison groups. Some form of random assignment is preferable. Randomization helps assure that the comparisons will be valid and greatly increases the strength of the evidence. Simple methods of random assignment include flipping a coin and using a table of random numbers. If the arrival of participants is essentially random, another method is to assign incoming participants alternatively to each of the practices.

3. As each participant enters the program, assign her or him to one of the two groups by the procedure you have just identified.
4. Record which participants are assigned to which procedures.
5. Track the outcomes for each participant in each practice over whatever period of time the program believes is necessary to indicate the outcomes of these practices fairly.
6. Tabulate the values on each outcome indicator for each of the two practices.
7. Compare the findings and make adjustments to program practices as appropriate. You may want to drop the practice that shows the poorer outcome. Alternatively, you might decide that you have not received a clear enough picture from the outcome data to make a decision, in which case you will continue the comparison for a longer time.

In the example given above, team activities show substantially better outcomes on increased interpersonal skills. Both types of activities show high levels of success related to increased verbal skills. The outcome for individual activities is slightly higher than that for team activities, but the difference falls below the 5 percent threshold for practical significance. The program might feel that the evidence from the test indicates that team activities have been more effective and put more emphasis on these in the future.

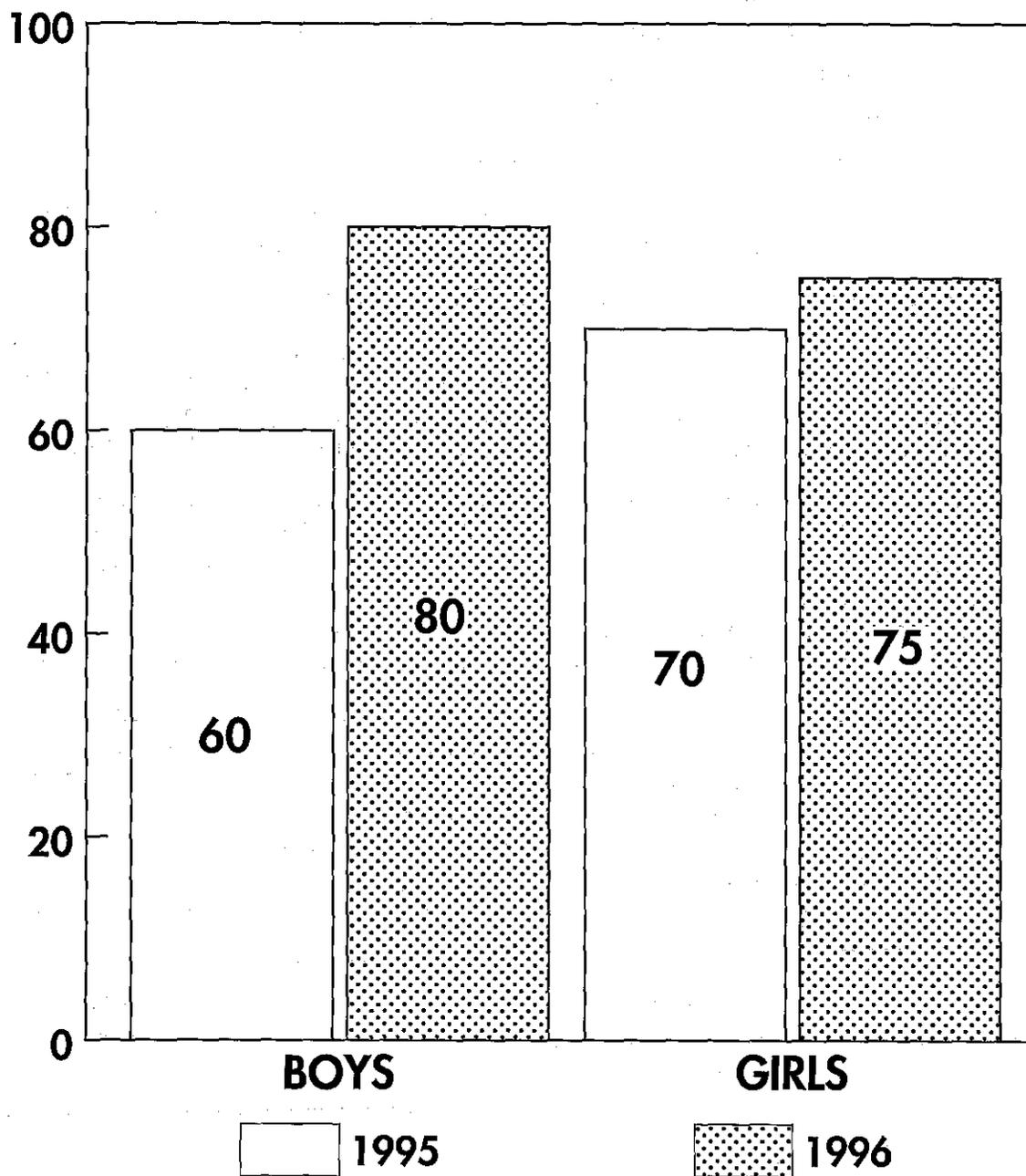
Note that outcome findings alone cannot guide program improvement choices. Other information, such as the resource requirements (cost) for each option, must be taken into account. Ultimately, decisions about program content are judgments by program officials. Outcome information better informs those decision makers before they make choices regarding which participants will be served under the new procedures and which will be served using the existing procedures.

- **Compare outcomes with those of previous years.** After a program has collected one or more rounds of outcome data, comparisons with its previous performance will be an important part of its program review and improvement activities. Exhibit 8-C, for example, compares 1996 outcome data presented earlier in Exhibit 6-E with data from the program's previous year.

Comparisons like these show the effect of program improvement efforts on the program's outcomes. In making such comparisons, it is important to identify and track outside influences that can affect program outcomes, such as changes in the local

Exhibit 8-C

**Percent of Participants With Pre- to Post-Test Increases  
of Six Points or More**



"At first we met with resistance from the staff, but now they have a greater sense of pride as they see the outcome of their work. The staff now talks about changes in the program that affect the outcome measures. They are really showing more initiative."

*Michael Moreau,  
Executive Director, Kingsley  
House and New Orleans  
Day Nursery Association,  
New Orleans, Louisiana*

economy, events that bring public attention to problems the program addresses, and variations in the program's funding levels.

■ **Compare outcomes with targets set for those outcomes.**

After a program has refined its outcome measurement system and is fairly confident of the data it is collecting, it may begin to set targets for future performance and then compare its actual performance with those targets. Caveats about setting outcome targets are presented later in this step.

Exhibit 8-D is an example of findings on three outcome indicators arrayed against targets and across

time periods. The findings indicate that the Child Abuse Public Awareness and Prevention Program of the Southside Children's Agency has improved its performance on two of the three indicators. However, it continues to fall short of its targets on two indicators, even after adjusting one of those targets downward.

Breaking out the target audience by profession (e.g., educators, child care workers, medical professionals, general public) would help the program know which audiences it is being more and less successful in reaching and which are more frequently reporting suspected abuse and neglect. This would help the program know which audiences may need more focused attention or more tailored messages.

"The information helps social workers to be very focused in their work with the kids and the more focused they are, the more positive and more effective are the outcomes we see. It is also beneficial to us administratively to know that we are accomplishing our objectives, and the payor sees results."

*John Steele, Executive  
Director, Catholic Family  
and Community Services,  
Schenectady, New York*

In addition to making the internal comparisons described above, a program may find it informative to seek outside sources of comparative data. These can include performance data of other agencies with similar programs, data from national surveys or other studies, and data on target populations similar to the program's participants who are not receiving services. The line graph in Exhibit 8-E shows the incidence of burglaries in the Southside Neighborhood and a nearby neighborhood before and after the Southside Children's Agency initiated its neighborhood organizing program.

Comparisons with other programs or target groups must be interpreted with care. Differences in population characteristics and resource levels are very likely to result in outcome differences that are not necessarily reflective of program quality or effectiveness.

