Evaluation Plan

PROGRAM	SCHOOLS/SITE	ADMINISTRATORS	#/TYPE PARTICIPANTS	INSTRUMENTS	ADMIN. TIME
CONNECTIONS (In-School Program) 16 Weeks – 10/26/19 8 Weeks - January	School 1 School 2 School 3 School 4 School 5 School 6 School 7	Lead T.A., T.A. School SW MSW intern (lead TA does attendance and class documentation)	Estimated 70 students 10 High School 60 Elem. School	Student Survey HS Student Svy **T.A. Class Documentation** Attendance Form Att. DataBase	Last or 2nd to last session Every session
ACTIVE LEARNING **January - April**	School 3		170 1st/2nd graders 280 3rd – 5th graders Every class has a group	Student Survey – Early Grades Student Survey – Upper Grades Teacher Survey	April April April
FAMILY CIRCLE 1 12 sessions, 10/4/19 – Dec., Visit - meal 16 Sessions, Jan – May, 2 Visits - meals	Site 1		9 – 12 men	Program Survey with event questions Other CG interviews	After visit Fall and Spring) Via phone - May
FAMILY CIRCLE 2 1. Mother's Day (event but no meal) 2. July Weekend (event + meal) 3. Holiday visit (like Mother's Day event)	Site 2		About 20 women, including those who are new to the program and continuing participants Special Assistants –2 to 3	Mother's Day event survey Weekend satisfaction svy Holiday event survey Other CG interviews	Spring 2019 July 2018 (done) July 2019 (for future reporting) December 2018 Via Phone - May
SHARING STORIES Programs start week of 9/17 - June	Site 2		25 to 30 (Everyone does 1 focus)		Day after performance (Thursday)

Making Data Collection Decisions

Method	Advantages	Disadvantages	Decisions
SURVEYS (Several commercially available, or unique instruments can be developed)	Easy to quantify and summarize results; quickest and cheapest way to gather new data rigorously, useful for large samples, repeated measures, comparisons between units and to norms/targets; **Good for studying attitudes and perceptions – can also collect some behavioral reports.**	Hard to obtain data on behavior, context shaping behavior (attribution). Not suited for subtle, sensitive issues. Surveys are impersonal and difficult to construct. Must address language and administration challenges; must avoid nonresponse, biased or invalid answers, over interpretation with small samples.	Who gets surveyed (sampling)? How will confidentiality be maintained? Validity of self-assessment? What are standards of desirability? Need for repeated measures - what intervals?
INTERVIEWS (Structured, semi-structured, intercept)	Readily cover many topics and features; can be modified before or during interview; can convey empathy, build trust; rich data; provide understanding of respondents' viewpoints and interpretations. **Good for studying attitudes and perceptions – can also collect some behavioral reports.**	Expensive, sampling problems in large programs; respondent and interviewer bias; non-comparable responses; time consuming to analyze and interpret responses to open-ended questions. Training and protocols required to conduct.	Who gets interviewed (sampling)? How will confidentiality be maintained? Validity of self-assessment? What are standards of desirability? Need for repeated measures - what intervals?
OBSERVATIONS (Participants during program sessions, participants in other settings)	Rich data on hard-to-measure topics (e.g., actual practices, behaviors). Behavioral data independent of self-descriptions, feelings, opinions; data on situational, contextual effects. **Good for studying program implementation and some behavioral changes.**	Constraints on access (timing, distance, objections to intrusion, confidentiality, safety); costly, time-consuming; observer bias, low inter-observer reliability; may affect behavior of people observed; hard to analyze, interpret, report data; may seem unscientific. Training and protocols required to conduct.	What subjects will be observed How many at which levels? Need for repeated measures - what intervals?
RECORD REVIEW (E.g., program records, school records, case management records)	Nonreactive; often quantifiable; repeated measures show change; credibility of familiar or standardized measures (e.g., birthweight, arrest incidents, drug test results, staff or parent assessment results); often cheaper and faster than gathering new data; can include data from other independent sources. **Good for determining (behavioral) status.**	**Access,** retrieval, analysis problems can raise costs and time requirements; validity, credibility of sources and measures can be low. Definitions must be determined prior to use, are often externally determined, can not be customized; need to analyze data in context; limited data on many topics.	Which documents? How can access be obtained? Need for repeated measures - what intervals?

