
Network Guide to Measuring Family Development Outcomes

**Brad Richardson, Ph.D.
Julie Spears, M.S.W., M.A.
Bill Thiesen, Ph.D.**

December 2003

**University of Iowa School of Social Work
National Resource Center for Family Centered Practice**

Acknowledgements

The authors wish to thank the Board of Directors of the Iowa Community Action Association, Julie Jakopic, Glenn Kamber, and Arnie Anderson for their reviews and input on earlier drafts, and the contribution of material from earlier work by Professor Miriam Landsman, Ph.D. and Judy McRoberts, M.S.W., of the University of Iowa School of Social Work and National Resource Center for Family Centered Practice. We would also like to thank Mark Dickey, Executive Director of the Iowa Community Action Association who served as project director and source of support in completion of the project; Nancy Graf of the University of Iowa School of Social Work, National Resource Center for Family Centered Practice; and the Department of Health and Human Services, Office of Community Services and our project officer Brandy RaNor-Hill who provided funding for the project under Grant No. OCS #90ET0241/01.

We dedicate this work to the memory of Rollie Grady who managed the Integrated Family Collection and Analysis Program known as IFCAP upon which the AAFP was founded and who brought a unique ability to bridge management information systems with the needs of its users.

This project was supported by the Office of Community Services, Department of Health and Human Services, Washington DC 20201 under Contract/Grant No. #90ET0241/01. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Department of Health and Human Services, Office of Community Services, the Iowa Community Action Association, or the University of Iowa.

Preface

This Network Guide to Measuring Family Development Outcomes serves as a review of information accumulated over the past two decades in work with the Community Services Network to define valid and reliable outcome measures. The background and history, progress with implementing Results Oriented Management and Accountability (ROMA) to date, and the most recent ROMA goals and outcome indicators are presented in Section 1, Chapter I. In Chapter II we present a review of some exemplary approaches by state and then provide Systems Assessment Criteria for use in comparing approaches. In Chapter III we present in more detail one exemplary approach to the measurement of family development outcomes: The Automated Assessment of Family Progress (AAFP).

Section 2 provides guidance on How to Develop Outcomes beginning with a review of general information related to the paradigm shift from focusing on service units to outcomes (Chapter IV). In Chapter V we examine how to make outcome evaluation efficient and effective, and then in Chapter VI explore defining and deciding on which outcomes to measure using a step-by-step approach.

Section 3 contains Outcome Development Tools which may serve as a curriculum for training in the development of outcome measures or as aids in work by agencies to determine outcome measures for family development. Chapter VII provides worksheet for use in defining outcomes followed by some case examples of their use. The final Chapters present a glossary of terms (Chapter VIII), some additional family development instruments that may be useful (Chapter IX), and Chapter X is a list of references cited throughout the document.

The Guide is designed to provide the reader with sufficient background and knowledge of ROMA to develop relevant outcomes for any family development program. The AAFP is presented in detail as one approach that could be adapted for use. Tools are then provided that can be used to train in the development of family outcome measures or aid in start up of a new system.

TABLE OF CONTENTS

<i>ITEM</i>	<i>Page</i>
Acknowledgments.....	2
Preface.....	3
Table of Contents.....	4
Short Abstract.....	5
Long Abstract.....	6
Section 1. ROMA and Its Implementation.....	8
I. Introduction.....	9
A. Background and History.....	9
B. ROMA Progress.....	10
C. ROMA Goals and Outcomes.....	12
II. Approaches to Measuring Family Development Outcomes.....	14
A. Approaches to Measuring Family Development Outcomes by States.....	14
B. Systems Assessment Criteria.....	18
C. Lessons Learned.....	20
III. Automated Assessment of Family Progress (AAFP).....	22
A. Use of AAFP Workers.....	23
B. Use of AAFP by Data Mangers and Analysts.....	23
C. Use of AAFP by Policy and Decision Makers.....	30
D. Lessons Learned.....	31
Section 2. How to Develop Outcomes.....	33
IV. The Outcome Measures Paradigm.....	34
V. Efficient and Effective Outcome Evaluation.....	44
VI. Step by Step Guide for Developing Outcomes for Family Development Programs.....	51
Section 3. Outcomes Development Tools.....	60
VII. Worksheets.....	61
VIII. Glossary.....	73
IX. List of Instruments.....	74
X. Bibliography/References.....	76
Appendix	
Appendix A: Automated Assessment of Family Progress (AAFP) Measures	82

Network Guide to Measuring Family Development Outcomes

Short Abstract

The University of Iowa School of Social Work, National Resource Center for Family Centered Practice, in collaboration with the Iowa Community Action Association and the State Division of Community Action Agencies have been involved in the design and implementation of Results Oriented Management and Accountability (ROMA) since 1995. The Network Guide documents a variety of exemplary approaches and describes how these approaches work. Lessons learned are provided for identification, implementation, use, reporting and improvement of outcomes of family development programs.

Long Abstract

At the national level, the Community Services Network (CSN) has been engaged in the design and implementation of Results Oriented Management and Accountability (ROMA) since 1994 in response to the Government Performance and Results Act (GPRA) of 1993. Along with the Management and Assessment Task Force (MATF) and the Scales and Ladders Subcommittee, the CSN in Iowa has also been engaged in the ROMA effort since 1995. Through meetings, planning sessions, collection and analysis of data, and tracking ROMA outcome measures the Iowa network has developed a system which enables service providers to better assist those who come to Community Action Agencies (CAAs) for services and also tracks outcomes. This work has also been coordinated with work in California including the reliability and validity assessment project conducted by Endres, Richardson and Sherman (1999) and the development of proposed cross validation studies. The Guide documents processes for developing and implementing ROMA outcome measures, exemplified by our work with the Automated Assessment of Family Progress (AAFP), the California Family Development Matrix (CFDM), and similar approaches outside the Community Services Network (c.f. Richardson, Graf and Huff, 2001; Kirk, et al., 2001). The Guide provides a complete method for identifying, implementing, using, improving and reporting family development outcomes.

The Network Guide to Measuring Family Development Outcomes presents approaches for collecting and analyzing results and linking services to family development and stability outcomes. This is no small task. This includes how to use the family development outcome measures in the service delivery process as part of case management for the benefit of clients, as well as a reliable outcomes measurement and data collection tool (Endres, Richardson and Sherman, 2000). This system is not simply “additional paperwork.” The next step will be to calculate return on investment (ROI, cf. Benson, 1999; Hunnemann and Richmond, 1996) for these services and use that information for institutional decision-making (Richardson and Huff, 2001). Further development of capacity and commitment to ROMA at a deeper and broader level in all CAAs is also needed. This will require making clear how family development data may be accessed and made more available, useful and relevant to CAA Boards and Directors and how they may serve to improve direct practice. Elaboration on the following issues will assist in building capacity generally, and specifically for family functioning and development outcomes:

- How to make the ROMA data useful to the CAA staff and managers
- How to make the ROMA data relevant to CAA Boards as a key ingredient in their strategic planning processes
- How to make the ROMA data accessible and available in a timely manner to CAA staff, boards and other members of the Community Action Network
- How to provide ROMA training to CAA Boards so that they can put ROMA outcome measures to use
- How to analyze ROMA outcomes and service utilization documenting their linkages
- How to increase the ease of use of outcome measures systems

- How to implement changes to ROMA outcome measures
- How to locate agency specific results in juxtaposition with statewide results
- How to report so that results are understandable and useful
- How ROMA works for OCS

The general orientation of “The Network Guide” is built on our experience with developing and implementing outcome measures outside the Community Services Network as well as within the CSN in our work on the AAFP, the California Matrix Model and with other outcome measures systems with which we have direct familiarity.

Section 1. ROMA and Its Implementation

- I. Introduction
 - A. Background and History
 - B. ROMA Progress
 - C. ROMA Goals and Outcomes
- II. Approaches to Measuring Family Development Outcomes
 - A. Approaches to Measuring Family Development Outcomes by States
 - B. Systems Assessment Criteria
 - C. Lessons Learned
- III. Automated Assessment of Family Progress (AAFP)
 - A. Use of AAFP by Workers
 - B. Use of AAFP by Data Managers and Analysts
 - C. Use of AAFP by Policy and Decision Makers
 - D. Lessons Learned

I. Introduction

Our orientation to the development of outcome measures systems is that they should be valid indicators that are integrated with best practices in working with families and aid in the improvement of work for those families. The merger between practice and outcome measurement is the value of measuring family development outcomes and defines what is needed in a guide. The most effective strategies are also able to adapt to change continuously improving and incorporating new outcomes, measures, practices and approaches into the system. The original goals for this guide were to document the following:

- How to make the ROMA data useful to the CAA staff and managers
- How to make the ROMA data relevant to CAA Boards as a key ingredient in their strategic planning processes
- How to make the ROMA data accessible and available in a timely manner to CAA staff, boards and other members of the Community Action Network
- How to provide ROMA training to CAA Boards so that they can put ROMA outcome measures to use
- How to analyze ROMA outcomes and service utilization documenting their linkages
- How to increase the ease of use of outcome measures systems
- How to implement changes to ROMA outcome measures
- How to locate agency specific results in juxtaposition with statewide results
- How to report so that results are understandable and useful
- How ROMA works for OCS

A. Background and History

The Results Oriented Management and Accountability (ROMA) approach developed largely in response to the Government Performance and Results Act (GPRA) of 1993 by the DHHS Office of Community Services (OCS). The Community Services Network (CSN) has been engaged in the design and implementation of (ROMA) since 1994. Along with the Management and Assessment Task Force (MATF) and the Scales and Ladders Subcommittee, the CSN in Iowa had been engaged in the ROMA effort since 1995. Through meetings, planning sessions, collection and analysis of data, and tracking ROMA outcome measures the Iowa network developed a system which enabled service providers to better assist those who come to Community Action Agencies for services and which also tracked outcomes. This effort became known as the Automated Assessment of Family Progress (AAFP). Also during this time we began work on the California Matrix Model which was coordinated with the work in Iowa through a reliability and validity assessment project funded by the Packard Foundation and conducted by Jerry Endres and Brad Richardson (Endres, Richardson and Sherman, 2000). Documenting the process by which these outcome measures systems developed, along with the understanding of similar approaches outside the CSN (c.f. Richardson, Graf and Huff, 2001) further the broad base upon which the Network Guide is founded. As a result, the current effort provides a state of the art method for

identifying, implementing, using, improving and reporting outcomes. We also take into account ways in which cost analysis, specifically return on investment (ROI, cf. Benson, 1999; Hunnemann and Richmond, 1996) for CAA services can be derived from the information for institutional decision-making (Richardson and Huff, 2001). Further development of capacity and commitment to ROMA at a deeper and broader level in all CAAs is also needed. This requires understanding of how family development data may be accessed and made more available, useful and relevant to CAA Boards and Directors and how they may serve to improve direct practice. We will elaborate on the goals of this guide generally and specifically as they relate to family functioning and development outcomes.

Community Action Agency Programs and their Participants

According to the CSBG Statistical Report for FY 2000, the CSBG network comprises more than 1100 local agencies, primarily CAA agencies that “create, coordinate and deliver a broad array of programs and services to low-income Americans.” Their purpose is to alleviate poverty by fighting its causes, e.g., unemployment, housing, nutrition, education, etc. (Power, Knowlton, Spade-Aguilar, 2003).

While the CSBG network serves an ethnically and racially diverse clientele, the report describes the typical CAA participant lived in a family with children, was white, non-Hispanic, was very poor, and had family members currently working or with work experience. Single mothers headed more than half the families; Nearly 40 percent of local participants were children (not including children in client families). Seventy percent of CAA participants had incomes below the current HHS poverty guideline of \$14,150 for a family of three. Fifty percent of CAA participants had incomes below 75 percent of the poverty guideline (\$10,610); Thirteen percent had incomes below 50 percent of the poverty guideline (\$7,075). TANF clients constituted only 16 percent of CAA participants. Power, Knowlton, and Spade-Aguilar (2003) note that this trend reflected the national experience of declining TANF caseloads. It also reflects a shift from CAA participants who are also TANF clients to CAA participants who rely on wages from employment, however low.

Data from the 2000 U.S. Census indicates that 9.2 percent of families and 12.4 percent of individuals are in poverty. These figures rise dramatically, however, when looking at children and families with single mothers. For single female headed households, more than one in four families (26.5%) are in poverty. Almost half of these families (46.4%) have children under 5 years old.

B. ROMA Progress

Since 1994 more than 935 public and private CAA have installed performance based management and evaluation processes associated with the Office of Community Services (OCS) six national goals. The Monitoring and Assessment Task Force (MATF) defined outcomes that 1) focused community action on helping people and families out of poverty; 2) mobilized programs to address multiple causes and conditions of poverty; and 3) improved capacity of community action to measure and

report improvements that resulted from its efforts. One important part of this process was the development of tools to help agencies monitor and assess their agencies, communities and low-income families in their efforts toward achieve the *national goals*:

1. Low-income people become more self-sufficient.
2. Conditions in which low-income people live are improved.
3. Low-income people own a stake in their community.
4. Partnerships among supporters and providers of services to low-income people are achieved.
5. Agencies increase their capacity to achieve results.
6. Low-income people, especially vulnerable populations, achieve their potential by strengthening family and other support systems.

OCS encouraged local agencies and state systems to develop models and processes for defining variables and measuring outcomes. This approach encouraged useful and innovative approaches to collecting and organizing outcome data utilizing a “bottom up” organizational strategy. As a result CAAs and states have developed many different models, measure different variables, measure similar variables in different ways and collect a variety of data regarding self-sufficiency and family strengthening.

The National Association for State Community Service Programs (NASCS) issued Performance Measurement Reports for FY 2001 (Kamber, February 2003). This document described information submitted by states in narrative form and with a data section. States were surveyed about the outcome data collected on participant, agency and community outcomes for the six national goals. More than 98 percent of CAA agencies and states reported at least one participant, agency or community outcomes--an increase from 85 percent in the previous year. This marked the first time that data were reported from all 50 states, Washington, DC and Puerto Rico.

The report concludes that while improvements are needed, ROMA implementation has made considerable progress. First, most CAAs and states (including Washington DC and Puerto Rico) understand the concept of ROMA. This is important for focusing attention on tracking and achieving intended results and measuring those results in the form of outcome indicators. Second, the number of CAAs and states using ROMA for program improvement has increased. This means CAAs are using measures of community action outcomes to assist in defining and adjusting efforts with program participants and communities to achieve better outcomes.