Exhibit 8-D

**Southside Neighborhood Children's Agency  
Child Abuse Public Awareness and Prevention Program:  
Actual Outcomes vs. Targets**

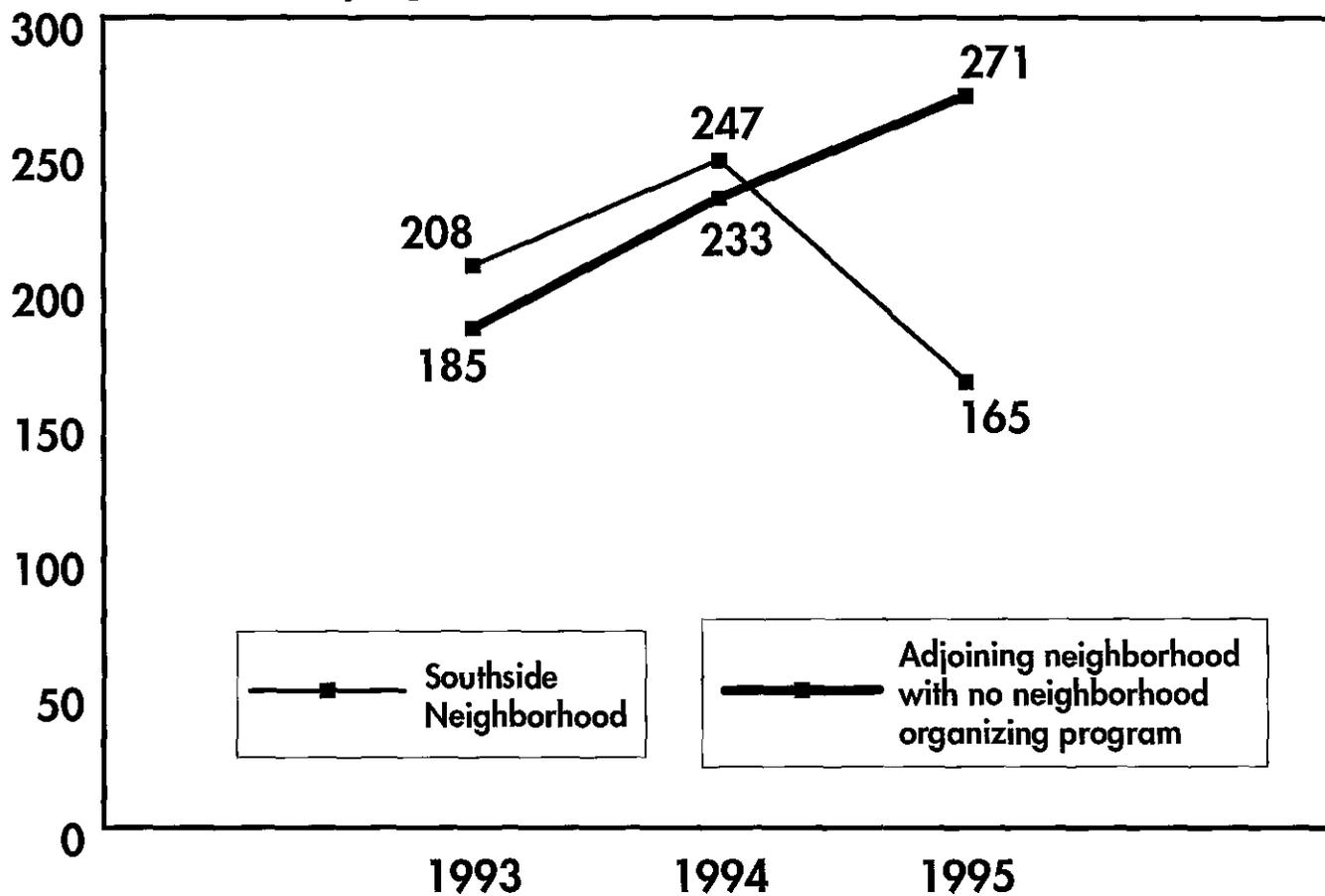
Outcome Indicator	Last Period			This Period		
	Target	Actual	Difference	Target	Actual	Difference
1. Percent of target audiences that recall content of PSAs, brochures, posters, or presentations	60%	40%	-20%	60%	45%	-15%
2. Percent of target audiences that can name 3 signs of abuse and neglect and can identify police or child protective services as appropriate contact	40%	30%	-10%	35%	40%	+5%
3. Percent of target audiences that say they have made such a report since hearing or seeing information.	20%	10%	-10%	15%	10%	-5%

Note: "-" means the actual outcome is worse than the target  
 "+" means the actual outcome is better than the target.

There are a few instances of standards for outcome indicators developed by a unit of government or a professional association. For example, the U.S. Department of Labor's performance measurement process for the Job Training Partnership Act includes annual calculation of national standards based on the most recent outcomes achieved by local programs. These standards address indicators such as adult employment rate 13 weeks after trainees complete the program. The Department of Labor recommends local adjustments to these standards to reflect local economic factors and participant characteristics. Another example of standards for outcomes are the targets established as part of the Healthy People 2000 effort. These examples notwithstanding, standards for outcome indicators are currently rare.

Exhibit 8-E

### Burglaries in Southside Neighborhood Before and After Implementation of Neighborhood Organizing Program (program instituted in November 1994)



## Support Annual and Long-range Planning

"This is definitely strategic planning fodder. It will certainly help longer term when the board has to make tough decisions."

*Royce Van Evera,  
Executive Director,  
Law, Order & Justice Center,  
Schenectady, New York*

The program analysis and improvement activities described to this point feed directly into both long-range strategic plans and annual action plans. Decisions about staff and volunteer training initiatives, expansion of effective services and delivery approaches, and new service needs all will be reflected in annual and long-range plans. Outcome information also can help policymakers track ongoing problem areas and decide whether to continue efforts at improvement or to discontinue a service.

An agency's planning activities should include plans for improving its outcome measurement system and perhaps expanding the system to additional programs. Program managers and the board also can consider what additional uses they can make of outcome measurement to help the agency monitor progress on its annual and long-range objectives.

## Guide Budgets and Justify Resource Allocations

The use of outcome data in annual and long-range planning ties naturally into decisions concerning funds and other resources. Findings that suggest a need for more emphasis on a certain outcome also help justify devoting additional resources to achieving it. Findings that prompt expanding a program that is successful with one population to reach a more-difficult-to-serve population also may prompt the allocation of more dollars per participant for the expanded service.

## Suggest Outcome Targets

A useful management planning tool is to ask program managers to establish targets for each outcome indicator for the upcoming reporting period. Generally, programs should not attempt this until they have analyzed at least one round of outcome data and have some basis for setting a target. Otherwise, targets are likely to be little more than guesses. In selecting targets, consider the following:

- **Outcome findings for previous reporting periods.** These usually will be a major factor in determining targets. To take account of differences among participant groups and to avoid the temptation to select participants based

"Personally, as a leader in a men's group, I am interested to have more than anecdotal information about where we're going and how we're doing. It will also have impact for programming."

*Larry Agle,  
Executive Director,  
Family Service of the East Bay,  
Oakland, California*

on the likelihood of achieving desired outcomes, set different targets for different participant groups.

- **Outcomes achieved by the most successful service unit** if the program has more than one unit that provides the same service. (This approach assumes that the most successful unit isn't dealing with an easier clientele.) For example, a program with the outcome data shown in Exhibit 8-A might select outcome levels of 80 percent and 25 percent as its next-year targets for participants entering the program with non-difficult and difficult problems, respectively. This assumes that the program believes the outcome findings are valid and these levels are reasonably achievable. A more conservative approach would be to set the targets at the average of the service units. For the program in Exhibit 8-A, this would establish targets of 77 percent and 12 percent, respectively, for participants entering with non-difficult and difficult problems.
- **Outcomes achieved by other programs**, whether in your community or elsewhere. Consider carefully how well other programs' service delivery approaches, resources, participants, and community context match your own in deciding how much to rely on their outcomes in setting targets for your program.
- **Your budget**, that is, the funds and other resources expected to be available. Reductions or increases may be a good reason for lowering or raising targets.
- **Any external influences of which you are aware that can be expected to affect future outcomes**, such as downturns or upturns in the economy.