Anita Baker, Evaluation Services www.evaluationservices.co

Making Data Collection Decisions

Method	Validity	Reliability	Available Resources	Cultural Appropriateness
SURVEYS	**LOW** No opportunity for clarification Participants often choose responses other than those provided. Participants may not want to report private behavior. Participants may not be aware of their own actions, behaviors or attitudes.	**HIGH** Administration is consistent from one individual to next. Standard response choices provide consistent range of responses. Little opportunity for data collector to influence results.	**ECONOMICAL** Mass distributed. Costs based on number of mailings, use of phone or mail, incentives.	**VARIED** Best for literate, middle class American-born populations. Particularly bad for immigrants and refugees.
INTERVIEWS	**HIGH** Can clarify questions and probe for more in-depth responses, Personal interaction can establish rapport for open discussion. Focus groups can foster discussion and sharing. Focus groups can clarify individual viewpoints through dialog with others.	**LOW** Interviews are unique based on comments of respondents; different questions and probes likely to be used.	**MODERATE** Individual interviews: moderate expense. Focus group: low to moderate expense.	**STRONG** Individualized interviews work well when paper formats are threatening or invasive and when behavior or attitudes pose a problem. Focus groups work well when the group opinion is the cultural norm.
OBSERVATIONS	**HIGH** Observers can directly observe behavior which may not be accurately reported otherwise. Observers can directly observe behaviors which have standards developed by professionals or institutions.	**MODERATE** Observers need structured protocols for coding their observations. Less structured observer formats reduce reliability because different observers may reach different conclusions.	**MODERATE - EXPENSIVE** Time is required in order to observe behaviors. This can be mitigated by using "natural observers."	**MODERATE** Cultural differences in behavior may be misinterpreted.
RECORD REVIEW	**LOW to MODERATE** Not really designed to measure, rather to document/record	**LOW to HIGH** Depends on whether there are standards for record keeping.	**ECONOMICAL** Data are part of the service delivery process and usually already exist. (Use of case records for evaluation requires up front planning). Some issues of access, confidentiality.	**VARIED** Depends on service delivery, appropriateness of program. May over or under-represent certain groups due to bias.

Survey Questionnaire Development and Assessment*: Step-by-Step

1. Identify the key issues you wish to explore or 'test' via the survey. Review available literature, including proprietary sources, to determine if there are good surveys or items that already exist to measure the key issues.

*2. Convert these key issues into questions and remember to:

- **State the question in very specific terms, using language that is appropriate for the target population.**
- **Use multiple questions to sufficiently cover the topic.**
- **Avoid 'double-negatives'**
- **Avoid asking multiple questions in one**
- **Be sure response categories match the question, are exhaustive and don't overlap.**

*3. Determine what other data are needed for analytical purposes. [Demographics, other background, contact information.]

4. Determine how the questions will be ordered and formatted and be sure to include directions for responses.

5. Have survey instrument reviewed by others including representatives from the target group

A Note about Response Choices

1) **There is no "right" number of choices, but keep it manageable. Even-numbers of responses prevent "waffling"**

2) **Answers typically are either forced choice (yes/no, descriptive) or likert-like "scales,"**

   Temporal *(e.g., never, sometimes, most of the time)
   ***Qualitative (e.g., poor, fair, good, excellent)**

3) **Answer order matters and should be consistent (pos to neg., or neg to pos.)**

4) **Response directions are critical.**

   *** MO (mark one)**

   *** MATA (mark all that apply)**

Designing surveys is complicated and time consuming! Use Caution.

TYPES OF SURVEYS

Mail Surveys (must have correct addresses and return instructions, must conduct tracking and follow-up). Response is typically low.

Electronic Surveys (must be sure respondents have access to internet, must have a host site that is recognizable or used by respondents; must have current email addresses). Response is often better.

Web + (combining mail and e-surveys). Data input required, analysis is harder.

Phone Surveys (labor intensive and require trained survey administrators, access to phone numbers, usually CATI software). Response is generally better than mail**,** but must establish refusal rules.