Third, the report suggests that the “bottom-up” approach to ROMA implementation appears to be working as evidenced by the 98 percent participation rate in reporting outcomes by CAAs and states. Aggregating CAA outcome information at the national level, however, continues to present challenges.

Of the 52 entities (50 states, Washington DC and Puerto Rico) comprising 935 CAAs reporting outcome data, some measures were more commonly used than others. For Goal 1, obtaining a job, increased income, employment for 90 days, improving literacy

or getting a GED, better housing and home ownership were the most frequently reported indicators. The report indicated 70,360 individuals in 42 states were reported to have obtained employment and 32,603 individuals in 23 states increased annual income through employment. The table below indicates the frequency with which the more common indicators were collected for Goal 1:

Table 1. Common indicators collected for Goal 1

	Participants	States
Obtained job	70,360	42
Employed 90 days	17,426	24
Safe, stable housing	33,795	26
Home ownership	1,861	16
Ending homelessness	12,662	23
Increased income	32,603	23
Literacy/GED	22,903	32
Post-secondary Ed.	12,846	22

Of the 52 government entities and 935 CAAs reporting outcome data, some measures were also most commonly used for Goal 6. Emergency needs met, maintained independent living, improved housing, reduced energy burden, child development, and maintained independent living (persons with disabilities) were the most frequently reported measures for Goal 6. The table below indicates the frequency of use for each of the measures for Goals 6.

Table 2. Common indicators collected for Goal 6

	Participants	States
Emergency needs met	2,764,132	43
Maintained Ind. Living (elderly)	816,008	28
Reduced Energy Burden	392,471	23
Maintained Ind. Living (disabled)	135,246	14
Improved family functioning	52,000	20

C. ROMA Goals & Outcomes

Programs have been developed to address a range of issues for families in poverty. Historically the successes of these programs have been documented anecdotally or not documented well at all. In responding to GPRA, the Management and Assessment Task Force created six national ROMA goals as listed earlier.

The specific focus of this guide is on goals one and six, though the processes and systems described may also apply to other goals and outcomes. The processes were also developed during a time when the number of indicators from which to choose under each Goal was much greater. Those longer lists of possible outcome indicators have recently been reduced to the following national performance indicators:

Goal 1: Low-income people become more self-sufficient.

1. Number and percentage of participants in community action employment initiatives who seek and obtain jobs.
2. Number and percentage of low-income households receiving community action services (i.e. housing, nutrition, child care, transportation, medical assistance) that achieve economic stability as measured by continuous employment for 90 days or more.
3. Number and percentage of low-income households in community action resource enhancement programs that experience an increase in income (from a variety of sources, including improved employment, child support, Earned Income Tax Credits or other tax relief).

Goal 6: Low-income people, especially vulnerable populations, achieve their potential by strengthening family and other support systems.

1. Number and percentage of senior citizens receiving services from community action that maintain an independent living situation as a result of those services.
2. Number of households served by community action that sought emergency assistance and the percentage of those households for which assistance was provided (i.e. cash, food, shelter, medical care, protection from violence, transportation, legal aid).
3. Number and percentage of all children and youth participating in agency development or enrichment activities that achieve program goals.

A variety of approaches have been undertaken by Community Action Agencies to measure family development outcomes. The focus of Section II is to compare, contrast and evaluate some of these approaches on their identification, implementation, use, improvement and reporting of outcomes in order to derive supportive evidence for recommendations about a best practices approach.

II. Approaches to Measuring Family Development Outcomes

In reviewing the approaches to measuring family development outcomes described below we developed standards by which a comparison of approaches may be conducted. The standards developed should be considered heuristic and take into account both the data systems and utility of the data for ROMA. While it is beyond the scope of this guide to conduct an exhaustive comparison and analysis of approaches, the criteria may prove useful for those considering adoption or adaptation of a model.

A. Approaches to Measuring Family Development Outcomes by States

The following brief descriptions summarize a variety of approaches undertaken by states implementing ROMA.

CALIFORNIA MATRIX MODEL

Overview

The California Matrix Model is a set of matrices intended to be used for assessment, case management and evaluation. The model consists of three distinct instruments: 1) the Family Development Matrix, 2) the Agency Development Matrix, and 3) the Community Scaling Tool. These instruments were created through the efforts of federal, state and county agencies interested in achieving results accountability. The Institute for Collaborative Studies, California State University Monterey Bay, along with collaborating partners including the National Resource Center for Family Centered Practice, further developed the Matrix Model to refine the language, structure and guidelines for use as well as subject the model to analyses of validity and reliability (Endres, Richardson and Sherman, 1999).

The family development matrix is intended to assist workers in health and human service organizations to better serve clients using a strengths based assessment that also serves as an evaluation tool. This process of using an assessment tool to empower and track individual and family outcomes over time is one of the most unique features of the model.

Measurement and Procedure

Twelve domains comprise the Family Development Matrix: Adult Education/Employment, Children's Care and Safety, Children's Education/Development, Community Relations, Family Relations, Finances, Food/Clothing, Health/Safety, Immigration/Resettlement, Shelter, Social/Emotional Health and Competence, and Transportation/Mobility. Agencies may choose to use all of the domains or a subset of the domains. Feedback is given to the client using the matrix to show progress since the last encounter between the worker and the client.

System Requirements

The software necessary for the Family Development Matrix is a simple spreadsheet. Some agencies have prepared Microsoft Access files for data entry, however, the basic output is a spreadsheet of data with multiple records where the date of contact indicates that the assessment has occurred. Simple analyses can be performed using the spreadsheet software. Additional analysis can be performed by importing the spreadsheet data into SAS, SPSS or other sophisticated statistical programs. Only basic PC hardware is required.

ILLINOIS

Overview

Illinois uses the Family Self-Sufficiency Scale. The scale was developed by the Bucks County Opportunity Council in Pennsylvania to help families move toward self-sufficiency. It has proven to be an effective instrument in measuring client progress toward self-sufficiency.

Measurement and Procedure

Initial contact between the potential client and the agency usually occurs through a phone inquiry or walk-in inquiry about services. During an initial screening, a worker schedules an intake interview and informs the potential client of what required documentation to bring. The initial assessment is made at Intake after which time the client and the staff person collaboratively develop a Goal Plan including both short and long term goals for the domain under which the client is requesting services. Quarterly follow-up assessments are conducted.

The Family Self-Sufficiency Scale comprises ten domains that indicate progress toward self-sufficiency, each domain with its own set of indicators. These domains include: Income, Employment, Education, Transportation, Parental/Social Functioning, Substance abuse, Mental Health, Health Insurance, Shelter and Subsistence. Each domain is scored on a continuum from total dependency (0) to total self-sufficiency (10). With a score of two and below, the client is considered “totally dependent.” A score between three (3) and five (5) renders a client participating in self-sufficiency activities, but still dependent. A score of six or more is considered in the successful range of self-sufficient. After combining the scale scores, a client achieving a total of sixty points on the ten domains would be considered self-sufficient so long as s/he is employed and not participating in a federal entitlement program.

MINNESOTA

Overview

Minnesota uses the ROMA Self Reliance Achievement Scale (SRAS) to help families move toward self-sufficiency and to measure client progress toward self-sufficiency. Minnesota’s training manual makes it very clear that the Self-Reliance Achievement Scale is not a participant needs assessment or a case management tool. Its sole purpose is to measure an individual’s progress across areas and to determine whether persons served by community action are becoming more self-reliant over time.

Minnesota piloted the SRAS with 7 seven CAAs in 2000; all 29 CAAs implemented use of SRAS in 2003.

Measurement and Procedure

The ROMA Self Reliance Achievement Scale comprises 15 domains which indicate progress toward self-sufficiency. Each domain has its own set of indicators, which are typically scored one (1) to four (4). The 15 domains are: Hours of Employment, Job Retention, Income, Education, Child Support, Earned Income Tax Credit, Adequacy of Income for Food and Shelter, Financial Credit, Housing Stability, Housing Affordability, Health Insurance, Childcare, English Proficiency Education, Transportation, and Informal Social Support.

Data used to complete the SRAS comes from complete and accurate documentation in participant case files. The observations of family progress must be scored at least twice and at least 30 days apart, but not necessarily every month. Also, a scale is completed for every adult participant in a household. Data are only collected on the services for which a participant is being served and for up to three services in each domain. Staff members complete the Self-Reliance Achievement Scale for every adult participant in the household engaged in services.

System Requirements

The hardware system for the Minnesota model is a networked information system. The software system for the Minnesota model was written in Access. The system is capable of organizing and tabulating the program outcome data. Agencies involved in Minnesota's pilot testing of the SRAS are currently testing an electronic version of the scale, with plans for statewide, agency-based electronic data entry and submission. In the future, Minnesota anticipates that the SRAS will become web-based therefore making data collection easier and more efficient.

The Minnesota model uses the SRAS instrument with at least two observations to input data that covers the calendar year beginning in January and submitted quarterly for data entry. Minnesota Community Action delivered an output report on the Pilot Project earlier in 2003. Reports on client outcomes included in this report are available to the public and posted on the website.

PENNSYLVANIA

Overview

Pennsylvania developed the Outcome Measures and Results matrix for use in its state to measure families progress toward self-sufficiency. Pennsylvania has 43 CAAs that participate in the ROMA tracking and reporting system. CAA participation in the development and design of the system promoted the eventual "buy-in" by local agencies. Particularly important in Pennsylvania is the concept of "benchmarks" for employment and education. Although the word is often used as a substitute for goal, in Pennsylvania the concept refers to "a snapshot of a statistic at a specific point in time." Pennsylvania's intent is to examine "results-based performance" or accountability.

Measurement and Procedure

Pennsylvania uses two matrices to help families move toward self-sufficiency and measure client progress toward self-sufficiency. Pennsylvania separately has created seven (7) domains for Goal 1 and eleven (11) domains for Goal 6.

The seven domains for Goal 1 indicate outcomes toward self-sufficiency. The seven domains are: Obtained Employment/Self-Employment, Maintained Employment 90 days, Increased Earned Income, Increased Household Resources, Increased ability to Manage Income/Assets, Housing, and Eliminated Barriers to Employment/self-sufficiency. Each domain has its own set of indicators, which are typically scored one (1) to four (4). The indicators are part of a six variable matrix indicating where clients locate on a level of poverty continuum.

The eleven domains for Goal 6 indicate outcomes toward strengthening family and other supportive systems. Again, each domain has its own set of indicators. The indicators are part of the six variable matrix indicating where clients locate on a level of poverty continuum.

IOWA

Overview

Iowa developed the Automated Assessment of Family Progress (AAFP) system (Appendix A). It is a matrix model intended to be used for assessment, case management and evaluation to assist workers in health and human service organizations to better serve clients using a strengths based assessment. It also empowers individuals and families as it tracks their outcomes as well as agency and community outcomes over time. Section III, below, provides a more complete description of the Automated Assessment of Family Progress.

The Iowa approach to ROMA uses the “strengths-based approach to practice.” Front line workers are critical to assessment, service delivery and use of outcome data. The system was designed to collect and report outcome measures that also provide feedback for practice improvement. In this context, AAFP data collected are measures useful to workers as they help families toward self-sufficiency, and that families themselves find meaningful. Case workers can provide clients visual feedback using simple graphics in spreadsheet programs to show families how they are progressing in various domains, as well as overall progress from visit to visit. Thus family development outcome measures become a part of service delivery not simply an add-on evaluation system.

Measurement and Procedure

Iowa identified ten domains for the AAFP instrument to measure family development outcomes based on the input of planners from numerous community action agencies. These domains were: Employment, Education, Community Involvement, Self-sufficiency, Household Management, Food/Nutrition, Health, Housing, Emergencies/Crises, and Household Linkages. The domains were originally scored as

scales and were revised as domains with multiple binary choice indicators for each domain (i.e., yes/no).

Following pilot testing it was decided that rather than measuring all 300,000 families that annually appear at Iowa CAAs, the AAFP would be formally administered to a sample to reduce the time and cost necessary for data collection. Staff agreed the sample should have a 95 percent confidence limit ($p < .05$) with a confidence interval of 4 percent. Data were reported twice a year at six month intervals. While this approach is consistent with the outcome measurement aspect of the model, the sampling procedure limits the practice and case management benefit for families not selected by the sampling procedure. AAFP was designed for flexibility in developing output reports even allowing for bundles of services and their associated costs to be related to the outcome measures.

System Requirements

Staff who participated in numerous planning sessions decided to use AAFP as a subprogram within the computerized information system being used to gather intake information. In other words it was simply an additional screen for workers already familiar with the system. This approach permitted each CAA to collect and store its own data, but also facilitated file exports and aggregation of common data elements at a central location (i.e., University of Iowa, National Resource Center for Family Centered Practice).

The hardware system varied based on CAA. Most systems using AAFP were networked desktop computers, however, at least one CAA was using a mainframe system. The only hardware requirements were that they be able to run one of the two operating systems. These systems include the Integrated Family Collection and Analysis Program (IFCAP) and the THO Software Systems. Data from those systems were exported as spreadsheet files.

B. Systems Assessment Criteria

In reviewing approaches to measuring family development outcomes, six descriptive criteria were identified that could be useful for assessment of data systems: participants, needs, inputs and outputs, and hardware and software. Participants refers to those involved in the assessment process and who are users of the outcomes data. Participants could be the families themselves, the CAA staff and direct workers, board members or CAA directors who make decisions related to the participants and programs at the local level, state agencies, or those at OCS or at other national or policy level. Needs refers to the specific information and utilization needs to which the data respond. Participants may have various needs for data so the extent to which those needs are met is an important criterion. Inputs refers to the measurement process, the instruments used to collect data and the protocol for so doing. Outputs refers to processes used to organize outcome data and produce reports. Hardware refers to the technology used and the required capacity (e.g., laptops with synchronization, mainframe, pencil and paper with keyed entry, etc.). Software refers to

the programs necessary to process the data and can range from a simple spreadsheet to relational databases with programmed results generation. The table below can be used to rate approaches based on these characteristics of data systems. Suggested rating criteria include simple binary codes for adequate or inadequate, or more sensitive satisfaction scales such as a Likert-type scales where 1 indicates unsatisfactory and 7 indicates very satisfactory.

Systems Criteria	Requirements Analysis for Systems Measuring Family Development Outcomes				
	Family	Worker	CAA	State	OCS
Participants					
Needs					
Inputs					
Outputs					
Hardware					
Software					

In addition to reviewing the systems requirements it is important to examine the utility of the system for responding to ROMA. We have maintained nine utility criteria for ROMA since receiving technical assistance in 2000. These nine assessment criteria are useful for rating the extent to which the system is utilization focused for users at the various organizational levels. These criteria capture the basic elements of data needs in practice for each of the five primary users.