**O**utcome data have proven especially effective as a tool for educating the program committee of the advisory board of **Associated Catholic Charities in New Orleans, Louisiana** about the agency's 57 programs. The information illuminates each program's progress and individual issues.

The committee reviews specified programs at each bi-monthly meeting. In advance of the meeting, members receive program summaries that include outcome data along with program mission, goals, values, and results. The outcome information provides substance for discussion and helps keep the committee focused on the actual workings of the programs. Members provide more informed feedback because they have a deeper understanding of program issues and obstacles. No longer giving blanket approvals, members now offer suggestions on measuring particular goals and provide leads on additional resources.

### **Focus Volunteers' Attention on Policy and Programmatic Issues**

Board members and other volunteers on key committees, such as strategic planning and program evaluation, will be involved at several points in the development of a program's outcome measurement system. The process of discussing, and perhaps approving, the list of desired program outcomes will help volunteer policymakers

clarify their understanding of the program's purpose and focus their attention on benefits for participants, rather than on day-to-day operations.

Outcome measurement findings give volunteer leaders explicit feedback on how well the program is doing for participants. When findings fall short of expectations, program managers can use the focus on outcomes to help volunteers weigh alternative solutions in terms of the program's goals. Many of the other issues and problems the board discusses also can be framed in terms of their relationship to desired program outcomes. In this way, program managers can help volunteer leaders direct their attention to strategic-level decisions and set coordinated policies based on a clear understanding of program aims.

"We shared the information with our board of directors and it put them closer to the customers they serve and helps connect them. The measurements have also proven to have implications for program development. For instance, because of the feedback on our questionnaire, we expanded the hours of operation at our main center."

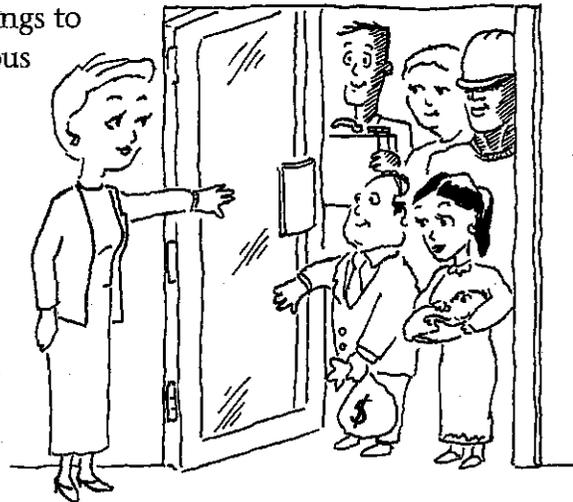
Marilyn Lange,  
Executive Director,  
Village Adult Day Center, Inc.,  
Milwaukee, Wisconsin

## External Uses of Outcome Findings

There are many ways to use outcome findings to enhance a program's interactions with its various publics. A few of these are discussed here.

### Recruit Talented Staff and Volunteers

Many potential staff will be attracted to a program that identifies and tracks desired outcomes and uses the information to improve its services. They will recognize that this program is serious about creating benefits



for participants

and is on the leading edge of a growing movement to increase accountability in human services.

"The board members feel more secure in their jobs since part of their job is to advocate in the community. Now they have documentation of results."

Ellie Gersten, Director,  
East Valley Catholic  
Social Services,  
Mesa, Arizona

Potential volunteers will appreciate a clear picture of what the program wants to accomplish for participants so they can determine if the program's aims match their values and priorities. Volunteers also are more likely to be drawn to a program that can describe the outcomes its participants experience than to one that describes its activities but not their results.

The inclusion of outcome information strengthens a program's staff and volunteer recruitment materials. In addition to describing good working conditions and competitive benefits, recruitment ads and brochures can include key outcome findings demonstrating that the program makes a difference for those it serves.

"Outcome measurement helped us better market our program to parents and youth. Parents of a Scout want him to develop skills for a lifetime. They want him to build self-esteem and leadership ability, and to carry out tasks in a timely manner. We have learned to say '65 percent of our Scouts learn skills that they will carry through life,' instead of '65 percent of them will earn merit badges.'"

*R. Keith Christopher,  
Director of Support Services,  
Boy Scouts of America,  
New Orleans Area Council,  
New Orleans, Louisiana*

program is a good match for the individual's needs and is likely to be effective. By sharing outcome findings with potential referral sources, a program helps clarify not only the problems it addresses, but also the solutions it seeks. It also offers reassurance that individuals referred to the program will receive services that promote results.

## Promote Your Program to Potential Participants and Referral Sources

When individuals or families are searching for a service, they want a provider that will be effective in helping them meet a goal or achieve a change. Outcome findings communicate the benefits other participants have gained in a program. This tells potential participants that the program is effective in these areas and helps them decide if the program is appropriate for their needs. Once participants decide to use a particular service, the outcome information helps define what they can expect from the program so that they and program staff are working toward the same objectives.

Similarly, when staff of an information and referral service or other referral source suggest a program, they want to know that the

**A**mong the programs offered by **Hale 'Opio Kauli, Inc., of Lihue, Hawaii** is a therapeutic group home program providing behavioral management to troubled adolescents. Each resident's progress is assessed by an outcome-focused plan developed with the youth. Milestones are structured in a logical progression towards outcome goals. Residents are keenly interested in their achievements because the attainment of each successive milestone earns privileges and responsibilities in the program, and graduation occurs when the outcome goals are reached.

## Identify Partners for Collaborations

Outcome findings can help programs identify various opportunities for collaborative relationships. A program may find, for example, that it is having difficulty achieving an outcome because of a key influencing factor that is beyond the program's scope. One response is to collaborate with programs focused on that factor, making their services

available to the first program's participants. For instance, a job skills training program may find that its outcomes are not reaching desired levels because participants are having trouble finding child care. The program may collaborate with a nearby child care program to make high-quality care available to children whose parents are receiving job training.

A program whose outcomes show that it is less effective with particular groups of participants may collaborate with programs that specialize in serving those participants. For example, a program whose outcome levels for elderly participants are lower than for younger individuals might create a partnership with a program that serves elderly clients successfully. The partner program need not offer the same type of service. The point is that its delivery approach, its staff, or some other program feature appears to interact more effectively with older persons. Managers of the partner program may provide insights that will help the first program improve its outcomes for older participants.

A program that is achieving good outcomes in its current service area may decide to extend the reach of its services. It might accomplish this by collaborating with programs already operating in the new service area. Programs that offer complimentary services and have made outcome measurement part of their management system would be good candidates for this relationship.

A shared focus on outcomes can prompt collaborations of programs with similar missions. Collaborators may, for example, compare outcomes of alternative services and service delivery approaches. They may work jointly to extend services into underserved geographic areas or to expand the scope of existing services. They may share the cost of training, technical assistance, or specialized resources to monitor and improve service quality. Particularly where funders are requiring collaborative efforts as a condition of grant awards, collaborations that use outcome measurement to demonstrate their value may become increasingly important to not-for-profit human services.

"There is a collaboration among groups, especially in the Prevention Partnership. Now we can compare not only internally, but across programs in the Tacoma area. There is the beginning of some common data collection and common outcome measurement. Not only can we see if we're making a difference, we also know if someone else is trying to do the same thing."  
*Karen Russell, Public Health Nurse Mentor/Project Manager, Healthy Start Program, Tacoma-Pierce County Health Department, Tacoma, Washington*

## Enhance Your Program's Public Image

Collectively, human services have fallen in credibility in the public eye. This loss of favor is expressed in many ways, but the fundamental complaint seems to be, "Too much money for too few results." Outcome findings can help programs separate themselves from this general indictment and demonstrate that they are in fact making a difference

**V**ocational Guidance Services of Cleveland, Ohio uses a simple but effective method of communicating the agency's progress to a broad variety of constituents. The agency publishes a quarterly report using a pictographic design style to summarize outcomes related to four major goals. Illustrations show, for example, the number of persons with severe disabilities who have obtained jobs and the economic benefit for the greater Cleveland community.