Staged Surveys (trained survey administrators required, caution must be used when collecting sensitive info). Can be administered orally, multiple response options possible, response rates very high.

Intercept Surveys (require trained administrators). Refusal is high.

STRATEGIES TO INCREASE SURVEY RESPONSE

Write a good survey and tailor administration to respondents.

Advertise survey purpose and administration details in advance.

Carefully document who receives and completes surveys.

Aggressively follow-up. Send reminders.

Consider using incentives.

Make response easy.

Remember: Non-response bias can severely limit your ability to interpret and use survey data

How Big Should Your Sample Be?

The number of program participants will determine whether to include everyone in the evaluation or select a sample, i.e., a smaller group who can represent everyone else and from whom we can **generalize**.

The sample should be as large as a program can afford in terms of time and money. The larger the sample size (compared to the population size), the less error there is in generalizing responses to the whole population -- i.e., to all cases or clients in a program.

1st RULE OF THUMB: if the population is less than 100, include them all (and strive to get an 80% response rate); if the population is bigger than 100 select a **probability sample**. (See your guidebook for sampling strategies.)

Probability samples allow you to calculate the likely extent of the deviation of sample characteristics from population characteristics. **Sampling Error** is the term used to refer to the difference between the results obtained from the sample and the results obtained if data had been collected from the entire population.

The objective when drawing samples is to decrease sampling error and to assure confidence that the results are reliable. 2nd RULE OF THUMB: a common standard for program evaluation is 95% confidence with a sampling error of ± 5%. In English that means that you believe that 95 percent of the time the results from your sample, would be off by no more than 5% as compared to the results you would have gotten if you had collected data from everyone.

<u>It is the absolute size of the sample rather than the ratio of sample size to population size that affects the sampling error</u> (Comer and Welch, 1988, p. 192). Sample sizes for varying population sizes and differing sampling error rates have been calculated (see following page). If you wish for more precision use the following calculation (for 95% confidence, 5% error).

$$n=385 \div ((1+ (385/N))$$

Example: If your population is known to have 472 members, then a sample of 212 would be necessary to ensure 95% confidence with no more than 5% error.
$385 \div ((1+ (385/472)) = 212$

Relationship Between Sample Sizes and Sampling Error

Sample Sizes (n) @ 95% Confidence, with 3, 5 and 10% Sampling Error

Population Size (N)	Sampling Error ±3%	±5%	±10%
100	92	80 (80%)	49
250	203	152 (61%)	70
500	341	217 (43%)	81
750	441	254 (34%)	85
1,000	516	278 (28%)	88
2,500	748	333 (13%)	93
5,000	880	357 (7%)	94
10,000	964	370 (4%)	95
25,000	1,023	378 (2%)	96
50,000	1,045	381 (<1%)	96
100,000	1,056	383 (<1%)	96
1,000,000	1,066	384 (<1%)	96
100,000,000	1,067	384 (<1%)	96

** Adapted from Reisman, 2000, A Field-Guide to Outcomes-Based Program Evaluation*

- As shown above, when a sample is comparatively large, adding cases provides little additional precision.
- As population sizes increase, the total size of the sample becomes proportionately smaller without affecting error.
- When the population size is small, relatively large proportions are required to produce reasonable error rates.
- A standard proportion (e.g., 33%) will not work as a sampling strategy for varying population sizes.
- 3rd RULE OF THUMB you must always draw a larger sample than what is planned for because of refusal. To do this, you need to estimate the refusal rate and then factor that into your calculation. Desired sample size ÷ (1- refusal rate) = TOTAL SAMPLE.

Anita Baker, Evaluation Services www.evaluationservices.co

Suggestions for Effective Interviewing (adapted from Patton 1987)

1. Select the type of interview (or combination of types) that is most appropriate to the purposes of the evaluation. Communicate clearly what information is desired, why that information is important, and let the respondent know how the interview is progressing.

2. Remember to ask single questions and to use clear and appropriate language. Check (or summarize) occasionally to be sure you are hearing and recording the respondent's responses accurately. Avoid leading questions.

3. Listen attentively and respond appropriately to let the person know he or she is being heard.

4. Understand the difference between a depth interview and an interrogation. Qualitative evaluators conduct depth interviews; police investigators and tax auditors conduct interrogations.