Utility Criteria	Requirement Analysis for ROMA Measures of Family Development Outcomes				
	Family	Worker	CAA	State	OCS
Ease of Use					
Timely Access					
Understandable					
Usefulness					
Relevance					
Ease of Change					
Specificity					
Training					
ROMA Linkages					

The ROMA assessment criteria refer to the interaction between the data systems and the ROMA information and decision-making demands placed on those systems. Ease of use is the level of effort necessary to collect the data. A worker needs to be able to easily gather and record data. Timely access is important for the various users of the information. A worker may need access to case data while an agency director may need timely access to reports of results for staff, the board, other members of the CAA network, and the state or federal government. Reports produced by the system need to

be understandable in ROMA terms. The data output must be in a format that is easily interpretable and linked to ROMA concepts, goals and measures. Tables or statistics which are difficult to read, understand or link to ROMA will not be understandable to workers, administrators or board members are not helpful for understanding how or why certain outcomes occurred and how to improve performance.

Usefulness for ROMA refers to the extent to which the measures address the ROMA goals and outcomes. The more measures directly capture ROMA outcomes the more useful they are; The farther removed as proxy measures the less useful they are. Relevance is critical for staff, administrators and boards to engage them and provide information that can be used in planning; the measures must have face validity for measuring activities in which users are involved. Ease of change refers to flexibility for collecting new data. New data could replace existing data or could be additional data though failing to maintain comparability of measures over time can compromise the usefulness of data. Our experience is that a balance must be found among all the assessment criteria.

Specificity may be related to scientific concepts of sensitivity and power. Measures must be clearly defined so that they are reliable, meaningful able to capture change. Training is directly related to the complexity of the system. Training may be measured in the amount of time and energy required to use the system for collection of data as well as processing and interpreting the results. ROMA linkages refer to the extent to which the measures capture information needed for state and national results.

LESSONS LEARNED

OCS promulgated six national goals and indicators, and the existing systems, despite their variation, attempt to address these as outcomes. The ROMA process permitted local agencies and state systems to develop independent models for measuring outcomes. Therefore the present system represents data elements selected by CAA staff, consultants and a host of others. This process focuses attention on individual cases and case management variables. As a result, more progress has been made on family development outcomes than other outcomes.

The current approach focuses attention on standardizing information needed for a few domains and generally respects the autonomy of CAAs and state systems. This approach treats the change as a small issue rather than attempting to change the "system" which might provoke large scale resistance. However, the massive amount of data collected do not rise to the level of common measures and are not generally efficient for larger state and OCS systems for summary outcome reports. There is surface similarity in the domains (variables) of different systems, but the reality is the plethora of independent models that measure domains and outcomes in different ways make it difficult to compare programs and outcome data from the various states.

The present outcome measures collection processes are expensive due to inefficiency. They tend to be labor intensive often using professional staff to perform clerical tasks and producing a plethora of measures that exceed the capacity for analysis. Typically staff fill out forms either “on-line” or on paper which are then collated and sent elsewhere for data entry. Following data entry, the output of results is produced and returned with considerable time delay and often to contact persons who do not provide the data or adequate interpretation and feedback to those who could directly apply the results. However, creating a web-based GPRA reporting system has not proved to be the panacea some could be tempted to believe (cf. <http://www.csat-gpra.samhsa.gov/index.cfm>).

While the present focus is on data collection, to paraphrase Deming (1986), data is not information; information is not knowledge. The larger ROMA system needs outcome data for OCS to respond to GPRA and audiences relevant to sustainability of programs. In this sense, there is need for greater clarity about OCS data needs. However, it is also clear that outcome data are not used in the larger ROMA system as intended and management by ROMA has not been fully implemented.

OCS needs standardized data from state systems. Any change should consider 1) what data OCS needs regarding the national goals and self sufficiency; and 2) how to analyze the reported data, i.e., what outcomes, tables and summary statistics are required to help OCS decision makers understand whether self sufficiency is being attained in terms of the national goals.

Though there are several similar models (e.g., California Matrix Model), the approach undertaken by the Iowa CSN is one model that we believe addresses many of the concerns raised above for measuring family development outcomes. It is simple and provides data that may be analyzed at a variety of organizational levels. The use of those data in decision-making processes, from working directly with families to decisions associated with program at the highest levels, can be provided from the analysis of a single aggregated dataset. The AAFP is presented in Appendix A and is more thoroughly described in Section III, below.

III. Automated Assessment of Family Progress (AAFP)

This section describes the approach undertaken by the Iowa CSN. The AAFP is an exemplary practice and outcome measurement tool that provides data for analysis at several levels of social organization: individual, agency, community, and program. For simplicity, its use is described for three primary audiences: caseworkers, data managers and analysts, and policy and decision makers. (The AAFP is presented in Appendix A.)

The AAFP is an automated client data collection and tracking system authorized for use by Iowa's Community Action Agencies through the DCAA and ICAA. Beginning in 1998 as a part of the standard intake process, workers began to ask additional questions of randomly selected participants in Community Action Agency programs at all intake sites across Iowa. These outcome measures scales are completed as part of the standard ongoing client interview process with questions asked corresponding to the outcome measures scales. The AAFP data locate families on scales covering a number of dimensions of the status of the family at the time they are first seen, and, over time, these scales are used to track family progress. Measuring the scales at intervals allows the system to track progress over time as well as provide community assessment information as new entrants to the system are added at each measurement time point.

The Automated Assessment of Family Progress developed from earlier work at MICA (Mid-Iowa Community Action) and the National Center for Services Integration (Bruner, 1992). The original matrix assessed family development concepts on one scale with three categories: thriving, safe, at-risk. Using these categories and family development dimensions as a starting point, and after many, many hours of meetings, drafts, reviews, negotiations and testing, the AAFP measures were determined. The major difference between the AAFP matrix and other systems is that the indicators are dichotomous resulting in greater reliability, and there are strengths-based indicators under each outcome area. The scales may be used for aggregate analysis at the individual, agency, community and program level.

Using the matrix, an initial family assessment is conducted on the following 10 domains:

- Employment
- Education
- Community Involvement
- Self-sufficiency
- Household Management
- Food/Nutrition
- Health
- Housing
- Emergency/Crisis
- Household Linkages

Subsequent assessments provide measures that can be compared over time. The analysis of these measures can show progress or change in status of the family development outcomes.

A. Use of AAFP by Workers

Outcome measurement can improve direct practice through strengths-based case management implementing ROMA at the micro or individual level. The AAFP relies on measures at the individual case level for performance-based management and improvement of service delivery. Direct workers are critical to the system because they must collect and report accurate assessment data for families. We also contend that the focus on accurate outcome assessment can also improve the relationship between the direct worker and the family, though the worker must bring family development and strengths based skills to the alliance with the family. Failing this, measurement becomes a reporting task that is not valued by the worker and consequently is similarly perceived by the families.

AAFP as Core to Strengths-based Practice

The AAFP was designed to improve direct practice through principles of strengths-based and solution-focused case management. AAFP questions and indicators are strengths-based rather than framed in deficit-based language reflecting this orientation.

The AAFP requires simple answers, generally “yes” or “no,” to questions about the family. Scales derive from the composite scores of items under the domains of the AAFP. The initial assessment provides workers and families with instant feedback on the status of the family on ten important life domains. Subsequent assessments provide measures of progress on each of the items and domains. In this way, using family development outcome measures becomes part of the service delivery process for the benefit of measuring clients’ progress. It also assists caseworkers in empowering families to identify strengths and how to build on those for general family development. Because outcome measures are written in straightforward “plain language” and are strengths-based, it is easier for the families to identify their strengths and see how they might build on those to make progress toward a specific goal.

B. Use of AAFP by Data Managers and Analysts

The AAFP is recognized as a valid and reliable outcome measures tool (Endres and Richardson, 2000). The approach used to develop the AAFP resulted in a simple and inexpensive computerized data collection system by which to examine data for families seeking assistance from Community Action Agencies. All CAAs used a computerized system to collect and store AAFP data from each county at each community action agency resulting in measurement at all CAAs and all counties in the state. Data managers at each site easily export data to flat files and then send those data to The University of Iowa School of Social Work, National Resource Center for Family Centered Practice, for processing, analysis using SPSS, and reporting of results.

AAFP Data Collection and Processing Procedure

Data collection is an essential element of the ROMA system. The AAFP data collection system is designed to collect data on outcome measures (scales) related to goals that inform the need for services and programs. The AAFP data are collected on an additional screen of the existing computer system used for gathering intake information at sites across the state. This is done so that the AAFP is integrated into the data collection already in place on CAA computer systems.

The data collection process begins when a family appears at a CAA for assistance. Each time the family appears for services at the CAA, the AAFP is completed. Generally, the family (or family member) will be sitting with a worker near a computer where data may be entered directly. If an application is completed on paper or processed at a location other than a county center, a prompt allows the program to exclude the household from the AAFP selection process; alternatively the worker may enter data into the system from the completed paper form at a later time.

Data are collected at the CAA and exported in flat file format (e.g., Excel) so that data are easily converted to a common format for statistical analysis. All CAAs send data to a central location at regular intervals; this creates a statewide system of common outcome measures. In Iowa, CAA data managers were required to send data twice a year at six-month intervals, though the original plan was for quarterly tracking for greater sensitivity to the seasonal effects common to CAA work.

Although the principles of case management were to be used for all cases, because of the added effort (and cost) associated with collecting and reporting data for all 300,000 annual cases in the state, a random sampling procedure was agreed upon and implemented. It was agreed that the sample should have a 95 percent confidence limit ($p < .05$), with a confidence interval of four percent (± 4). The sampling procedure involved each county estimating how many households they serve during a 12-month period and reporting this number to a selection setup program. (The estimate was to be the number of families that the CAA will certainly exceed, but is as close as possible to the actual total number that will be seen over the year.) The selection setup program then creates a table based on the number of households that will be served. Agency file numbers (family identification numbers) are then selected based on a table of random numbers. Families are either selected or not selected based on the random number and the percentage of families needed for the sample. For smaller counties an increased sample proportion of 20 percent was used for statistical reliability. The sampling procedure ensures that every household has the same chance of being selected regardless of when the family appears (those appearing late in the year have the same chance of selection as those appearing early in the year), or how often they appear (whether once a year or several times each month).

While ROMA training and technical assistance has increased capacity, as we have worked in the communities, it has become clear that the most effective mechanism for sustaining ROMA is through work with those who are nearest the program participants of the CAAs. For this reason, we provided software and trained data managers to

analyze their own agency data for more specific and potentially thorough local analysis. Providing each agency with the tools to analyze data for their CAA, and each of the counties located within a CAA service area, facilitates interpretation within the context of local service provision. Outcomes data can be used by planners with the assistance of direct workers to further identify program and resource needs. Local cost analysis can also be performed and when state level data become available, agency specific results can be compared with the state level data. This procedure also provides data managers at the CAA with useful data for engaging direct workers, managers and decision makers in a timely manner about how the CAA is doing compared to the state.

Compiling AAFP Data

To create a statewide assessment system of ROMA measures, data are combined from the multiple spreadsheets of each CAA into one dataset for analysis using more advanced statistical software programs. SPSS (Statistical Packages for Social Sciences) was used in Iowa but numerous other programs exist (e.g., SAS, Stata, etc.).

We found some data cleaning to be necessary, especially during the first two to three rounds of data reporting from the CAAs. However, during that time a set of decision rules, cleaning scripts and instructions were created to accelerate the process. Some examples of decision rules included:

- Only accepting data within the agreed upon date range, e.g., Oct 1 to March 31 and April 1 to Sept 30.
- Converting string data to numeric.
- Only accepting data from the first visit to the CAA and last visit in subsequent reporting periods.

Once cleaned, cases are filtered so that each case that appeared in previous AAFP datasets is “tracked” while “new” cases are identified as initial assessments. Merging the new dataset with previously received datasets at each interval accomplishes this task. When a new case matches an old case, it becomes part of the “matched pairs” dataset (further described below). Where no previous match is found, those cases are “initial assessments.” In our ongoing analysis, after the initial assessment families that appear more than once during a sampling interval have their most recent AAFP scores used. However, more thorough analysis could include data from every AAFP assessment to detect incremental changes occurring from visit to visit. In our ongoing analysis, we defined a third group who appeared in serial intervals as a “tracking study.” To insure that a selection bias is not created, it is imperative that cases in “tracking studies” be followed whether or not they appear for services. We found that tracking only those who appear for services represents needier families and results in assessments that don’t appear to be improving. To accurately capture outcomes, there is a need to follow a sample of cases that cease to present for services to correct for the selection bias.

Thus three basic stratification levels are derived from the outcomes reported:

1. *Initial assessments:*

Analysis of these data in aggregate produce a community needs assessment based on the strengths and needs of families appearing at CAAs who have not previously accessed CAA services. Aggregation based on program, county, or other units of analysis can yield similar needs assessments for other purposes. The initial assessment data also illustrate strengths and where family development outcomes can be improved in individual work with families.

2. *Matched Pairs:*

Analysis of these data produce measures of change over time for families who have received services from CAAs more than once. In aggregate, these data are relevant for both summative and formative evaluation (i.e., how well CAAs doing, adjustments that might improve effectiveness, etc.). These data also illustrate achievement of family development outcomes for work with individual families.

3. *Tracking Study:*

Analysis of these data provide continuous outcome measures for families seen at serial (i.e., each) AAFP data collection and reporting intervals. This sample represents those with long term contact or ongoing service needs and is perhaps the most difficult segment of the target population to analyze. Families who consistently access services at the CAA can be relatively intractable, and their outcomes may appear to indicate that progress is flat or even deteriorating between some assessment intervals. This can be due to internal and/or external (uncontrollable) circumstances, such as the economy, unemployment, etc. For program evaluation purposes, it is important to follow and measure at least a sample of those seen in the tracking study who then do not return for services. This allows for the analysis of the bias attendant to the results from only measuring those who continually appear for services. The tracking data may also be useful in practice, however, by illustrating to families where there are strengths and helping to build on them.

AAFP Data Analysis

As described above, the data were stratified into three groups for analysis: initial assessments, matched pairs, and tracking. Data processing queries that generated sets of tables “at the touch of a button” each time data were processed and sent from the CAAs were made available to each of the CAAs. In fact, simple queries were altered so that CAA staff could locally determine how their own CAA data were analyzed.