The agency mails the report to more than 5,000 individuals and organizations, including board, volunteers, local media, state politicians, referral sources, customers, and employers that hire disabled and disadvantaged people placed by the agency. Vocational Guidance Services, who places over 2,000 challenged people into jobs each year, believes the tool is effective because readers can understand their messages instantly. Response from recipients is uniformly positive, with unsolicited comments offered to the local media by audiences as diverse as potential volunteers and board members.

Public dissemination of outcome findings must be done with care. Although having 40 percent of participants achieve a desired outcome may be excellent for some types of programs, a public looking for a 90 percent success rate may regard 40 percent as a failure. Programs can provide a context for outcome findings by including comparative data, as discussed earlier in this step. Testing possible outcome-based messages on a sample of intended recipients is highly advisable.

## **Retain and Increase Funding**

Many human service programs are implementing outcome measurement to meet the requirements of funders. Although outcome measurement offers many other benefits for programs, its importance in funding decisions is growing as more and more funders ask for documentation of results. This is true of corporate funders, as well as private foundations and state and federal governments. Moreover, funders who do not ask for outcome information often are impressed by programs who provide it.

Programs can respond to this increased concern with accountability by including information on outcomes, measurement approaches, and findings in their funding

for individuals and families. Annual reports, newsletters, informational brochures, news releases, public service announcements, radio interviews, television sound bites, recognition breakfasts and dinners, and donation solicitation materials are among the vehicles programs can use to send messages about the benefits their participants are experiencing.

"Effectiveness data is wonderful for communicating with funding sources and the community at large. We are able to say, 'This program works.' We use the information in speaking engagements, in flyers distributed at fairs, and in our direct-mail campaign literature. We can point to specific areas where we have made a difference. It has led to good PR, and we have seen more contributions from corporations as well as from new contributors."

*Susan Getman, Executive Director, Big Brothers Big Sisters of Salem County Inc., Salem, New Jersey*

"Fundors are impressed. Not many nonprofits are doing outcome measurement."

*Kay Coffin,  
Executive Director,  
Big Brothers Big Sisters  
of Fresno County,  
Fresno, California*

requests and progress reports. Caveats given above about including comparison figures in messages to the public and in Step 6 about including explanatory information in reports are important considerations in reports to funders. Programs can increase the effectiveness of their communications by using terminology matched to that of individual funders. Reports to the program's board may talk of "outcomes," for example, while reports for some funders may label the same information as "results" or "impact."

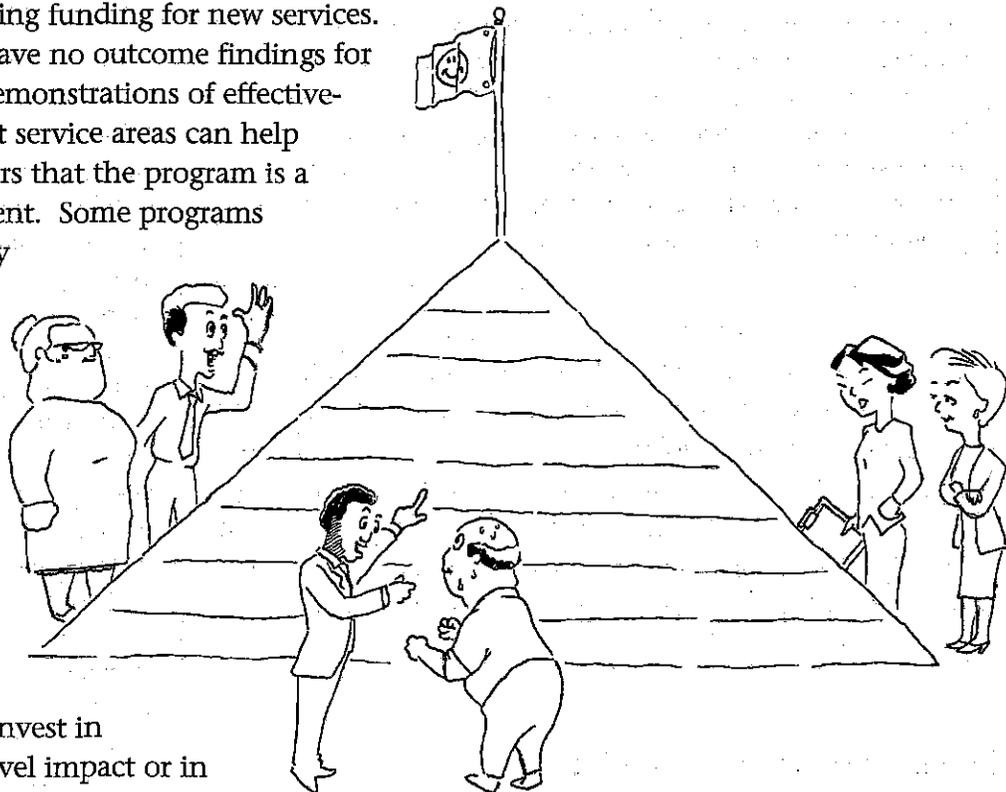
Outcome findings can strengthen the case of programs seeking funding for new services. Even if they have no outcome findings for this service, demonstrations of effectiveness in current service areas can help reassure funders that the program is a good investment. Some programs may benefit by

targeting their grant-seeking efforts on funders interested in the outcomes the program measures.

For funders

who want to invest in community-level impact or in long-range individual change, programs can use reports from

controlled, long-term studies by policy and research organizations to substantiate the link between their outcomes and the funder's larger interests.



# Appendix:

## Issues in Developing Data Collection Instruments and Procedures

Step 4 of *Measuring Program Outcomes: A Practical Approach* provides general guidance on preparing instruments and developing procedures for collecting data on program outcomes. This appendix gives more detailed information on designing three types of instruments—record extraction forms, questionnaires, and trained observer rating scales—and establishing procedures for their use.

### Record Extraction Forms

Programs use record extraction forms to collect information from their own written records or those of other agencies. This section of the appendix first discusses development of the form itself and then describes issues to be addressed in data collection procedures related to record reviews.

#### Instrument Development

Information from your own records or records of another program that you will use in measuring outcomes needs to be copied (“extracted”) from those records. Information to extract includes:

- Data related to the outcome indicator(s).
- Data on factors that may influence participant outcomes (for instance, age, sex, race, length of time in program, date of program completion, level of problem severity at intake; see Step 3).
- Data needed to link the record with other data about the participant, if applicable (for instance, the social security number or case number).

If records of your program and others involved are computerized, extraction can be done electronically. If records are not computerized, you will need a record extraction form that provides a place for recording each data item from each record of interest. Exhibit 1 is a sample record extraction form for youths enrolled in a General Equivalency Diploma (GED) program.



With data from the form, the program can measure:

- The number of youth who completed the program (all those who passed the GED exam the first or second time).
- The percent of youth who completed the program (the number of youths who obtained a GED divided by the number who entered the program).

The program can break out the outcome data by sex, age, race or ethnicity, and, for those who passed the GED exam, by the length of time from the date the youth first entered the program to the date of the successful exam. The data also allow the program to look at the relationship between the entry level score (the level of difficulty) and the outcome. Did more of the youth with lower pretest GED scores leave the program prematurely? Did it take longer for students with lower pretest scores to obtain a GED? If so, how much longer?

Not all data extracted from records will be quantitative (numerical) or easily categorized. However, qualitative types of data may be very useful. For example, case records may include case notes on participant progress. A procedure can be created whereby certain statements or phrases in the case records can be categorized to represent certain categories of outcomes. For example, a mentoring program could develop a procedure that categorizes youth-and-mentor matches as "positive," "problematic," or "unclear," based on notes in the case file on the mentors' comments. When using such qualitative ratings, it often is a good idea to have a second rater review a sample of the case records to ensure that the judgments are clear and consistent.

### **Procedural Issues Specific to Record Extraction**

In addition to the procedural issue of insuring consistency among raters when extracting qualitative data, there are procedural concerns in record extraction when reviewing records from other agencies. Procedures for extracting data from another program's records need to accommodate their requirements. Talk with staff from that program about issues related to accessing the data. Topics to discuss include:

- Are there any restrictions on the information they will release? If you were counting on obtaining data they cannot release, you will need to go back to your outcome measurement framework to identify an alternative source, or perhaps to modify the indicator or even the outcome you will measure.
- Do your definitions of key terms coincide with theirs? If not, be sure you know their terms or data categories for the information you want.