5. Recognize when the respondent is not clearly answering the question and press for a full response.

6. Maintain neutrality toward the specific content of response. (You are there to collect information not to make judgments about that person.)

7. Observe while interviewing. Be aware of and sensitive to how the person is affected by and responds to different questions.

8. Maintain control of the interview.

9. Treat the person being interviewed with respect. Keep in mind that it is a privilege and responsibility to peer into another persons' experience.

10. Practice interviewing. Develop your skills.

11. An interview is not a conversation. You should not interrupt the respondent (unless you need to regain control or move the interview along), and you should not share your opinions about the questions or the person's response. You need to cover all the questions on your protocol and you need to deliver them in an order that makes sense.

FOCUS GROUPS

Focus group interviews: these are usually conducted with a semi-structured protocol or list of topics (focuses). The respondents are asked to discuss the topics and the interviewer facilitates, records and later analyzes the discussion. THE FOLLOWING ARE NEEDED.

1. Careful recruitment of participants

   Systematic recruitment procedures
   5 to 10 people per group
   Similar types of people, but not close friends
   3 to 4 groups per topic

2. Proper meeting environment

   Neutral setting
   Circle seating
   Tape recorded

3. Skillful moderator

   Trained, has adequate knowledge of the topic, appears like the participants,
   Has a smooth and snappy introduction that includes a welcome, overview and ground rules
   Uses pre-determined questions, uses pauses and probes (such as "Would you explain further")
   Uses an assistant moderator to handle logistics and take notes
   Establishes a permissive environment, controls verbal and nonverbal reactions to participants
   Uses subtle group control (manages experts, dominant talkers, shy participants, ramblers)
   Uses 3-step conclusion -- summarizes with confirmation, reviews purpose and asks for missing data, thanks participants.

4. Appropriate analysis and reporting

   Systematic analysis
   Verifiable procedures
   Appropriate reporting

5. Helpful References

Morgan, David L., ed. *Successful Focus Groups: Advancing the State of the Art*. Sage Focus Edition.

Stewart, D.W., and Shamdasani, P. N. *Focus Groups: Theory and Practice.* Applied Social Research Methods Series, Volume 20.

Documenting Service Delivery /Assessing Implementation

Defining and Assessing Implementation/Service Delivery

Implementation involves following a design to deliver planned strategies. To assess implementation and/or program delivery, you must be able to accurately describe what a program looks like in operation. You may also want to determine if the description matches the intended program design. For outcomes evaluation, it is important to document program strategies, and/or assess program implementation, so that you can address the relationship between program outcomes and program services.

Collecting Implementation Data/Documenting Program Strategies{tc \l2 "Collecting Implementation Data/Documenting Program Strategies}

Assessment of implementation involves use of all the evaluation data collection strategies described in this guide. Specifically, the following should be undertaken to accurately describe program implementation/service delivery.

- Review documents (program descriptions, proposals -- remember, labels are not good enough)
- Conduct Observations (to determine fidelity and quality)
- Conduct Interviews (ask about the features described below)
- Collect self-reported data (construct surveys or activity reports, collect or inspect logs, participation records)

Focus on the following when collecting implementation data/ documenting program strategies:

✓ **Background and Contextual Information about the Program**

Origin of the program
Nature of the program sites (demographic characteristics, breadth of participation)
How need for the program was determined
Historical background of the program
Background, qualifications and activities of program personnel
Administrative features (including finances where appropriate)

✓ **Critical Features of the Program**

Target group
Activities, schedule, organization
Frequency/duration

Barriers or Problems associated with implementation

DOCUMENTING SERVICE DELIVERY: Planning

Organization Name: _____

Program Name: _____
Notes:

1. What program documents should you review?

2. What program observations should you conduct? (Be sure to specify, how many, how often, and possible dates).

3. Who must you talk to about program implementation? What will you ask? (Fill in the following chart)

WHO MUST YOU TALK TO?	WHAT SHOULD YOU ASK?
	Who are the target participants?
	How are participants recruited for this program?