To create outcome scales and standardize reporting, we converted to binary (0,1) format those questions not already dichotomously coded. This means that in instances where the AAFP asks a question that is in other than yes or no format, the data were recoded. The result sought is an outcome expressed as a mean or percentage reporting a positive or strength outcome. For example, in the Emergency & Crisis

Domain the AAFP asks: "Is household in an emergency or crisis situation?" The three possible responses are:

1. = Crisis now exists
2. = Crisis potential exists
3. = No crisis at this time

Our procedure was to score 1 and 2 as zero and 3 as 1. In contrast, on the Food & Nutrition domain the AAFP asks: "Is household food supply adequate?" and the two possible responses are "Yes" coded as 1 and "No" coded as 0 (i.e., no recoding was necessary). Using these "standardized" scores allowed for the construction of scales under each of the domains while maintaining direct comparability of the individual items.

Case example: 100 families are interviewed using the AAFP. For the Food Supply question above, 65 families respond "Yes" that the food supply is adequate. This means that the average (mean) number of families with adequate food supply is 0.65 which can be expressed as 65%. The results indicate that 65 percent of respondents reported having adequate food supply upon initial appearance at CAAs.

AAFP Data Reporting

The results were presented in an easy to read one page format. Table 1 below is an example of the results after the data reporting in one interval in 2003. This table shows the percent assessed with positive outcomes for each measure in each of the ten domains. The first reporting period for initial assessment ended in April of 1999 (initial assessment column header, 1998), each subsequent "year" represents a six month data collection and reporting interval for a total of eight intervals for years 1998 to 2002.

Reporting Made Simple

Reports should be clear and understandable for both direct workers, managers and policy and decision makers. Tables such as the one presented below are readily understood with limited explanation. CAAs can reduce the amount of report writing by ensuring data are presented concisely in understandable and useful tables. Charts can also easily illustrate the differences made over time on each domain (cf. Automated Assessment of Family Progress, 1998-2002b et al.).

Table 1 below presents two types of outcomes data: Initial Assessments and Matched Pairs. Under Initial Assessments, the table shows the percentage of families assessed with strengths (positive outcomes) for the ten AAFP domains at first appearance for each outcome item in each year of the study presented.

Case example - Initial Assessment: Employment, Job Status

For those families initially appearing in 1998 at CAAs across the state, 32 percent reported that adults in the household were employed. This percentage fluctuated over the four year time span: 34 percent in 1999, 31 percent in 2000, 30 percent in 2001, 34 percent in 2001a, 34 percent in 2002, 31 percent in 2002a and 24 percent in 2003. These data indicate that from 1998 to 2002a, approximately one-third of those initially appearing at CAAs have adults with employment as a strength on which to build. In 2003, the employment percentage dropped to 24 percent which could be the

appearance of the effect of external conditions affecting employment such as the declining U.S. economy.

Case example - Matched Pairs: Emergency/Crisis, Frequency of CAA Emergency Assistance to Household

In the "Matched Pairs" section of Table 1 comparisons are reported for 1) those initially assessed in the reporting period for April 2002 compared to their assessment one year later in the reporting period for April 2003, and 2) those initially assessed in the reporting period for October 2002 compared to their assessment six months later in the reporting period for April 2003. The smaller number of cases as compared to the Initial Assessments indicate that while 2031 families were assessed for the first time in 2003, another 280 who appeared were seen one year prior (2002) and another 171 were seen during the previous reporting interval six-months earlier.

As an example of interpretation of the results, the second outcome measure under the Emergency/Crisis domain using the matched pairs columns shows 62 percent with no emergency assistance visits compared to 60 percent for those same families one year earlier (an improvement of 2 percent). For families that appeared in April 2002 and again in October 2002, there was a sharp, statistically significant decrease in the percentage of families visiting CAAs for emergency assistance, from 30 to 63 percent.

Table 1. Automated Assessment of Family Progress Initial Assessments & Matched Pairs Outcomes by Year

Domain	Initial Assessments								Matched Pairs			
	1998	1999	2000	2001	2001a	2002	2002a	2003	2002 to 2003		2002a to 2003	
Employment (n= # of cases)	6484	2131	6586	1905	6036	2516	5116	2031	280		171	
Job Status	32%	34%	31%	30%	34%	34%	31%	24%	24%	32%	30%	31%
Benefits	73%	72%	60%	69%	62%	80%	61%	66%	61%	61%	63%	68%
Job Training	80%	66%	84%	73%	85%	75%	84%	68%	79%	77%	86%	83%
Job Goal	91%	89%	90%	85%	89%	88%	87%	77%	84%	85%	83%	83%
Education												
Training Education	73%	64%	78%	68%	79%	69%	77%	63%	72%	73%	78%	81%
Basic Education	87%	75%	90%	82%	88%	82%	90%	84%	87%	87%	82%	87%
Education Goal	88%	84%	91%	87%	91%	86%	91%	82%	90%	93%	87%	91%
Community Involvement												
Community Activities	60%	58%	61%	62%	62%	65%	61%	55%	57%	65%	68%	64%
Community Involvement Goal	91%	83%	94%	87%	92%	88%	92%	84%	90%	89%	95%	93%
Self-Sufficiency												
FIP Exit Plan	19%	23%	15%	18%	13%	22%	13%	18%	30%	26%	30%	20%
Self-Sufficiency Program	7%	16%	6%	8%	7%	8%	7%	7%	8%	14%	3%	4%
Self-Sufficiency Goal	90%	85%	93%	87%	93%	92%	93%	85%	94%	94%	91%	86%
Household Management												
Savings Account	33%	33%	37%	32%	35%	35%	32%	23%	30%	26%	28%	22%
Bills Paid on Time	71%	58%	74%	61%	67%	50%	62%	54%	56%	47%	62%	62%
Basic Needs Met	36%	21%	38%	30%	41%	24%	34%	22%	16%	18%	23%	23%
Quarterly Income	\$2,771	\$2,587	\$2,803	\$2,843	\$3,060	\$3,149	\$2,972	\$2,627	\$2,912	\$2,980	\$3,445	\$3,219
Debt Load	61%	54%	58%	56%	53%	47%	54%	57%	53%	51%	55%	61%
HH Management Goal	85%	78%	89%	82%	87%	75%	85%	75%	78%	79%	85%	85%
Food/Nutrition												
Adequate Food Supply	51%	50%	64%	54%	63%	46%	55%	45%	42%	41%	32%	34%
Food Assistance Benefits	47%	51%	42%	49%	37%	46%	42%	48%	61%	52%	64%	64%
Food/Nutrition Goal	92%	86%	94%	88%	92%	85%	90%	82%	61%	52%	64%	64%
Health												
Health Insurance	81%	74%	82%	78%	80%	80%	78%	70%	85%	83%	90%	92%
Preventive Health Habits	90%	82%	92%	91%	92%	92%	89%	88%	93%	94%	97%	95%
Health Affecting Other Areas	74%	64%	73%	66%	75%	69%	76%	68%	75%	69%	74%	74%
Health Goal	89%	89%	94%	88%	91%	88%	91%	84%	88%	88%	92%	85%
Housing												
Afford Housing Available	86%	81%	88%	83%	88%	82%	86%	78%	86%	86%	83%	82%
Current Housing Affordable	86%	72%	88%	82%	85%	71%	83%	75%	80%	80%	78%	89%*
Current Housing Adequate	92%	86%	95%	93%	95%	91%	94%	90%	90%	84%	93%	95%
Housing not in High Crime Area	8%	18%	7%	9%	5%	8%	5%	7%	8%	15%	9%	9%
Housing Goal	91%	85%	95%	88%	93%	85%	92%	83%	91%	85%	91%	93%
Emergency/Crisis												
Household not in Crisis	74%	55%	74%	53%	66%	42%	64%	48%	40%	46%	76%	63%
Freq of CAA Emerg Asst to HH	81%	64%	85%	70%	86%	78%	87%	80%	60%	62%	30%	63%*
Emergency/Crisis Goal	88%	81%	93%	84%	90%	80%	88%	81%	82%	83%	90%	86%
Household Linkages												
Transportation Needs Met	78%	66%	81%	69%	76%	65%	77%	70%	71%	73%	76%	83%
Legal Needs Met	86%	82%	88%	82%	82%	78%	83%	82%	81%	81%	90%	94%
Family/Friends not in Crisis	86%	77%	88%	80%	87%	78%	87%	82%	83%	81%	85%	88%
Child Care Needs Met	90%	80%	91%	83%	88%	80%	90%	84%	84%	87%	92%	90%
Household Linkages Goal	92%	84%	94%	88%	93%	86%	92%	83%	91%	86%	90%	91%

C. Use of AAFP by Policy and Decision Makers

Outcome measurement is integral with the philosophy and concept of ROMA. We have addressed the value of AAFP data for families, workers and direct practice, above. The AAFP also provides valuable information for local CAA boards and directors, state administrators, and at the federal level as well. The process provides outcome measures of programs and services, the extent to which they make a difference, and how much of a difference. Data from the measurement of increments of outcome by the AAFP can also be linked to cost data to provide return-on-investment information. In aggregate, AAFP data provide measures of change and point-in-time assessment at several levels of analysis. These data may be analyzed by geographic area (counties, agency service areas, statewide), programmatically by services received, over time (time series), or for assessing validity and reliability.

1. Local CAA Boards

AAFP family development data are readily available, useful and relevant to CAA boards and directors. Local CAA boards can use the Initial Assessments of the AAFP to illustrate a robust picture of the community needs. Because data can be broken down by county and by CAA program in each of the CAA areas, data analysis can also be performed at these levels. This will allow local CAA boards to make decisions about the types and amount of services needed in a particular county or for the entire CAA. Decisions of CAA boards and directors can be based on routinely collected needs assessments data rather than point in time assessments which soon become dated. By collecting the same measures over time trends can also be seen.

2. State Administrators

The calculation of return-on-investment (ROI) can be calculated for services (ROI, cf. Benson, 1999; Hunnemann and Richmond, 1996) and achievement of desirable outcomes can also be analyzed for institutional decision-making (Richardson and Huff, 2001). Cost data may be associated with services provided and linked to outcomes to obtain cost per increment of outcome estimates. As stated above, the outcomes are included in Iowa's state data collection systems. Iowa's CIS currently monitors service delivery and outcomes comprise the Integrated Family Collection and Analysis Program (IFCAP) and the THO Software Systems Inc. Bundles of services and their associated costs may also be related to outcomes measured by the AAFP. This strategy promotes tracking of change at the family, county, agency, program and state level and also facilitates linkages between services, cost information and outcomes. Thus, all CAAs using this approach to goals and outcome measures can access the inputs and outputs as well as outcomes (cf. Friedman, 1995). With additional training and technical assistance, CAAs can track and analyze incremental change and cost at all levels of analysis from individual and program level, to county level, to CAA level, to statewide analysis. Though CAAs are unlikely to be responsible for the statewide level of analysis, having statewide data available provides essential comparative information. Developing statewide systems connect outcome measures and scales with services, programs, and fiscal management information systems and decision making based on results is ROMA.

Many funding sources now require outcome measures in some form or another. The AAFP provides GPRA data on outcome measures that may be tailored to these requirements. This facilitates the ease with which state administrators can report GPRA data and OCS can compare these data.

3. Federal Administration

Linking services to family development and stability outcomes is critical to the mission of CSBG. The key to the development of an outcome-based service system is maintaining a focus not exclusively based on the measurable outcomes, but on the linkages between the outcomes, the presenting needs and the services connected to the outcomes. Once trends are established through collection and reporting of the data, the real work of improving the life circumstances of those who seek the assistance of agencies can be better informed. With better and more easily obtainable outcome information, results can be improved. While the AAFP provides common measures across agencies in one state, common measures across states seems equally important. Though beyond the scope of the present work, a crosswalk of outcomes collected by each state seems an important task to be undertaken for federal oversight.

D. Lessons Learned

In working with the AAFP and outcomes over the last two decades, some truths have emerged. The following items represent a beginning list of considerations when faced with adopting or adapting a system for measuring family development outcomes. Where appropriate, other sources attributable to the derivation of the truth are noted.

- Don't confuse issues of data analysis with those of data collection.
- Like all data systems "everyone is in charge, and no one is in charge" (Arnie Anderson, 2003).
- "Bad information" on time is better than good (i.e., perfectly accurate) information late (Arthur Shostak, Michael Patton, circa 1980).
- Know your audience: Who makes decisions? What are their information needs? And keep your audience involved in the process (Jakopic, 2003).
- If the audience doesn't like the results, be careful they don't try changing all the measures.
- If you do change measures, don't throw the baby out with the bathwater (Jakopic, 2003).
- If changing the measures doesn't satisfy the audience, be especially careful - the audience may want to change the messenger.
- Data collection is never done.

- Take a method and try it; if it fails, admit it, and try another. But above all, try something. (Franklin Delano Roosevelt, circa 1932).
- Resistance is a problem, for example, “Should the lastname field be 15 or 23 characters?”
- Refinement is never done.
- Don’t get mesmerized by technology.
- There is no such thing as perfect; perfect for you is not perfect for me (Jakopic, 2003).

Section 2. How to Develop Outcomes

- IV. The Outcome Measures Paradigm
- V. Efficient and Effective Outcome Evaluation
- VI. Step by Step Guide for Developing Outcomes for Family Development Programs

IV. The Outcome Measures Paradigm

The terms *outcomes* and *outcome measures* became popularized in the late 20th century as statistical results became part of everyday business. The paradigm shift was essentially one from the level of activity or services provided to what was achieved as a result of those activities. Interest waned in funding work activity replacing a focus on services to the outcome associated with those efforts. The term *results* is associated with academic journals and technical literature so *outcomes* became preferable and more acceptable in non-technical settings. However, along with *outcomes* it is important to measure the activities associated with programs or services intended to achieve outcomes. The distinction between process (activities) and results (outcomes) is an important one.

Outcomes are the desired changes that a program is trying to effect--whether the change is in individuals, families, neighborhoods, states, etc. Outcomes are why a program is being funded. Activities are the services provided and interventions used with the intention of bringing about the outcomes.

Traditionally in human services, accountability meant reporting what (all the hard work) you did (program activities or processes), not what happened (results or outcomes). For example, human service agencies might report:

- # of children served by the program
- # of hours of service provided
- # of sessions attended
- # of people participating in classes.

Now funders are asking for something more. Answers to the following questions have become commonplace in reports to funders:

- What did the program achieve?
- And what measures can you provide to demonstrate what you achieved?