- Are their records computerized? If so, what are the procedures for obtaining an electronic file with the needed information? You may want to develop an inter-agency agreement with another program regarding sharing of confidential data.
- If the extraction must be done manually, who will be allowed to review the records and copy the information? It may be that a member of their staff will have to do this. Or perhaps a member of your staff can have access, but not a volunteer or student intern.
- What information is available to confirm that someone in their records with the same name as your participant is in fact the same person?

## **Questionnaires**

Questionnaires are the most common type of data collection instrument. Programs can use them to gain data from participants, staff, and others who know and work with participants, such as their parents or employers. Programs also can use questionnaires to obtain information from neighborhood or community residents or from target populations that the program has served on a general, rather than a one-by-one, basis.

## **Instrument Development**

There are a number of critical issues in the construction of effective questionnaires. The first consideration is content—the information the questionnaire seeks to elicit. Also important in obtaining useful information is how questions are worded and questionnaire format and style.

### **Content**

Begin the task of developing a questionnaire by referring to your outcome measurement framework to identify the specific data it is to obtain. This will include information on one or more outcomes; it also may include factors that may influence participant outcomes. Questions might ask about:

- Participants' needs, expectations, knowledge, attitudes, self-perception, skill level, behavior, life situation, status, experiences in the program, and results of those experiences. Questions on these topics can be asked of participants themselves and of others in a position to comment, such as family members, counselors, and employers.

- The amount and type of service received. Although this information is probably available from program case records, the program may want to obtain such information from the participant. The participant's perception of the services received may provide useful information to the program and will avoid requiring the program to link two different sources of information.
- Why respondents gave particular answers or ratings. Ask especially for explanations about poor ratings. Keep in mind, however, that questions requiring more than a check mark or a few words to answer take more time to collect and more time to analyze.
- Demographic and other information, such as age, sex, income, race/ethnicity, and location, needed to break out the outcome findings. Some respondents are uncomfortable answering some demographic questions. To increase the likelihood that they will answer the other questions, put demographic questions at the end of the form and explain briefly why the data are needed.

---

**Note:** In addition to reporting on outcomes, program participants often are asked to provide information and feedback on aspects of the quality of the program which may effect the achievement of outcomes. These can include:

- |   |   |
|---|---|
| ■ Awareness of program services                   | ■ Condition of facilities participants see    |
| ■ Timeliness of service                           | ■ Convenience/accessibility of location       |
| ■ Convenience/accessibility of hours of operation | ■ Ease of reaching someone to talk to         |
| ■ Accuracy of help                                | ■ Ratings of specific service characteristics |
| ■ Pleasantness/friendliness of staff              | ■ Overall satisfaction with program           |
|   | ■ Suggestions for improving service           |

Participants also can be asked about their needs, to help with needs assessment, and about what additional help they would like. Such information may not be directly related to outcome measurement, but it is a side benefit of participant surveys. Be careful, however, that you do not lengthen the questionnaire so much that you discourage participants from responding.

---

Exhibit 2 illustrates questions that can be included in participant questionnaires. These questions are from questionnaires that the Family Service Association of America used to assess the outcomes of family counseling services.

The first question will provide information on an *outcome indicator*: number and percent of participants who report improved outcomes after receiving services. The second question illustrates that participant surveys also can be used to obtain information on the participant's perception of the extent to which the services *affected the outcomes* that the participant reports. A program can combine the data from responses to the first

and second questions to calculate an additional outcome indicator: number and percent of participants that indicate improved outcomes *and* report that the services they received contributed to that improvement. As noted in the Introduction and Step 1 of this manual, outcome information generally does not tell *why* the outcomes occurred. If such questions as this one are included in participant surveys, the program can obtain some information on the "whys."

## Exhibit 2

### Survey Questions on Family Counseling

1. Since you started at the agency, has there been any change for better or worse in the way the members of your family get along with each other? Would you say you now get along:

- Much better
- Somewhat better
- The same
- Somewhat worse\*
- Much worse\*

\* Please explain: \_\_\_\_\_

2. How do you feel the service provided by the agency influenced the changes you have reported?

- Helped a great deal
- Helped somewhat
- Made no difference
- Made things somewhat worse\*
- Made things much worse\*

\* Please explain: \_\_\_\_\_

3. How satisfied were you with the way you and your counselor got along with each other?

- Very satisfied
- Somewhat satisfied
- No particular feelings one way or the other
- Somewhat dissatisfied\*
- Very dissatisfied\*

\* Please tell us why you felt this way: \_\_\_\_\_

The third question will provide information on one aspect of service that may affect whether the outcome occurred: satisfaction with the participant's counselor.

Note that if participants give unfavorable ratings in response to a question, they are asked to explain "why." Summaries of participants' responses should be made available to program personnel, keeping individual responses anonymous. Such information can provide guidance on how the program can be improved, especially to the extent that particular problems are reported by many participants.

You also might ask "why" of respondents who give favorable ratings to questions. This would be particularly useful for programs where the specific service activity varies among participants, such as camping, mentoring, or scouting.

## Wording

Questions must be clear, unambiguous, and neutral in their presentation. It is surprisingly easy to use biased or unclear wording. For instance, questions that provide four response categories, of which three are favorable and one is unfavorable, bias the response in a favorable direction. The question, "*When* do you experience this event?" may mean "at what time" or "under what circumstances." Be specific.

Most questions should ask respondents to check or circle the appropriate response from a fixed set of answers (i.e., *fixed-choice* questions), rather than asking them to compose an answer (i.e., *open-ended* questions). A self-administered questionnaire might include two or three open-ended questions. Interviews can include a few more. In either case, open-ended questions are harder to analyze. They also take longer to complete. This may discourage respondents from finishing the survey, although it also may encourage responses from people who have strong reactions about the program. Exhibit 3 gives examples of fixed-choice response sets.

To write effective questions:

- Keep each question short. Use simple sentence structures.
- Use basic vocabulary.
- Ask only one question in each question. For example, "In disciplining your child, do you state the rule clearly and explain the consequences of breaking the rule?" is two questions.

Exhibit 3

**Examples of Fixed-choice Response Sets**

Question Type	Examples
Yes or No	Have you used this new skill in the past week? <input type="checkbox"/> Yes <input type="checkbox"/> No
Agree/disagree scale	<p>How much do you agree or disagree with the following statement:</p> <p>"The public service announcements about the effect of second-hand smoke gave a very powerful argument against smoking around small children."</p> <p><input type="checkbox"/> Strongly agree                      <input type="checkbox"/> Disagree  <input type="checkbox"/> Agree                                      <input type="checkbox"/> Strongly disagree</p>
Frequency scale	<p>How often do you read to your child?</p> <p><input type="checkbox"/> At least once a week                      <input type="checkbox"/> Once a month or less  <input type="checkbox"/> At least twice a month                      <input type="checkbox"/> Never</p> <hr/> <p>Last week, how many evenings did you read a bedtime story to your child(ren)?</p> <p><input type="checkbox"/> 1   <input type="checkbox"/> 2   <input type="checkbox"/> 3   <input type="checkbox"/> 4   <input type="checkbox"/> 5   <input type="checkbox"/> 6   <input type="checkbox"/> 7   <input type="checkbox"/> None</p> <hr/> <p>What percentage of your income do you put into a savings account?</p> <p><input type="checkbox"/> None    <input type="checkbox"/> 6 to 10 percent  <input type="checkbox"/> 1 to 5 percent                                      <input type="checkbox"/> More than 10 percent</p>
Comparative response scale	<p>Compared to other programs that your child(ren) take part in, how important is the after-school tutoring program?</p> <p><input type="checkbox"/> Very important   <input type="checkbox"/> Somewhat important   <input type="checkbox"/> Unimportant</p>
Identification response	<p>Below is a list of services provided by our program. Please check the ones your child participated in since the beginning of the year:</p> <p><input type="checkbox"/> After-school program                      <input type="checkbox"/> Sports program  <input type="checkbox"/> Tutoring program                              <input type="checkbox"/> Summer camp</p>

Adapted from: Irwin Epstein and Tony Tripodi. *Research Techniques for Program Planning*. © 1977 by Columbia University Press. Used with permission of the publisher.