WHO MUST YOU TALK TO?	WHAT SHOULD YOU ASK?
	Is program retention sufficient? How is retention defined? If there are problems holding participants, what is done?
	How are services delivered?
	How is information tracking handled for this program? What's available?
	OTHER

4. Are there any other data you should review?

Evaluation Bibliography

Cousins, J. B. & Whitmore, E. (1998). *"Conceptualizing participatory evaluation". In E. Whitmore (Ed.). Participatory Evaluation Approaches, New Directions in Evaluation. No. 80. (pp. 5-23). San Francisco : Jossey Bass.*

Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods*, 3rd Edition, (Sage Publications)

Patton, M. Q. (2008) *Utilization-Focused Evaluation: 4th Edition* (Sage Publications)

General Evaluation Books
The following all provide good introductions to, and practical guides on, evaluation (books with an "*" are strongly recommended).

Community Impact Evaluation, by Nathanie; Litchfield (University College, London Press -- UCL Press, 1996)

Empowerment Evaluation: Knowledge and Tools for Self-assessment and Accountability, by D.M. Fetterman, S. Kaftarian, and A. Wandersman, (Sage Publications,1996)

Empowerment Evaluation: Principles in Practice, by D. M. Fetterman and A. Wandersman (Sage Publications, 2004)

Evaluation, 2nd Ed., by Carol H. Weiss (Prentice Hall, 1997)

Evaluation: A Systematic Approach, 7th Ed., by Peter Rossi and Howard Freeman (Sage Publications, 2004) (textbook)

Evaluation for the 21st Century, by Eleanor Chelimsky and William R. Shadish (Sage Publications, 1997) (A reader)

Evaluation Strategies for Communicating and Reporting: Enhancing Learning in Organizations, by R. Torres, H. Preskill and M.Piontek, (Sage Publications, 2004)

Evaluator's Handbook, by Joan L. Herman, L. L. Morris, and Carol Taylor Fitz-Gibbon (Sage Publications, 1987) (Very basic)

**Handbook of Practical Program Evaluation*, by Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer, Editors, 2nd Edition (Jossey-Bass, 2004) (A reader)

Impact Analysis for Program Evaluation, by Lawrence Mohr (Sage Publications, 1995)

**Program Evaluation: Alternative Approaches and Practical Guidelines*, by B.R. Worthen, J.R. Sanders, and J. Fitzpatrick (Addison, Wesley, Longman, 2003) (textbook)

Practical Guide to Program Evaluation Planning: Theory and Case Examples, Holden, D.J., Zimmerman, M. A. (Sage Publications, 2009)

**Program Evaluation: Methods and Case Studies*, 7th Edition, by Emil J. Posavac and Raymond G. Carey (Prentice Hall Humanities/Social Sciences, 2006)

The Logic Model Guidebook: Better Strategies for Great Results, Knowlton, L., and Phillips, C. (Sage Publications, 2009)

What Counts as Credible Evidence in Applied Research and Evaluation Practice? Donaldson, S.I., Christie, C.A., and Melvin, W.M., eds. (Sage Publications, 2009)

Evaluation Guide Books

A Field Guide to Outcome-Based Program Evaluation, by Jane Reisman, 1994. The Evaluation Forum, 1932 First Avenue, Suite 403 Seattle, Washington 98101 206 269 0171

Assess for Success: Needs Assessment and Evaluation Guide, Girls Incorporated. 30 East 33rd Street, New York, NY 10016 (212) 689-3700.

Evaluator's Handbook, by Joan L. Herman, L. L. Morris, and Carol Taylor Fitz-Gibbon (Sage Publications, 1987) (Very basic)

How to Manage and Analyze Data for Outcomes-Based Evaluation, by Marc Bolan, Kimberly Francis, and Jane Reisman, 2000. The Evaluation Forum 1932 First Avenue, Suite 403 Seattle, Washington 98101.

Key Steps in Outcomes Management, 2003. The Urban Institute Press 2100 M. Street, Washington, D.C., 20037.

Measuring Program Outcomes: A Practical Approach, (The United Way)

Outcomes for Success! 2000 Edition. The Evaluation Forum 1932 First Avenue, Suite 403 Seattle, Washington 98101.