Outcomes at different levels

Outcomes can occur at many units of analysis. For example, at the program level, outcomes may be achieved by participants in the program (e.g., individuals, families, groups). However, outcomes may also be achieved at the community, neighborhood, county, or state level. Knowing from which level outcomes are expected is important so that the level used for writing outcomes is appropriate for the expected impact of the program. For example, you might expect community-wide effects for larger, universal programs that are available to the entire community (i.e., home visiting for all newborns). Smaller programs available to a particular high-risk group should most likely effect only participants in the program. An example of this is a program provided to all middle school youth might have expected impacts on all youth of that age range in the school district. The table below illustrates this comparison.

<i>Program level outcomes</i>	<i>Community level outcomes</i>
For specific programs serving a limited number of individuals/families: <ul style="list-style-type: none"> • % of children in the program who have received their full series of immunizations • % of high risk teens in the program who have remained in school • % of families in the program who did not have a subsequent confirmed child abuse/neglect report 	For county/district-wide initiatives: <ul style="list-style-type: none"> • % of children in the county who have received their full series of immunizations • % of high risk teens in the district who have remained in school • births to teens as a % of all births in the county

Levels of Focus for Measuring Change

Change can be measured at any or all of the following levels:

- Level I: Individuals (children, parents, families)
- Level II: All participants in a program
- Level III: All programs within an agency
- Level IV: Systems of care within a community
- Level V: Communities
- Level VI: Aggregates of communities

Key Points:

- In planning an evaluation, be sure to identify the level(s) of focus.
- As you move up the levels, effecting change requires more time to achieve.
- As you move up the levels, effecting change requires more resources.
- Outcomes must be measured at the appropriate level of focus.

Adapted from Nancy K. Young, Sidney L. Gardner, and Soraya M. Coley (1994). Getting to Outcomes in Integrated Service Delivery Models. In National Center for Service Integration, Resource Brief 7, Falls Church, Virginia.

In the current service environment, programs are now asked to be accountable for their outcomes, not just their level of activity. Human service providers need to devote some attention to measuring, understanding and reporting the outcomes of their services. Human service providers also need tools for developing a manageable outcome system that will not create an overwhelming burden on staff time and agency finances.

The pages that follow in this section provide guidance and may also serve as training and review materials for general audiences concerned with issues related to the development of valid and reliable outcome measures for human service programs. First we present guidance on the pros, cons and limitations of program evaluation under the current outcomes paradigm followed by “The ‘Up’ Side of Self Evaluation” and “The ‘Down’ Side of Self Evaluation,” and then the “Top 10 Ways to Make Outcome Evaluation as Useless as Possible.” We then address the process of defining outcomes or operationalization, some common evaluation designs, sampling, and some simple statistics and the concepts of reliability and validity.

Pros, cons and limitations of program evaluation

The following list provides some pros, cons and limitations of program evaluation for you to consider while planning the evaluation of your family development program.

Pros:

- Reduces micro-management
- Concrete evidence of results
- Helps agencies collaborate
- Less wasted effort
- Feedback on effectiveness
- Exposes inadequate resources

Cons

- Inadequate time to show results
- May exclude difficult cases
- Effects not always obvious
- Can lead to labeling of clients
- Less individual agency responsibility

Limitations

- Problem identification but not resolution
- Cannot prevent unethical practice
- Program changes often required

Adapted from Lisbeth Schorr (1994). The Case for Shifting to Results-Based Accountability. In National Center for Service Integration, Resource Brief 7. Falls Church, VA.

The “Up” Side of Self Evaluation

1. Builds staff's investment in the evaluation process (staff become key stakeholders).
2. May keep agency and staff focused on outcomes that the program was designed to achieve.
3. Saves on costs of an external evaluator.
4. Enhances the evaluation's relevance to the program.
5. Can create an ongoing mechanism to monitor/improve the program.
6. Can create developmental opportunities for staff.
7. Results may be presented in more user friendly ways.
8. May create mechanisms for more community involvement.

The “Down” Side of Self-Evaluations

1. Creates extra work load for staff.
2. Requires ongoing commitment to be successful.
3. May not have sufficient expertise.
4. Likely to have reduced scientific rigor.
5. Differences among staff may impede evaluation progress.
6. Internal pressure to “bury” results that make the program look bad.
7. May lack credibility that the external evaluator brings.
8. May not be sufficient for some grants that require an independent evaluation.

Top 10 Ways to Make Outcome Evaluation as Useless as Possible

10. Do not by any means involve people who know anything about the consumers or the services in planning the evaluation (including front line workers or consumers themselves.)
9. Collect more data than you can possibly enter and analyze in a lifetime.
8. Make the data collection procedures so complicated that no one (except *maybe* the evaluator) can remember them.
7. Give your staff mountains of new forms specially designed and only to be used “for the evaluation.”
6. Don't bother to provide any staff training in the data collection procedures or forms...after all, you need to keep costs as low as possible.
5. It follows that you should not factor into staff workloads any additional responsibilities required by the evaluation.
4. Compile all data only when a program has ended, and then hire an evaluator to tell you whether it worked out or not.
3. If you found out that it “didn't work,” either bury the report and say that “results were inconclusive due to small sample sizes and various methodological limitations,” or change the measures that are being used so that long-term comparisons can never be made.
2. If you found out that it “did work,” report results so that only a Ph.D. in statistics can understand them.
1. Since you tried it once and weren't satisfied with the results, conclude that evaluation is too much trouble and that all that really matters are a few “good stories!”

Courtesy M.J. Landsman, 1997, National Resource Center for Family Centered Practice, University of Iowa School of Social Work, Iowa City, Iowa.

Operationalization

A step beyond conceptualization of outcomes, **operationalization** involves the development of operational definitions. It is a process of redefining a concept in specific, precise and measurable terms. Operationalization is a specification of the steps, procedures, or operations that you will go through in actually measuring and identifying the variables you want to observe.

There can be many operational definitions for a particular concept. When defining operational definitions of outcomes, they should reflect the objectives of the program. Operational definitions should also be formulated at the appropriate level of analysis.

For example, it would probably not make much sense to operationally define the success of a local self-sufficiency program as “reducing the state’s FIP enrollment rate by ten percent.” More appropriate of various outcomes of such a program might include:

1. Attendance at 7 of 10 sessions and a score of at least 80% on the final exam.
2. Improvement in understanding of key concepts presented in the program, based on a comparison of pre- and post-test scores (E.g., understanding self-sufficiency).
3. Evidence of utilization of information learned in the program, collected by phone or mail survey of participants one year after completion of the program (i.e., family not on FIP, family not in crisis, family member employed full-time).

Common Evaluation Designs

Here we review five of the most common evaluation designs. An **experiment** occurs when subjects are randomly assigned to treatment and control groups and the same measures are obtained on both groups. When done faithfully, experiments have the strongest ability to make firm conclusions about program’s effects. **Quasi-experiments** use comparison groups to approximate an experiment rather than using random assignment; therefore, they are not as valid as true experiments. Quasi-experiments are popular in field settings due to their key advantage of greater feasibility.

Descriptive studies without comparison or control groups are very common in program evaluation and are called **one-group designs**. The only comparison that can be made is within the group over time. One-group designs are considered the least powerful of designs, but are also the least costly. These evaluations are appropriate for obtaining descriptive information about programs in early stages of development, but lack the ability to draw firm conclusions about program’s effects. **Single system** designs follow a small number of cases with many data collection points over time. These evaluation designs are useful for depicting clinical change. Some variations provide evidence for strong conclusions about program effects. Finally, **case studies** provide an in-depth study of one or more cases. This design is rich with descriptive detail, intends to provide insight into a phenomenon and generate theory where none exists.

Sources: Babbie, (1986), The Practice of Social Research. and Pietrzak, Ramler, Renner, Ford and Gilbert. (1990). Practical Program Evaluation.

Sampling and How to Use It

Sampling is the process of selecting a portion of a population for an experiment. There are two types of sampling called probability and non-probability sampling.

Probability sampling, often known as “random sampling,” means that there is a known probability for each member of the population to have a chance of being selected. This type of sampling leads to greater generalizability of findings, because there is a greater confidence that the sample represents the population from which it is drawn. One thing to remember is that you need to identify the entire population first.

With **non-probability** sampling, there is not a known probability for each member of the population to have a chance to be selected. For that reason, non-probability sampling is easier to use, because it does not require knowledge of the entire population. This type of sampling is often preferred in the study of hidden populations (i.e. drug dealers, stalking victims, gay and lesbian groups) and new areas of research where the goal is exploratory. The drawback to non-probability sampling is that you cannot generalize findings to the larger population with confidence.

Probability

How to use Systematic Random Sampling:

Start with list of names that comprise the population, select every Kth name from the list. K is determined based on the desired sample size relative to the size of the population. Make sure that the list is not ordered in an identifiable way.

How to use Stratified Random Sampling:

Reorganize list based on stratifying variable, then select through random sampling (see above). This is used to obtain a greater degree of representativeness on key subgroups of interest. Disproportionate stratified sampling may also be used to “over sample” from underrepresented groups.

Non-Probability

How to use Snowball Sampling:

Begin with a small number of subjects. Then, ask those subjects to provide information needed to locate additional subjects (e.g., in their social network).

How to use Quota Sampling:

Decide matrix identifying categories and proportions for the sample. Then, collect data from individuals having characteristics of each cell until all cells are full.

Some Useful Simple Statistics

Here are some basic useful phrases and definitions used in statistics. A **frequency distribution** displays of all the values of a variable, represented in numbers and/or percentages.

There are three ways of describing “typical” cases also known as Measures of Central Tendency. These are the mean, median and mode. The **mean** is the same as the mathematical average or the sum of values divided by number of cases. The **median** is the midpoint that divides all cases exactly in half (50th Percentile). The **mode** is the most frequently occurring value of a variable.

Measures of Dispersion are ways of describing the spread of values. These include the range and the standard deviation. The **range** is difference between the highest and lowest value. The **standard deviation** is the position of values relative to the mean and is usually reported with the mean.

Reliability and Validity in Measurement

Reliability refers to consistency in measurement:

Does an instrument yield the same results upon repeated measurement?

Validity refers to accuracy in measurement:

Does the instrument measure what it purports to measure?

Types of Reliability

- Interrater Reliability:** Extent to which an instrument yields similar results using different observers.
- Test-Retest:** Instrument administered to same individual at two points close in time will yield similar results.
- Alternate Form:** Different versions of an instrument yield similar results.
- Split Halves Method:** Divide items into two halves, assess correlation of total scores of the two halves.
- Internal Consistency:** Average correlation of all items in scale (Cronbach's Alpha)

Types of Validity

- Face Validity:** Degree to which instrument appears to measure the concept of interest.
- Content Validity:** Degree to which a measure covers the range of items included within a concept.
- Criterion-Related:** Degree to which a measure corresponds to a second (criterion) measure
- Two Types:*
- Concurrent: Occurring at the same time
 - Predictive: Occurring at a future point in time
- Constructive Validity:** Degree to which a measure relates to other variables theoretically related
- Two-Types:*
- Convergent: Results similar to other means of measuring same concept
 - Discriminant: Measure does not correspond to other measures of different concept

V. Efficient and Effective Outcome Evaluation

Our work on the Local Outcomes Technical Assistance project has led to the development of materials that can assist agencies in creating the most efficient and effective family development outcomes. Outcome measurement can serve multiple purposes including reporting progress to funding agencies that want verification that outcomes are being achieved, using data for other fundraising activities (e.g., grant proposals), reporting progress to stakeholders in the community, using outcomes to keep your program “on track,” and using data for self-evaluation and program improvement (formative evaluation). It is important, however, to maintain the distinction between activity and outcome. The following two examples help articulate these concepts.

Example #1: Happy County Juvenile Diversion Program

<i>Activity (process)</i>	Happy County’s juvenile diversion program provided five hours of counseling per week to high risk youth.
<i>Result (outcome)</i>	85% of high-risk youth participating in the Happy County juvenile diversion program were free of juvenile court involvement.

Example #2: Happy County Community Action Partnership

<i>Activity (process)</i>	50 families came to the HCCAP
<i>Result (outcome)</i>	95% of families served this quarter accessed needed community services.

Activities are relatively easy to measure and report. These are usually counts of what has been done. Outcomes are more difficult to report because we need to identify the desired result, identify a measure of the result, collect data, and then report the result. Some common places to get “stuck” are:

- Understanding differences between outcomes and program activities
- Identifying good measures that are obtainable
- Figuring out how to collect data
- Figuring out how to organize data once you have it
- Figuring out how to report data

The Language of Outcomes

Unfortunately, there is no universal language expressing the concepts related to outcomes. There are differences in terminology by textbook, training program used, field (i.e., social work, education, health), and by funding agency. This means that we need to be multi-lingual when it comes to outcomes. The terms listed below do not have single definitions:

- Performance measures
- Performance outcome measures
- Process measures
- Measures
- Outcomes
- Short term outcomes
- Intermediate outcomes
- Long Term Outcomes
- Impacts
- Benchmarks
- Indicators
- Targets
- Objectives
- Goals

In outcomes training conducted throughout the U.S., we attempted to simplify, some say oversimplify, the terminology above which is frequently associated with logic models. Two side-by-side cells labeled “What happened,” and if something happened, “what did you do?” have been proposed (Richardson, 2002). After all, who’s interested in “what you did” if nothing happened? If nothing could be shown to have happened, we could further reduce the model to only cell #1.

Example #3: Abbreviated Logic Model

1	2
<i>What happened?</i>	<i>What did you do?</i>
<i>Result (outcome)</i>	<i>Activity (process)</i>

Mark Friedman has conducted training for years on developing outcomes. His model separates outcomes into two groups: community-wide outcomes and client (customer) outcomes. These outcomes are distinct from each other, however, they are, *and should be*, related. To illustrate, client outcomes should feed into the desired community level outcomes. In Friedman’s language, outcomes and measures refer to the well-being of whole populations, otherwise known as the community level. On the other hand, performance measures refer to the well-being of client populations, or program level. Some of the terms Friedman uses are listed below (Adapted from Marc Friedman (1996). Workshop: Results and Performance Accountability, Decision-making and Budgeting. Fiscal Policy Studies Institute.)

Measure

A measure which helps quantify the achievement of a result

Examples of this include:

- Rate of low birth-weight babies
- Rate of high school graduation
- Employment

Outcome

A condition of well-being for children, adults, families, or communities

Examples of this include:

- Children born healthy
- Children succeeding in school
- Self-sufficient families

Note that in this terminology, outcomes are not written as measures, but in more global terms.