- Think carefully about words and phrases that may have double meanings. This is especially true if the question writer is of a different age, ethnic or cultural background, or educational level, or is from a different part of the country than the respondents. (Exhibit 4-D in Step 4 describes some cultural considerations in questionnaire development.)
- Make sure nothing in the question, including the response choices, suggests a “right” answer.
- Be specific about time frames. For instance, say “within the past three months” or “since March 1” instead of “recently.”
- Make response categories as specific as possible. Choices like “excellent; good; fair; poor” and “frequently; infrequently; rarely” are subject to wide interpretation. An alternative frequency scale might be “every day, once a month, fewer than 12 times a year.” An alternative qualitative scale might be “have used the skill already; haven’t used it but am confident I can; not sure I understand how to do this; don’t understand how to do this.”
- Read questions aloud. This practice will pinpoint some awkward phrases and unclear meanings.
- Have someone experienced in questionnaire development review the questionnaire before it is finalized.

### Format and Style

Once the questions are written, you are ready to arrange and format your questionnaire. Make it clear and easy to answer. Here are some suggestions on format and style, especially for self-administered questionnaires:

- Group related questions together. Start with the least-personal and most obviously relevant set of questions.
- Be sure instructions are short and explicit.
- If you must use *skip patterns* (e.g., if respondents answer “no” to a question about whether they have children and you want them to skip the next three questions because they concern children), be sure the instructions are easy to follow. Consider using arrows to lead visually to the appropriate next question. As an alternative to skipping questions, consider adding a “not applicable” response. For example, in the

three questions about children, add a response category that says “not applicable; no children.”

- Make the questionnaire easy to read. Use a clear typeface. Leave plenty of “white space.” Use left justification only. If interviewers or respondents are older persons, use a larger typeface.
- If you plan to use colored paper, pick a color that photocopied well and is easy on the eyes.

If substantial numbers of persons with limited English proficiency will be respondents, translate the questionnaire into their native language. Then, test the accuracy of the translation by having a second person translate the foreign-language version back into English to see if the questions still say the same thing.

## **Procedural Issues Specific to Questionnaires**

Step 4 of this manual summarized several general issues related to data collection procedures. Collection of data via questionnaires involves a number of additional issues, discussed below.

### **Administration**

The primary methods of administering questionnaires are interviews and self-administration. Generally, self-administered questionnaires are either given to respondents at the program site or mailed to their homes. Interviews generally are administered either by telephone or in person. In-person interviews can be conducted at the respondent’s home, at the program site, or at another accessible facility. Considerations related to these options are as follows.

- **Mail.** Mail surveys usually are inexpensive; however, they require special procedures to obtain acceptable response rates. Mailed questionnaires need to be short and simple. They are not useful for respondents with very low levels of literacy. For those with difficulty understanding English, the questionnaire can be translated into other languages. Note that many communities with non-English speaking residents have a diverse ethnic make-up and translation into only one additional language may not suffice. Getting adequate response rates can be a major problem with mailed questionnaires. You will need multiple mailings, and possibly telephone follow-ups, to encourage a response or to conduct a phone interview. Additional suggestions are given later in this appendix.

- **Telephone.** Telephone surveys are a less expensive alternative than in-person interviews and can get good response rates. However, they require considerable interviewer time, and interviewers need training to conduct telephone interviews. If the agency has staff or volunteers that have time to do telephone interviews, this option may be a very good one. Note, however, that to avoid biasing the answers and to ensure treating participants ethically, the interviewers should not be the same persons that provided services to the participants. If a substantial percentage of program participants do not have a telephone, then clearly this is not a good data collection mode for the program.

Surveys focused on low-income populations do need to seek responses from adequate numbers of those populations. One possibility is to conduct a telephone survey and then conduct in-person interviews in those neighborhoods known to have few telephones. If substantial numbers of persons speaking limited English are expected, bilingual interviewers will be needed.

- **In person at the person's home.** Generally, this is quite expensive and not likely to be feasible for regular data collection. However, it is sometimes the best option for elderly homebound participants and other special populations. For instance, if you are interested in obtaining information about experiences with attendant care provided to elderly persons and persons with disabilities, you might consider in-home interviews with a sample of attendant-care recipients.
- **In person at the agency or other readily accessible facility.** Usually, this is quite inexpensive but not useful for obtaining post-service information unless the participant is coming for other services (for instance, parents who have completed a parenting class but still need to see their case worker on a regular basis). The survey can be administered through an interview or by asking the respondent to check off the answers to a written questionnaire.
- **Combinations of the above.** For example, a meals-on-wheels program could have drivers distribute post card questionnaires to new recipients, then use interns or volunteers to tabulate responses and do telephone follow-up with recipients who do not respond.

Key concerns in selecting a method of administration include response rates, accuracy of responses, and cost. Choosing a method involves trade-offs among these concerns. Exhibit 4 provides a summary comparison of the various methods of questionnaire administration.

Most programs will find that mailing questionnaires will be the most practical approach because of its low cost. Programs that undertake second mailings and possibly telephone reminders to achieve return rates of 50 percent or more should achieve

sufficiently accurate data for outcome measurement. However, if the program has a cadre of staff, volunteers, or students it can depend on to undertake regular telephone interviews at a low cost to the program, telephone interviewing would be the preferred method.

For programs that are seeking data on outcomes which occur early in the service delivery process or immediately thereafter, and/or where the program does not obtain the names or addresses of participants, then in-person administration at the location where the service is delivered probably would be the appropriate method of questionnaire administration. For example, this method might be appropriate to obtain data on the knowledge gained in a parenting education workshop or skills obtained in a CPR course.

In circumstances where a program can reach participants only in their own environment, in-person interviewing there may be necessary. For example, to reach homeless individuals who do not come into shelters, a program providing homeless services might need to attempt interviews on the street. Another example is that some participants may be home-bound, without a telephone, and unlikely to be able to respond to a mailed questionnaire. At-home interviews might then be the only option for interviewing those participants.

Exhibit 4

### Summary Comparison of Questionnaire Administration Options

Characteristic	Method of Questionnaire Administration			
	Mail	Telephone	In Person at Home	In Person at Facility
Response rate	Low/Medium*	High	High	High
Amount of information obtainable	Low	Medium	High	Medium
Ability to obtain post-service information	High	High	High	Low
Speed in getting responses	Slow	Fast	Medium	Fast
Cost	Low	Medium	High	Low/Medium

\*Low, if only one mailing. Medium, if multiple mailings and possibly telephone follow-ups are done.

## **Voluntary Completion**

Answering questionnaires, regardless of how they are administered, must be voluntary. Persons asked to respond have the right to refuse and must not be pressured to respond. Programs need to be certain that declining to answer in no way penalizes participants in receiving future services. Further, participants should be told explicitly that choosing not to respond will not affect their eligibility for services.

## **Participant Confidentiality**

Depending on the type of program you operate, the fact that the participant obtained services from your program may be confidential. One cannot, for example, call someone's home and leave a message asking the person to call her or his former substance abuse or mental health counselor. For programs in which this is an issue, telephone interviewers must be sure to reach the participant directly; mail questionnaires probably are not appropriate. For other programs, such as child care or youth development, confidentiality may be less of an issue.

## **Anonymity**

If participants are interviewed in person, their responses can be kept confidential, but the participants are not anonymous because the interviewers know who they are. This generally is true of telephone interviews as well. Mail questionnaires, however, can be returned without an identifying name or code, and thus be anonymous. The key advantage to anonymity is that respondents may be more honest with their responses.

There are several disadvantages. One is that the responses cannot be linked with data from program records. Therefore, all information that is needed about date of entry into the program, type and amount of services received, and other outcomes and influencing factors needs to be asked on the questionnaire. Another disadvantage is that the program cannot follow up with non-respondents because it has no way of knowing who has responded and who hasn't. This means either a lower response rate or increased costs because a second mailing needs to go to all respondents.

A third disadvantage of anonymous questionnaires is that there is no way of linking data that are collected from participants at multiple points to track changes over time. Programs sometimes try to overcome this problem by having participants put a personal code, such as their mother's maiden name, on each questionnaire they complete. The effectiveness of such strategies depends on each participant's code being unique and on participants' memory of the code they used on previous questionnaires.

## Times for Contacting Potential Interview Respondents

Decisions about days and times that potential respondents are contacted for interviews depend to a great extent on the characteristics of the respondents. Elderly and retired individuals may be called during daytime hours. Employed parents who work day shifts are best contacted in the evening. Avoid religious holidays and those days respondents celebrate as the Sabbath. Set time limits on when contacts may be made on other days (e.g., not before 10:00 a.m., during the dinner hour, or after 9:00 p.m.).

## Response Rate

The smaller the response rate, the greater the possibility that the sample is not representative of the population from whom the program is seeking responses. Whatever administration method you use, as long as the respondent selection method provides a representative sample, aim to get completed questionnaires from at least 50 percent of those from whom you are seeking responses. (The federal government seeks a 75 percent rate for federal surveys, but such a target is likely to be too costly for regular surveys by most private, non-profit agencies.) With mail surveys, second and third mailings or telephone reminders to non-respondents usually will be needed to obtain response rates high enough to provide reliable information. For telephone surveys, multiple calls at various times of the week are likely to be needed to reach respondents. Exhibit 5 gives tips for increasing response rates on mail surveys.

## Costs

Here are some ideas for saving costs on questionnaire administration.

- Use already-available questionnaires.
- Use volunteer technical consultants to help design the questionnaire and procedures. Some otherwise well-paid consultants will accept a lower rate (or provide free time) when working for a non-profit organization. Although technical help *may* cost more initially, it is likely to save you the cost of collecting unreliable, unusable, or insufficient data.
- Use mail surveys (but plan on doing second and even third mailings to obtain an adequate response rate).
- If your program has several hundred participants, survey a sample rather than everyone.
- Use volunteers to conduct interviews. Ask for help from a local university or college.

## Exhibit 5

### **Tips for Improving Mail Survey Response Rates**

1. Establish a procedure to inform new participants about the survey process and encourage them to respond if and when they are surveyed later.
2. Make questionnaire as short and easy to use as possible.
3. Respondents are more likely to complete and return questionnaires that look attractive and simple. Go for a professional, polished look—even if you produce the questionnaire in-house. Use a printer or high-quality copier and good-quality paper. Check for spots, blurs, and crooked copies. If using colored paper, pick a color that is easy on the eyes. Avoid expensive-looking paper, two-color printing, or other features that may give the impressions you “wasted” money on production. If the questionnaire is to be completed by children, make it as much fun as possible to do.
4. To the extent possible, avoid surveying during vacation and holiday seasons.
5. Use an advance post card to notify participants that they will be receiving a mailed questionnaire shortly.
6. Include a carefully worded transmittal letter signed by an agency official, such as the executive director or board chair, or a notable community figure encouraging response. Emphasize the agency’s need for the information to improve future services. Explain that each question is important. Guarantee that responses are confidential and will not be attributed to individuals in any way.
7. If possible, personalize the letter and address it to a specific individual.
8. Consider offering inexpensive incentives for responding, such as certificates for a fast-food store drink.
9. Enclose a postage paid, pre-addressed return envelope with the questionnaire at each mailing.
10. Send post card reminders that thank those who already have responded and encourage others to complete and return the questionnaire.
11. Re-mail the questionnaire if the response rate is too low.

## Trained Observer Ratings

One of the values of trained observer ratings is that they are more objective than self-reports and can be more consistent than casual observations. Achieving a high degree of accuracy and reliability in observation ratings depends on having clearly defined rating categories used by carefully trained observers. The more observers that are involved, the more critical the rating scales are.

### Instrument Development

Rating systems used by trained observers include:

- **Written descriptions.** This is the most basic type of rating system. It relies solely on written descriptions of each grade in the rating scale. These descriptions must be quite specific about what behaviors or characteristics merit each rating. This is necessary so that different raters give approximately the same rating for a particular condition. Exhibit 6 provides an abbreviated example of written guidelines for the condition of a playground connected to a child care program in an inner-city neighborhood.
- **Photographic rating systems.** These systems use pre-selected photos to represent the various grades on the rating scale. Photographic scales are one of the most useful forms of rating systems. They have been used to assess the cleanliness of streets, the condition of parks and playgrounds, and external housing conditions.
- **Other visual scales.** Drawings, sketches, video clips, or other visual alternatives to photographs can be used to depict each grade on a rating scale.

### Procedural Issues Specific to Trained Observer Ratings

Achieving a high degree of accuracy and reliability in observation ratings depends on having:

- Clearly defined rating categories.
- Adequate training and supervision of the observers and the process.
- A procedure for checking the quality of the ratings periodically. The goal is for different observers, perhaps at various times, to give the same ratings to comparable conditions.

Observers can be program agency personnel, college or graduate school students, volunteers, or community members. More technical ratings will require persons with more professional training. Avoid having staff or volunteers observe persons or places with whom they work.

Exhibit 6

**Sample Rating Scale for Trained Observer Rating  
of a Playground for Preschool Children**

Rating 1: **Clean.** The area is completely, or almost completely, clean; up to two pieces of paper litter are permitted. None of the litter is broken glass, drug paraphernalia, or other dangerous materials.

Rating 2: **Moderately clean.** The area is largely clean; a few pieces of isolated litter are observable. None of the litter is broken glass, drug paraphernalia, or other dangerous materials.

Rating 3: **Heavily littered.** Scattered litter in several areas and/or blown all along the fence. However, none of the litter is broken glass, drug paraphernalia, or other dangerous materials.

Rating 4: **Dangerously littered.** The litter includes broken glass, drug paraphernalia, or other dangerous materials. The area therefore is not safe for children to use.

Training of observers is important. For behavioral ratings, videotaped examples of various levels of the behavior often are effective. Training should include opportunities to practice applying the rating scales and then compare and discuss the ratings with other observers.

Establish procedures for supervising observers, recording the data they collect, and transcribing and processing the data. Also establish procedures for systematically checking the ratings to assure quality control and reliability of the process. For example, the supervisor might check a small sample of the ratings done by each observer.



# Bibliography

## Performance Measurement

*The Accountable Agency.* R.K. Carter. Thousand Oaks, CA: Sage Publications, 1983.

*Building Strong Foundations: Evaluation Strategies for Family Resource Programs.* (Knowledge Transfer Project, Part 3) J.H. Littell. Chicago, IL: Family Resource Coalition, 1986.

*The Case for Performance Monitoring.* J. Wholey & H.P. Hatry. In *Public Administration Review*, vol. 52, no. 6. Washington, DC: American Society for Public Administration, November/December 1992.

*Cultural Influences on Research Methods and Statistics.* D. Matsumoto. Pacific Grove, CA: Brooks/Cole, 1994.

*Developing Client Outcome Monitoring Systems: A Guide for State and Local Social Services Agencies.* R. Millar & A. Millar (editors). Washington, DC: The Urban Institute, 1981.

*Eleven Ways to Make Performance Measurement More Useful to Public Managers.* H. Hatry, C. Gerhart, & M. Marshall. In *Public Management*, September 1994.

*Emerging Strategies in Evaluating Child and Family Services.* In *The Evaluation Exchange*, vol. 11, no. 1. Cambridge, MA: Harvard Family Research Project, Winter 1996.

*Evaluating Your Agency's Programs.* M.J. Austin, G. Cox, N. Gottlieb, J.D. Hawkins, J.M. Kruzich, & R. Rauch. Thousand Oaks, CA: Sage Publications, 1989.

*Evaluation: A Systematic Approach.* P.H. Rossi & H.E. Freeman. Thousand Oaks, CA: Sage Publications, 1987.

*Evaluation Research.* C. Weiss. Old Tappan, NJ: Prentice-Hall, 1972.