Another example

Your family development program has set certain outcomes in its plan. These might include: improving self-sufficiency, family stability, health, and safety. These are comparable to Friedman's use of the term "results." Your family development program may also have established measures for the community as a whole, for example, reduce rates of child abuse and neglect, improve housing, and so on. Service providers are asked to provide outcomes for the programs for which they are requesting funding. These program-specific outcomes would be performance measures.

How do they fit together?

Program outcomes fit within the outcomes established for the community. In this way, the services play a role in contributing to community improvement.

Performance measure

A measure of how well the agency or program service delivery is working

Examples of this include:

- Percent of customers served in their language
- Percent of clients off alcohol and drugs
- Percent of families in housed status
- Percent in school or with jobs

The following are some important considerations in choosing outcome measures. Outcome measures should be relevant to the program's activities and interventions, be meaningful to stakeholders, reflect the most important goals of the program rather than every goal, be measurable, have data that are available or accessible within the program's means, have reliable measures (using the same measure over and over again will yield consistent outcomes), have and valid measures (they actually measure what you want them to measure).

Evaluating the Effectiveness of Your measures (SEE WORKSHEET #3).

Friedman proposes three powers or considerations in choosing a measure. Each of the considerations can be rated as either having high or low power:

1. Communication power: does it speak to a range of stakeholders? Does the measure communicate to a broad range of audiences?
2. Proxy power: does the measure present important information about the key goals? Does the measure say something of central importance about the result? Does the measure bring along the rest of the data “herd”?
3. Data power: will the data be available to you as often as you need it? Is high quality data available on a routine and regular basis?

Sample Menu of Outcomes and Measures (SEE WORKSHEET #4)

Outcomes	Measures
A. Improved health of young children.	<ol style="list-style-type: none"> 1. % of children receiving full series of immunizations 2. % of program participants with babies born at low birth weight 3. % of children and families with medical insurance
B. Improved economic well-being of families.	<ol style="list-style-type: none"> 1. % of adults employed 2. % of families living at 100 and 200% of the poverty level 3. % of families living in adequate housing
C. Preventing unnecessary out of home placements.	<ol style="list-style-type: none"> 1. % of children remaining safely at home 2. % of children in foster care who are reunified with their families
D. Achieving permanency for children.	<ol style="list-style-type: none"> 1. % of children with a goal of adoption who are adopted 2. % living with long term guardianship arrangements 3. % of children who are reunified with their families
E. Improved chances for teen parents' social and economic well-being.	<ol style="list-style-type: none"> 1. % of teen parents who do not experience a subsequent pregnancy 2. % of teen parents completing high school or GED 3. % of teen parents who do not use alcohol, drugs, or tobacco
F. Preventing substance use among teens.	<ol style="list-style-type: none"> 1. % of youth who do not use substances 2. % of teens who complete prevention program
G. Improved school attendance.	<ol style="list-style-type: none"> 1. % increase in school days attended 2. % decrease in unexcused absences
H. Improved academic functioning of youth.	<ol style="list-style-type: none"> 1. % improvement in grade point average 2. % youth graduating from high school on time 3. % youth performing at grade level on standardized math and reading tests 4. % of youth remaining in school
I. Reduced juvenile court involvement.	<ol style="list-style-type: none"> 1. % youth who do not experience subsequent involvement with juvenile court 2. % youth who are not adjudicated delinquent
J. Improved school readiness.	<ol style="list-style-type: none"> 1. % children identified with disabilities who receive needed services 2. % eligible children enrolled in accredited child care centers, preschools, Head Start programs
K. Improved child safety.	<ol style="list-style-type: none"> 1. % children who do not experience a founded assessment of child abuse or neglect 2. % of children living in safe housing 3. % of children living in families without domestic violence
L. Improved school readiness.	<ol style="list-style-type: none"> 1. % children identified with disabilities who receive needed services 2. % eligible children enrolled in accredited child care centers, preschools, Head Start programs

**IDENTIFY AND EVALUATION PERFORMANCE MEASURES
(SEE WORKSHEET #5)**

The following are some types of consumer self-reports that can be used in outcomes data collection:

1. Pre and post tests
2. Exit questionnaires
3. Follow-up mail surveys
4. Follow-up phone contacts or visits

Finding data

There are many different ways of obtaining data for outcomes, such as, consumer self-reports, staff reports, agency records and public records, community surveys. Types of agency records include case files, public health data, school records, juvenile court records, Workforce Development records, DHS records. Types of staff data include observations, pre and post rating scales (of clients) by staff, observational measures of families & children, exit questionnaires completed by staff for each client, interviews with staff, follow-up telephone contacts with staff. Some types of community surveys include surveys of community residents (mail or phone), school based surveys, surveys of consumers of a specific service, surveys of a specific population subgroup.

Performance measures

	Quantity	Quality
<i>Input</i>	How much service did we deliver?	How much effect/change did we produce?
Output	How well did we deliver service?	What quality of effect/ change did we produce?

	Quantity	Quality
<i>Input</i>	Number of low-income families served	Percent of families needing self-sufficiency program
Output	Number completed Self-sufficiency program	Percent completed Self-sufficiency program

Write your own Performance Measures (SEE WORKSHEET #6) and Plan Plan for Data Collection (SEE WORKSHEET #7)

Keeping data collection simple

Data collection can be as simple or as burdensome as you want it to be. Here are some guidelines for keeping it simple:

- Collect only what you need, not everything you can think of.
- When you choose your measures, keep accessibility of the data as an important priority.
- Know your reporting requirements, including how often you will be asked to report outcomes to your funders.
- Prepare a written plan for data collection, including the who, what, when, how, and where. Don't leave this to memory alone.
- Learn to use a simple spreadsheet to manage the collection and reporting of data.

Planning for working with the data

When planning for working with the data, begin by knowing your reporting requirements. If you are using the data to report to multiple funders, decide ahead of time what information you are going to present to each funder. If you are unsure of how to set up a data management system, ask for help from someone who has done this before. Setting up a system right the first time will prevent problems down the road.

Reporting Data (SEE WORKSHEET #8)

What kinds of data can you present?

To summarize performance measures for a program, simply report percentages (%) for categorical variables. Categorical variables are measures that fit into distinct grouping such as: is not currently employed full-time or is currently employed full-time (Yes/No). Alternatively, for continuous variables, report means and standard deviations. Continuous variables are measured along a continuum, such as household income.

Examples

If your performance measure is: % of children receiving full series of immunizations, you should report: 75% of children participating in this program received the full series of immunizations by age two.

If your performance measure is: average length of time in residential treatment, you should report: The average length of time for youth in this residential treatment program was 3.4 months (standard deviation= 1.4).

VI. Step by Step Guide for Developing Outcomes of Family Development Programs

DO COMMUNITY NEEDS ASSESSMENT



CLEARLY DEFINE PROJECT GOALS



CLARIFY RESOURCE ALLOCATION



DETAIL PROGRAM ACTIVITIES

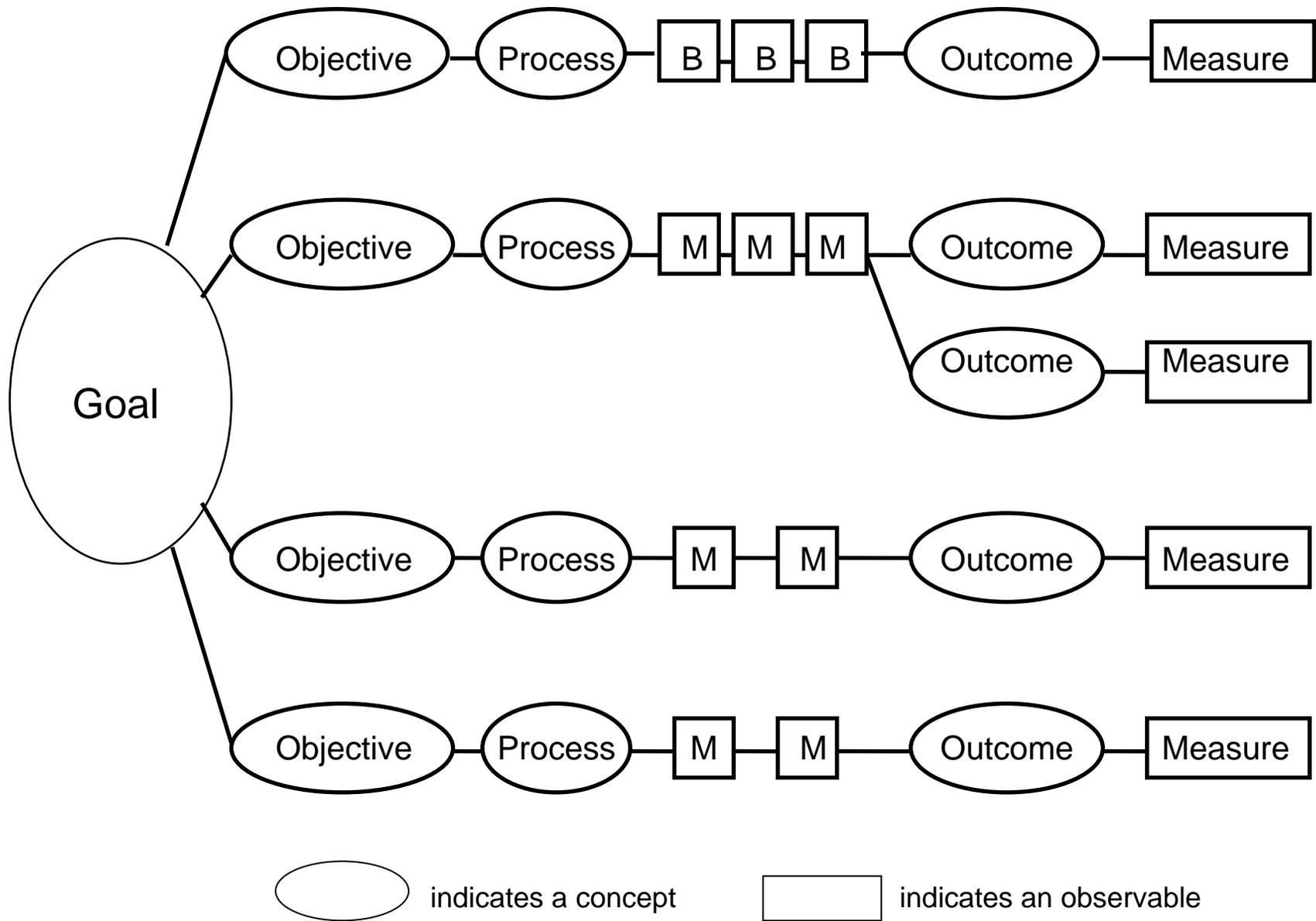


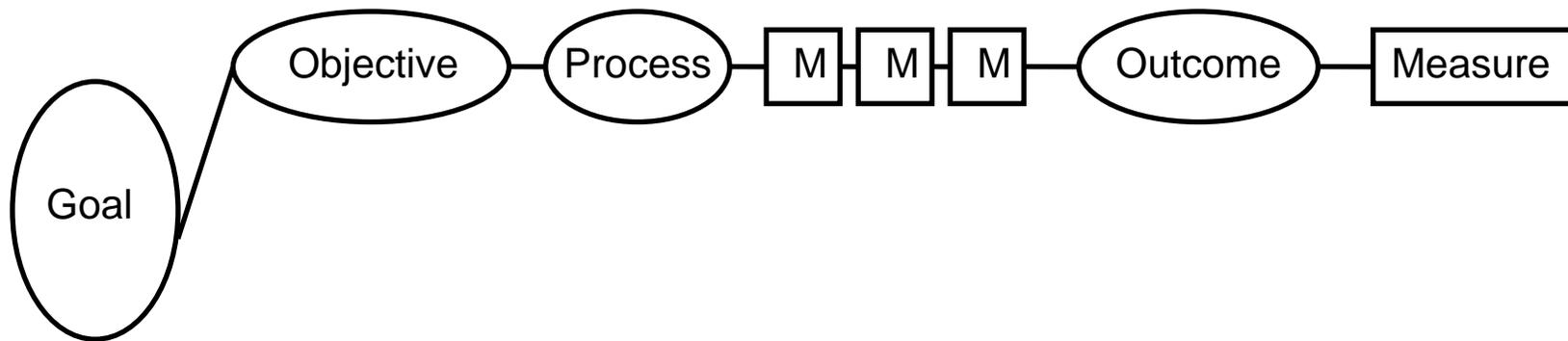
IMPLEMENT PROGRAM



EVALUATE PROGRAM

Adapted from Nancy K. Young, Sidney L. Gardner, and Soraya M. Coley (1994). Getting to Outcomes in Integrated Service Delivery Models. In National Center for Service Integration, Resource Brief 7. Falls Church, VA.