*How to Conduct A Client Follow-Up Study* (2nd ed.). D.F. Beck & M.A. Jones. New York, NY: Family Service Association of America, 1980.

*How to Use Qualitative Methods in Evaluation.* M.Q. Patton. Thousand Oaks, CA: Sage Publications, 1987.

*Measuring the Performance of Human Service Programs.* L.L. Martin & P.M. Kettner.  
Thousand Oaks, CA: Sage Publications, 1996.

*Misused Statistics: Straight Talk for Twisted Numbers.* A.J. Jaffe & N.F. Spierer. New York,  
NY: Marcel Dekker, 1987.

*Program Evaluation Kit.* (nine-volume series on planning and conducting program  
evaluation) J.L. Herman (series editor). Thousand Oaks, CA: Sage Publications, 1991.

*Say It With Charts: The Executive's Guide to Successful Presentations in the 1990's* (2nd  
ed.). G. Zelazny. Irwin, IL: Business One, 1991.

*Service Efforts and Accomplishments Reporting--Its Time Has Come: An Overview.* H.P.  
Hatry. Norwalk, CT: Governmental Accounting Standards Board, 1990.

*Social Services: What Happens to the Clients.* M. Koss, H. Hatry, A. Millar, T. van Houten,  
K. Craddock, W. Bright, G. Burlean, & J. Reardon. Washington, DC: The Urban  
Institute, 1979.

*Using Performance Measurement in Local Government.* P.D. Epstein. New York, NY: Van  
Nostrand Reinhold Co., 1984.

*Utilization-Focused Evaluation.* M.Q. Patton. Thousand Oaks, CA: Sage Publications, 1986.

## **Data Collection**

*Asking Questions: A Practical Guide to Questionnaire Design.* S. Sudman & N.M.  
Bradburn. San Francisco, CA: Jossey-Bass, 1982.

*Citizen Surveys: A Special Report on Designing, Conducting, and Understanding Citizen  
Surveys.* T.I. Miller & M.A. Miller. Washington, DC: International City/County  
Management Association, 1991.

*Developing and Using Questionnaires.* Washington, DC: U.S. General Accounting Office,  
GAO/PEMD-10.1.7., October 1993.

*Focus Groups: A Practical Guide for Applied Research.* R.A. Krueger. Thousand Oaks, CA:  
Sage Publications, 1988.

- Handbook for Excellence in Focus Group Research.* M. Debus. Washington, DC: Academy for Educational Development, 1988.
- How to Conduct a Citizen Survey.* Planning Advisory Service Report Number 40. T. van Houten & H.P. Hatry. Chicago, IL: The American Planning Association, 1987.
- How to Conduct Interviews by Telephone and In Person.* J.H. Frey & S.M. Oishi. Thousand Oaks, CA: Sage Publications, 1995.
- How to Conduct Surveys: A Step-by-Step Guide.* A. Fink & J. Kosecoff. Thousand Oaks, CA: Sage Publications, 1985.
- How to Measure Attitudes.* M.E. Henerson, L.L. Morris, & C.T. Fitz-Gibbon. Thousand Oaks, CA: Sage Publications, 1987.
- Improving Survey Questions.* F.J. Fowler, Jr. Thousand Oaks, CA: Sage Publications, 1995.
- Introduction to Survey Sampling.* G. Kalton. Thousand Oaks, CA: Sage Publications, 1983.
- Measuring Client Satisfaction: Developing and Implementing Good Client Satisfaction Measurement and Monitoring Practices.* Ottawa, Canada: Office of the Comptroller General, Evaluation and Audit Branch, August 1992.
- Research Methods for Social Work* (2nd ed.). A. Rubin & E. Babbie. Pacific Grove, CA: Brooks/Cole, 1992. (see Chapter 7, Constructing Measurement Instruments, and Chapter 8, The Logic of Sampling)
- The Survey Kit.* (nine-volume series on conducting and analyzing surveys) A. Fink (series editor). Thousand Oaks, CA: Sage Publications, 1995.
- Survey Questions. Handcrafting the Standardized Questionnaire.* J.M. Converse & S. Presser. Thousand Oaks, CA: Sage Publications, 1986.
- Telephone Survey Methods: Sampling, Selection, and Supervision* (2nd ed.). P.J. Lavrakas. Thousand Oaks, CA: Sage Publications, 1993.
- Using Structured Interviewing Techniques.* Washington, DC: U.S. General Accounting Office, GACO/PEMD-10.1.5., October 1993.
- Your Guide to Measuring Client Satisfaction.* Ottawa, Canada: Office of the Comptroller General, Evaluation and Audit Branch, Communications and Coordination Directorate, 1992.

## Performance Indicators

*Assessing Drug Abuse Among Adolescents and Adults: Standardized Instruments.* Clinical Report Series. Rockville, MD: National Institute on Drug Abuse, Public Health Service, U.S. Department of Health and Human Services, 1994.

*Assess for Success: Needs Assessment and Evaluation Guide.* New York, NY: Girls Incorporated, 1991.

*ATOD Prevention Program Outcomes and Instrument Selection System.* D.C. Gottfredson & M.A. Harmon. Washington, DC: Center for Substance Abuse Prevention, U.S. Department of Health and Human Services. Draft, May 26, 1995.

*Evaluating Family Programs.* H.B. Weiss & F.H. Jacobs. New York, NY: Aldine de Gruyter, 1988.

*Family Assessment: A Guide to Methods and Measures.* H.D. Grotevant & C.I. Carlson. New York, NY: Guilford, 1989.

*Handbook of Family Measurement Techniques.* J. Touliatos, B.F. Perlmutter, & M.A. Straus (editors). Thousand Oaks, CA: Sage Publications, 1990.

*Handbook of Research Design and Social Measurement (2nd ed.).* D.C. Miller. Thousand Oaks, CA: Sage Publications, 1991.

*KIDS COUNT Data Book: State Profiles of Child Well-Being.* Baltimore, MD: Annie E. Casey Foundation, 1995.

*Maternal Substance Use Assessment Methods Reference Manual: A Review of Screening and Clinical Assessment Instruments for Examining Maternal Use of Alcohol, Tobacco, and Other Drugs.* Washington, DC: Center for Substance Abuse Prevention, U.S. Department of Health and Human Services, 1993.

*Measurement in Prevention: A Manual on Selecting and Using Instruments to Evaluate Prevention Programs.* K.L. Kumfer, G.H. Shur, J.G. Ross, K.K. Bunnell, J.J. Librett, & A.R. Millward. Washington, DC: Center for Substance Abuse Prevention, U.S. Department of Health and Human Services, 1993.

*Measures of Personality and Social Psychological Attitudes.* J.P. Robinson, P.R. Shaver, & L.S. Wrightsman. San Diego, CA: Academic Press, 1991.

- Measuring for Clinical Practice.* K. Corcoran & J. Fischer. New York, NY: The Free Press, A Division of Macmillan, 1987.
- Mental Measurements Yearbook* (12th ed.). J.C. Conoly and J.C. Impara (editors). Lincoln, NE: Buros Institute of Mental Measurements, 1995.
- The Odyssey Project: A Descriptive and Prospective Study of Children and Youth in Residential Treatment, Group Homes, and Therapeutic Foster Care.* Washington, DC: Child Welfare League of America Research and Evaluation Staff and The Odyssey Project Consortium, August 1995.
- Outcome Based Performance Measures.* Towson, MD: The Accreditation Council on Services for People with Disabilities, 1993.
- Outcome Based Performance Measures: A Procedures Manual.* Towson, MD: The Accreditation Council on Services for People with Disabilities, 1995.
- Outcome Measures for Child Welfare Services.* S. Magura & B. Moses. Washington, DC: Child Welfare League of America, 1986.
- Performance Measurement in Selected Public Health Programs.* Washington, DC: Department of Health and Human Services, Public Health Service, Report for 1995-1996 Regional Meetings.
- Program Evaluation: Guidelines for Development and Implementation.* Bethesda, MD: Goodwill Industries International, 1994.

...and the ... of ...

...and the ... of ...