Goal: To help communities to strengthen families

Objective: Increase child well-being by reducing the number of adolescent parents

Process: Establish a teen pregnancy prevention program

Measures: Program active

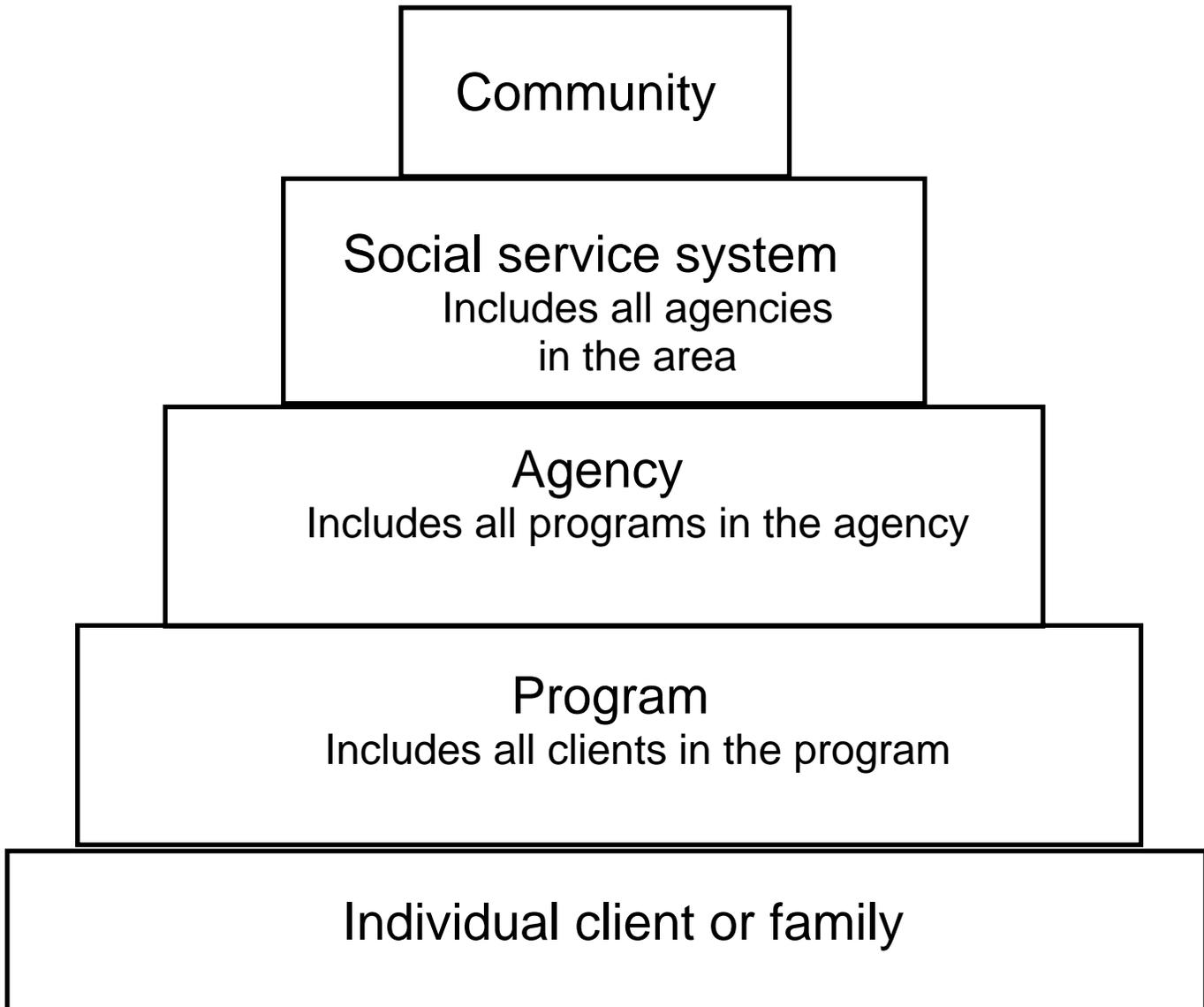
Number enrolled in program

Rate of return to program

Outcome: Fewer families with teenage mothers

Measures: Number of births to women less than 18 yrs old

Levels of Focus



Adapted from Nancy K. Young, Sidney L. Gardner, and Soraya M. Coley (1994). Getting to Outcomes in Integrated Service Delivery Models. In National Center for Service Integration, Resource Brief 7. Falls Church, VA.

1. Sample Outcomes and Measures by Level

	Individual	Program	System	Community
Outcomes	Individual engagement	Community-wide family well-being	Agency cooperation	Cohesiveness
Measure	In-home visitations criteria met Attendance at parent education class	Percentage of at-risk population enrolled Client return rate	Number of referrals Agency participation levels	Resident awareness of services Community group participation rates

2. Sample Outcomes and Measures by Level

	Individual	Program	System	Community
Outcomes	Increased child well-being	Parenting program served client need	Coverage of client needs	Community-wide family well-being
Measure	Child immunizations up to date	Clients improvement in parenting skills Client satisfaction survey	Client satisfaction Provider's assessment met and unmet needs	Reduction in homelessness Increase in school attendance

Characteristics of Good Measures

- It is closely related to the process or outcome it is intended to measure.
- There is consensus within the planning group on its importance.
- Data are easy to identify and obtain.
- If it is a measure, it can be shown to be related to an outcome.
- It is the least ambiguous measure available.
- It cannot be used to negatively effect clients.
- Collection of data provides a positive experience for users and staff.

Adapted from Lisbeth Schorr (1994). The Case for Shifting to Results-Based Accountability. In National Center for Service Integration, Resource Brief 7. Falls Church, VA.

Data Collection Methods

Existing data

- Government statistics
- School records
- Court records

Existing questionnaires/interviews

- Self-administered questionnaires
- Closed-ended interviews
- Open-ended interviews

Constructing questionnaires/interviews

Focus groups

Observation

- Participant
- Observer

Elements of the Evaluation Report

- I. **Executive Summary**
 - A. Short overview of the evaluation
- II. **Introduction**
 - A. Goal of the project
 - B. Purpose of the evaluation
- III. **Methods and Measures**
 - A. Evaluation design
 - B. Outcome measures
 - C. Data collection procedures
- IV. **Results**
 - A. Qualitative and/or quantitative
 - B. Charts and tables
- V. **Discussion**
 - A. Interpretation of results
- VI. **Summary and Recommendations**
 - A. Conclusions
 - B. Future status of the program
 - C. Suggested future research

Useful Introductory Phrases

“Thanks are due to Joe Blotz for assistance with the data analysis and Gail Frink for valuable discussions.”

Translation: *Blotz did the work and Frink explained to me what it meant.*

“It has long been known...”

Translation: *I didn't look up the original reference.*

“Of great theoretical and practical importance...”

Translation: *Interesting to me.*

“It is believed that...”

Translation: *I think.*

“It is generally believed that...”

Translation: *A couple of other people think so too.*

Useful Outcomes and Summary Phrases

“A definite trend is evident.”

Translation: *These data seem practically meaningless.*

“Three of the cases were chosen for detailed study.”

Translation: *The other(s) made no sense.*

“Multivariate analyses showed few additional effects...”

Translation: *Prolonged and expensive quantitative analyses were then conducted to obscure the fact that the outcomes have no generalizable meaning for anyone.*

“While it has not been possible to provide definite answers to these questions...”

Translation: *This was an inconclusive evaluation, but I still hope to get paid for it.*

“A highly significant area for exploratory study...”

Translation: *A totally useless topic suggested by the evaluation advisory committee.*

“It is clear that much additional work will be required before a complete understanding of the phenomenon is possible...”

Translation: *I don't understand it.*

Adapted from Natalie Kishchuk and Shelley Borys. (1995). Analysis of Qualitative Data in Program Evaluation. Workshop, International Evaluation Conference.

Section 3. Outcomes Development Tools

VII. Worksheets

1. Assessment of Current Evaluation Activities
2. Assessment Using Evaluation Outcomes
3. State your program, goals & target population
4. Situation your program in the community context
5. Identify and evaluate performance measures
6. Write your own performance measures
7. Planning for data collection
8. Reporting data
9. Cultural competency questions
10. Case study #1
11. Case study #2

VIII. Glossary

IX. List of Instruments

X. Bibliography/References

VII. Worksheets

Worksheet 1: Assessment of Current Evaluation Activities

Description

1. Briefly describe your current evaluation activities.

2. Who designs or has input into your current evaluation activities?

Program managers/directors?	Y	N
Program supervisors?	Y	N
Line staff?	Y	N
Consumers (service recipients)?	Y	N
An external evaluator?	Y	N
An advisory committee?	Y	N
Others?	Y	N

3. Are you using:

Interviews or focus groups with consumers?	Y	N	Consumer surveys?	Y	N
Interviews or focus groups with providers?	Y	N	Provider surveys?	Y	N
Management information system data?	Y	N	Case record reviews?	Y	N
Case studies?	Y	N	Pre and post tests?	Y	N
Control or comparison groups?	Y	N	Other methods?	Y	N

Strengths

4. What works well with your current evaluation?

5. What is problematic or what would you like to change about your current situation?

Worksheet 2: Assessment Using Evaluation Outcomes

How do you currently use your evaluation results?

1. Within the agency:

Distribute internal reports to:

Line staff	Y	N
Supervisors	Y	N
Program Managers	Y	N
Discuss evaluation results at staff meetings	Y	N
Others:	Y	N

2. With consumers/community:

Share results with consumers	Y	N
Disseminate reports throughout community	Y	N
Disseminate results through newsletters	Y	N
Public forums	Y	N
Others:	Y	N

3. With funders/sponsors:

Progress reports for family development programs	Y	N
Reports to funders	Y	N
Grant development (potential funders)	Y	N
Others:	Y	N

4. How would you want to change the way(s) in which you currently use evaluation results?

5. What would you need in order to accomplish these changes?

Worksheet 3: State your program, goals and target population

Use this worksheet to develop your program, its goals and target population:

1. What is the **program** you're working on today to develop outcomes?

2. What are the most important **goals** of this program? (Don't use measures here, write more global goals). For example:

- a. To increase family self-sufficiency
- b. To increase access to affordable housing
- c. To decrease family crisis

3. What is the **target population** served by the program? (Who are you trying to reach?)

Worksheet 4: Situate your program in the community context

1. What family development program outcome best fits your program? (For example, improving food availability or reducing hunger.)

2. What measure is the Community Action Agency using to measure community level progress toward this outcome? (For example, reduce rates of food assistance benefits.)

3. What performance measure might your program use to demonstrate your program's contribution to this community level outcome? (For example, 80% of program participants will have adequate food supply.)

Worksheet 5: Identify & evaluate performance measures

Worksheet 5 allows you to view the performance measure in terms of three types of power: communication, proxy and data.

Performance measure	Communication power (speaks to range of stakeholders)	Proxy power (important information)	Data power (data are available as needed)
1.			
2.			
3.			

Worksheet 6: Write your own performance measures

Using the performance measures for your program, fill in the Input, Output, Quantity and Quality.

	Quantity	Quality
<i>Input</i>		
Output		

	Quantity	Quality
<i>Input</i>		
Output		

	Quantity	Quality
<i>Input</i>		
Output		

Worksheet 7: Planning for data collection

Use this worksheet to help make a comprehensive plan for data collection including the what, where, who, and how.

1. **What** data are you collecting? (The performance measure)

2. **Where** can the data be found?

3. **Who** will collect the data?

4. **When** will the data be collected?

5. **How** will the data be collected?

6. **Where** do the data go once they are collected?

Worksheet 8: Reporting data

Use this worksheet to practice how to report your programs' outcomes and evaluation report.

1. Your performance measure is:

2. In your report you will state: (It's okay to make up the number for this exercise!)

Worksheet #9: Cultural Competency Questions

The following questions about process planning, implementation and outcomes are designed to address cultural competency in your program evaluation.

A. Process Planning

- | | | |
|---|---|---|
| 1. Are culturally diverse groups involved in the program planning process? If so, to what extent? | Y | N |
| 2. Does the agency (or agencies) support cultural diversity through mission statements, organizational structure, staffing, etc.? | Y | N |
| 3. Are services located in culturally diverse communities? | Y | N |
| 4. Has the program built relationships with key community members and organizations supporting minority groups? | Y | N |

B. Process Implementation

- | | | |
|---|---|---|
| 1. Does the staff composition reflect the community profile? | Y | N |
| 2. Does the staff receive training in cultural competency and/or diversity? | Y | N |
| 3. Are culturally diverse groups referred or recruited as clients/consumers? | Y | N |
| 4. Do culturally diverse groups complete the program or drop out at higher rates? | Y | N |
| 5. When collecting data, are clients asked directly about their racial/ethnic identity or does staff make assumptions on the basis of physical characteristics, surnames, etc.? | Y | N |

More Process implementation Questions...

- | | | |
|---|---|---|
| 6. If standardized instruments are used, have they been checked for previous use with culturally diverse populations? | Y | N |
| 7. Does the evaluation process have bilingual capabilities (instruments, interviewers, etc.)? | Y | N |
| 8. Are the programs's philosophical or treatment approaches consistent with the values of culturally diverse groups? | Y | N |

C. Outcomes Evaluation

- | | | |
|--|---|---|
| 1. Are defined program outcomes appropriate to diverse cultural groups? | Y | N |
| 2. Is achievement of program outcomes similar across different racial/ethnic groups? | Y | N |
| 3. Are differences in program outcomes attributable to differences in starting points for culturally diverse groups? | Y | N |
| 4. Is consumer satisfaction similar across different cultural groups? | Y | N |

Evaluation Design

Case Study #1

A community action agency offers a family self-sufficiency program for low-income families. A caseworker works with these families in the home or agency to teach XY and Z.

The evaluation plan calls for three components: 1) a family self-sufficiency inventory is administered at the time the family is referred and at the time the family completes the program; 2) household income reports are collected on a quarterly basis; and 3) at the last scheduled service visit, the caseworker gives the family a satisfaction questionnaire to fill out and return in a stamped, addressed envelope.

Results at the end of the first year show that scores on the self-sufficiency inventory increased significantly from referral to program completion and that household income increased by 5%. Thirty percent of the families returned the client satisfaction survey, and the most frequent response was “somewhat satisfied.”

Based on these evaluation results, what can you say about the program’s effectiveness? Write down the points that support the program’s effectiveness or ineffectiveness?

Evaluation Design Case Study #2

The same self-sufficiency program had been operating for several years and staff felt confident about it. So when a well known university professor invited the agency director to participate in an experimental study to scientifically test the program's effectiveness the director agreed (despite the objections of direct service staff who felt this to be unethical, and besides were not consulted about the plan until after the fact). Under the experimental design, families who were referred and determined to be eligible for the program would be randomly assigned to the family aide program or to the agency's casework services, which were believed to be less intensive. A program supervisor with responsibility for both the family aide and casework services was given the task of overseeing the random assignment of families to these programs. The same self-sufficiency inventory was administered to both groups at referral and six months later, and income was tracked for the same period of time.

Evaluation results after one year showed that the experimental and control groups improved to about the same degree and had similar rates of XYZ. Those in the caseworker groups reported higher levels of satisfaction with the program. The agency director was quite distressed over the findings, as the research had proven the highly publicized self-sufficiency program to be no more effective than the traditional self-sufficiency program that had been in operation at the agency for about twenty years.

Do agree with the agency director's conclusions about the evaluation results? Has the program been proved ineffective, or are there other factors that might explain these results?

VIII. Glossary

Goal: A broad statement of what we plan to do about an identified need. It is not intended to list specific outcomes.

Objective: A more specific part of the overall goal. It breaks the goal down into smaller pieces.

Process: The means by which goals and objectives are reached. This includes elements such as program design, implementation plans, and resource development.

Process Measure: An observable event or behavior that can be used to measure a process such as effectiveness of design and timeliness of implementation.

Outcome: A specified effect which is expected at the end of a project. It is concrete and measurable.

Outcome Measure: An observable, usually countable, event or behavior that can be used to measure an outcome.

IX. Sample List of Available Instruments

Family Functioning:

Family Environment Scale

Population: Families

Format: 90 item self-report questionnaire

Source: Consulting Psychologists Press, 3803 East Bayshore Road, Palo Alto, CA 94303

Family Assessment Device

Population: Families

Format: 60 item self-report questionnaire

Source: Family Research Program, Butler Hospital, 345 Blackstone Boulevard, Providence, RI 92906

Index of Family Relations

Population: Family members (12 and over)

Format: 25 item self-report questionnaire

Source: Nurius, P. S. and Hudson, W. W. (1993). *Human Services Practice, Evaluation and Computers*. Pacific Grove, CA: Brooks/Cole Publishing

Child Well-Being Scales

Population: Families

Format: 43 items, administered by case worker or can be used with case records

Source: Magura, S. and Moses, B. S. (1986). *Outcome Measures for Child Welfare Services*. Washington, D. C.: The Child Welfare League of America.

Parent Functioning:

Adult-Adolescent Parenting Inventory

Population: Parents

Format: 32 item self-report questionnaire

Source: Family Development Resources, Inc., 767 Second Avenue, Eau Claire, WI 54703

Child Abuse Potential Inventory

Population: Parents

Format: 160 item self-report questionnaire

Source: PSYTEC Inc., P.O. Box 564, DeKalb, IL 60115

Parenting Stress Inventory

Population: Parents of children up to 10 years old

Format: 101 item self-report questionnaire

Source: Pediatric Psychology Press, 320 Terrell Road West, Charlottesville, VA 22901

Adapted from Peter J. Pecora, Mark W. Fraser, Kristine E. Nelson, Jacquelyn McCroskey, and William Meezan (1995). Evaluating Family-Based Services. New York: Aldine De Gruyter.

Social Support:

Family Relationship Index

Population: Families

Format: 30 item self-report questionnaire

Source: Consulting Psychologists Press, 3803 Bay Sore Road, P. O. Box 10096, Palo Alto, CA 94303

Milardo Social Support Inventory

Population: Families

Format: 25 item rating scale administered by caseworker

Source: Professor Robert Milardo, Merrill Hall, School of Human Development, University of Maine at Orono, Orono, Maine 04469

Family Support Scale, Inventory of Social Support

Population: Families

Format: Two short instruments administered by caseworker

Source: Dunst, C. J., Trivette, C. M., and Deal, A. G. *Enabling and Empowering Families: Principles and Guidelines for Practice*. Cambridge, MA: Brookline.

Child Development:

Denver Developmental Screening Test

Population: Children age 0-6

Format: 105 item "pick and choose" test

Source: Ladoca Publishing Foundation, Laradon Hall Training and Residential Center, East 51st Avenue and Lincoln Street, Denver, CO 80216

Self-esteem:

Rosenberg Self-Esteem Scale

Population: Individuals (high school through adult)

Format: 10 item Guttman Scale

Source: Dr. Morris Rosenberg, Department of Sociology, University of Maryland, College Park, MD 20742

Substance Abuse:

CAGE Questionnaire

Population: Individuals

Format: 4 item self-report inventory

Source: Ewing, J. A. (1984). Detecting alcoholism: The CAGE questionnaire. *Journal of the American Medical Association*, 252, 1905-1907.

X. Bibliography:

Outcomes in Family Development

Alter, C., & Murty, S. (1997). Logic modeling: A tool for teaching practice evaluation. *Journal of Social Work Education*, 33(1), 103-117.

American Humane Association. (1998). *Assessing outcomes in child welfare services: Principles, concepts, and a framework of core outcome indicators*. Englewood, CO: The Casey Outcomes and Decision-Making Project. (Available at <http://www.caseyoutcomes.org>)

Babbie, E. (1986). *The practice of social research*. Belmont: Waldworth.

Benson, D. (1999). *Return on Investment: Guidelines to determine workforce development impact*. Worthington, OH: Appropriate Solutions, Inc.

Berk, R.A. & Rossi, P.H. (1990). *Thinking about program evaluation*. Newbury Park, CA: Sage Publications, Inc.

Bruner, C., et al. (1994). A framework for measuring the potential of comprehensive service strategies. Des Moines, IA: National Center for Service Integration, Resource Brief 7, p. 31.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D., & York, R.L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

Deming, W.E. (1986). *Out of the crisis*. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Dubrow, M. & Richmond, F. (1997). *Return on investment*. LeMoyne, PA: Center for Schools and Communities.

Endres, J., Richardson, B., & Sherman, J. (2000). *Assessment, case management and evaluation with a network of family resource centers using the California matrix*. Family Support America Annual Conference. Chicago, IL.

Endres, J., Richardson, B., & Sherman, J. (1999). *Testing the reliability/validity of the California matrix*. Palo Alto: Packard Foundation, Final Report to the Organizational Effectiveness Program.

Franklin, J.L., & Thrasher, J.H. (1976). *An introduction to program evaluation*. New York: John Wiley and Sons, Inc.

Friedman, M. (1995). *From outcomes to budgets--an approach to outcome based budgeting for family and children's services*. Washington, DC: Center for the Study of Social Policy.

Friedman, M. (1996). Workshop: *Results and performance accountability, decision-making and budgeting*. Fiscal Policy Studies Institute.

Fulbright-Anderson, K., Kubisch, A.C., Connell, J.P. (1998). *New approaches to evaluating community initiatives. Vol. 2: Theory, measurement and analysis*. Washington D.C.: Aspen Institute.

Gore, A. (1991). *The Gore report on reinventing government*. New York: Times Books.

Government Performance and Results Act: Background Information. Retrieved August 7, 2003 from <http://www.ombwatch.org/article/articleview/406/1/90/>.

Guba, E.G., & Lincoln, Y.S. (1989). *Fourth generation evaluation*. California: Sage Publications.

Hernandez, M. (2000). Using logic models and program theory to build outcome accountability. *Education and Treatment of Children*, 23(1), 24-40.

Hunnemann, E. & Richmond, F. (1997). *Positive outcomes*. Harrisburg, PA: Missouri Community Services Block Grant Colloquium.

Hunnemann, E. & Richmond, F. (1996a). *Positive outcomes: What every board member needs to know about outcomes*. Harrisburg, PA: Community Action Association of Pennsylvania Management and Technical Assistance Publication No. 2.

Hunnemann, E. & Richmond, F. (1996b). *Positive outcomes. The transition to results oriented management and accountability: The issue is survival, not compliance*. Harrisburg, PA: Community Action Association of Pennsylvania Management and Technical Assistance Publication No. 4.

Jakopic, J. (1999). *Implementing results oriented management and accountability for the community services block grant (CSBG): A preliminary report on state efforts*. Washington, DC: National Association for State Community Service Programs.

Jakopic, J. (1999). MMDC Communication.

Julian, D.A., Jones, A., & Deyo, D. (1995). Open systems evaluation and the logic model: Program planning and evaluation tools. *Evaluation and Program Planning*, 18(4), 333-341.

Kamber, G., Knowlton, G., Alwin, R., Murphy, S. (2003). *First annual report of performance outcomes from the community services block grant program (FY 2001)*. Washington DC: National Association for State Community Services Programs.

Kirk, R. S., & Ashcraft, K. R. (1998). *User's guide for the North Carolina Family Assessment Scale (NCFAS): Version 2.0*. Chapel Hill, NC: Jordan Institute for Families, School of Social Work, University of North Carolina.

Kishchuk, N. & Borys, S. (1995). Analysis of Qualitative Data in Program Evaluation. *Workshop, International Evaluation Conference*.

Landsman, M.J. (December 1999). Presentation: *Working With Outcomes*. University of Iowa School of Social Work. Iowa City, Iowa: National Resource Center for Family Centered Practice.

Landsman, M. & Richardson, B. (1998). Outcomes Consultation: Lessons from the Field (Part I). *Prevention Report (2)*, 9-11.

Magura, S., & Moses, B.S. (1986). *Outcome measures for child welfare services: Theory and applications*. Washington DC: Child Welfare League of America.

Millar, R. & Millar, A. (eds.) (1981). *Developing client outcome monitoring systems: A guide for state and local social service agencies*. Washington, DC: The Urban Institute.

Monitoring and Assessment Task Force, Department of Health and Human Services. (1999). *ROMA guide: results oriented management and accountability*. Washington DC: Office of Community Services, Department of Health and Human Services.

Mordock, J.B. (2002). *Managing for outcomes: A basic guide to evaluation of best practices in the human services*. Washington DC: Child Welfare League of America Press.

National Resource Center for Family Centered Practice. (1997). *The Downside to Evaluation*. University of Iowa School of Social Work.

National Resource Center for Family Centered Practice. (1997). *The Upside to Evaluation*. University of Iowa School of Social Work.

Neigher, W., Ciarlo, J., Hoven, C., Kirkhart, K., Landsberg, G., Light, E., Newman, F., Struening, E.L., Williams, L., Windle, C., & Woy, J.R. (1982). Evaluation in the community mental health centers program: A bold new reproach? *Evaluation and Program Planning*, 5, 283-311.

Patton, M.Q. (1997). *Utilization-focused evaluation: The new century text*. (3rd ed.). Thousand Oaks, CA: Sage Publications Inc.

Patton, M.Q. (1982). *Practical evaluation*. Newbury Park, CA: Sage Publications, Inc.

Pawson, R. & Tilley, N. (1997). *Realistic evaluation*. Thousand Oaks, CA: Sage Publications, Inc.

Pietrzak, J., Ramler, M., Renner, T., Ford, L. & Gilbert, N. (1990). *Practical program evaluation*. Newbury Park, CA: Sage.

Power, M., Knowlton, G., Spade-Aguilar, M. (2003). *The community services block grant statistical report: FY 2001 executive summary*. Washington DC: The Center for Community Action Research.

Available at: <http://www.acf.dhhs.gov/programs/ocs/csbq/html/3.htm>

Rice, J.M. (1897). The futility of the spelling grind. *Forum*, 23, 163-172.

Richardson, Brad (2002). *Using outcomes measures*. Presentation to the Florida Department of Children and Families, Permanency and Adoption Division, Fort Meyers, FL, June 13-15.

Richardson, B. (2001). *Outcome based programs and evaluation*. Workshop presented at the Florida Department of Children and Families District 8 Promoting Safe and Stable Families Conference, Fort Myers, FL.

Richardson, B., Graf, N., Huff, D. (2001). *Final report: Local outcomes technical assistance project*. Iowa City, IA: University of Iowa School of Social Work, National Resource Center for Family Centered Practice.

Richardson, B. & Huff, D. (2001). Using outcomes in decision making. *Prevention Report (2)*, 8-12.

Richardson, B. & Landsman, M. (2001). *Stress Free Outcome Evaluation*. Presentation to the Fourth National Training Institute. National Resource Center for Family Centered Practice: Developing Families and Communities. Deerfield Beach, FL.

Richardson, B., & Landsman, M. (1999). Outcomes consultation lessons from the field (Part II): The automated assessment of family progress. *Prevention Report (1)* 26-29. Iowa City, IA: The University of Iowa, School of Social Work, National Resource Center for Family Centered Practice.

Richardson, B., & Landsman, M. (1996a). *Evaluating a Model of Head Start and Other Concurrent Interventions in Poverty and Homelessness: Partnerships among Researchers, Practitioners and Policymakers. Proceedings of Head Start's Third National Research Conference*. Washington, DC: Administration on Children, Youth and Families, Dept. of Health and Human Services, with Columbia School of Public Health.

- Richardson, B., & Landsman, M. (1996b). *Head Start Outcomes for Homeless Families and Children: Evaluation of the HACAP Homeless Head Start Demonstration Project*. Iowa City, IA: University of Iowa, National Resource Center for Family Centered Practice.
- Rivlin, A.M. (1971). *Systematic thinking for social action*. Washington DC: Brookings Institution.
- Roethlisberger, F.J. & Dixon, W.J. (1939). *Management and the worker*. Cambridge, MA: Harvard University Press.
- Rossi, P.H., & Freeman, H.E. (1982). *Evaluation: A systematic approach*. Beverly Hills, CA: Sage Publications.
- Rutman, L. (Ed.) (1977). *Evaluation research methods: A basic guide*. Beverly Hills, CA: Sage Publications, Inc.
- Schalock, R.L. (2001). *Outcome-based evaluation*. (2nd ed.). New York: Kluwer Academic/Plenum Publishers.
- Schene, P. (1995). *Outcome measures in public child welfare*. New York: The Edna McConnell Clark Foundation.
- Schorr, L. (1994). The Case for Shifting to Results-Based Accountability. In *National Center for Service Integration, Resource Brief 7*. Falls Church, VA.
- Scriven, M. (1967). The methodology of evaluation. *AERA Monograph Series in Curriculum Evaluation*, 1, 39-83.
- Sichel, J.L. (1982). *Program evaluation guidelines: A research handbook for agency personnel*. New York: Human Sciences Press, Inc.
- Smith, E., & Tyler, R.W. (1942). *Appraising and recording student progress*. New York: Harper & Row.
- Stake, R.E. (1967). The countenance of educational evaluation. *Teachers College Record*, 68, 523-540.
- Suchman, E.A. (1967). *Evaluation research: Principles and practice in public services and social action programs*. New York: Russell Sage Foundation.
- TA Brief Volume 2, Number 2. (Spring 1996.) The Technical Assistance Center for the Evaluation of Children's Mental Health Systems, Boston, MA: Judge Baker Children's Center.

Thompson, M.S. (1982). *Decision analysis for program evaluation*. Cambridge, MA: Ballinger Publishing Company.

The Urban Institute. (2003). *Developing Community-wide Outcome Indicators for Specific Services*. Series on Outcome Management for Nonprofit Organizations. Washington, DC: The Urban Institute.

U.S. Department of Health and Human Services, Centers for Disease Control and Prevention (1999). Framework for program evaluation in public health, *Morbidity and Mortality Weekly Report*, 48 (RR-11).

U.S. General Accounting Office (1975). *Exposure draft: Evaluation and analysis to support decision making*. Washington, D.C.: U.S. General Accounting Office.

United Way of America (1996). *Measuring program outcomes: A practical approach*. Alexandria, VA: United Way of America.

Weiss, Carol H. (1972). *Evaluation research: Methods for assessing program effectiveness*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Wholey, J.S., Scanlon, J.W., Duffy, H.G., Fukumoto, J.S. & Vogt, L.M. (1970). *Federal evaluation policy: Analyzing the effects of public programs*. Washington, DC: The Urban Institute.

Young, N.K., Gardner, S. L., & Coley, S.M. (1994). Getting to Outcomes in Integrated Service Delivery Models. In *National Center for Service Integration, Resource Brief 7*, Falls Church, Virginia.