Performance Measurment, by Harry Hatry, 1999. The Urban Institute Press 2100 M. Street, Washington, D.C., 20037.

The Managers Guide to Program Evaluation. By Paul Mattessich (Wilder Research Center, 2003)

W.K. Kellogg Foundation: Evaluation Handbook. 1998. One Michigan Avenue East; Battle Creek, MI 49017-4058. 616 968-1611

Focus Groups

Morgan, David L., ed., 1993, *Successful Focus Groups: Advancing the State of the Art*. Sage Focus Edition.

Stewart, D.W., Shamdasani, P. N., Rook, D. 2007. *Focus Groups: Theory and Practice 2nd ed.* Applied Social Research Methods Series, Volume 20.

Quantitative Data Analysis{tc \l2 "**Important References Quantitative Data Analysis}**

Fitz-Gibbon, C.T., and Morris, L. L. 1987. *How to Analyze Data.* Newbury Park, Californa: Sage Publications.

Morris, L.L., Fitz-Gibbon, C.T., and Lindheim, E. 1987. *How to Measure Performance and Use Tests.* Newbury Park, Californa: Sage Publications.

Welch, S., and Comer, J. 2006. *Quantitative Methods for Public Administration: Techniques and Applications*. Chicago: Dorsey Press.

Wright, D.B. 1997. *Understanding Statistics: An Introduction for the Social Sciences.* Newbury Park, California: Sage Publications.

Evaluation Web Sites

The Internet is a great place to get information about evaluation. The following sites on the Internet offer a range of information and resources for evaluation. Many have links to other evaluation-related sites:

General Sites

http://www.eval.org/

The Home Page of the American Evaluation Association, an international professional association of evaluators devoted to the application and exploration of program evaluation, participatory evaluation, personnel evaluation, technology, and many other forms of evaluation. The Participatory evaluation page has links to several "how-to" guides.

http://www.evaluationcanada.ca/site.cgi?s=1

The Home Page of the Canadian Evaluation Association (La Société Canadienne D'évaluation), which is dedicated to the advancement of evaluation for its members and the public (Dévouée à l'avancement de l'évaluation pour le bien de ses membres et du public).

http://www.civicpartnerships.org/

The Center for Civic Partnerships' mission is to provide leadership and management support to build healthier communities and more effective nonprofit organizations. It is a support organization that strengthens individuals, nonprofits, and communities through facilitation and leadership development and by fostering community engagement. The site has an extensive resources section including many tools for identifying indicators and information about logic models and theories of change.

http://www.cyfernet.org
Practical, research-based information from leading Universities. CYFERnet is designed to be used by anyone who needs comprehensive children, youth, or family information including: educators, researchers, parents, youth agency staff, community members, human services and health care providers, students, policy makers, youth, media. CYFERnet's Evaluation section includes practical tools that you can use to evaluate community-based programs; information on how community programs can be sustained; and assessments of organizational support for work in the areas of children, youth, and families.

http://www.ericae.net - A clearinghouse on assessment and evaluation listing many education-related links for assessment and evaluation.

http://www.geofunders.org
Grantmakers for Effective Organizations is a coalition of grantmakers committed to building strong and effective nonprofit organizations.

http://www.hfrp.org/evaluation/the-evaluation-exchange
Harvard Family Research Project, The Evaluation Exchange. *The Evaluation Exchange* is a one-of-a-kind periodical that contains new lessons and emerging strategies for evaluating programs and policies, particularly those focused on children, families, and communities. Since it was launched in 1995, as a key resource on the HFRP Website, *The Evaluation Exchange* has become a nationally known and significant force in helping to shape evaluation knowledge and practice.

http://www.liveunited.org/outcomes/
The United Way's Resource Network on Outcome Measurement: A guide to resources for measuring program outcomes for health, human service, youth- and family-serving agencies. Order their manual, Measuring Program Outcomes: A Practical Approach, here.

http://www.innonet.org/
Innovation Network, Inc., (InnoNet) is an organization dedicated to helping small- to mid-sized nonprofit organizations successfully meet their missions. The purpose of their site is to provide the tools, instruction, framework to create detailed program, evaluation and fund-raising plans.

http://www.theoryofchange.org/
The Theory of Change interactive website was developed and is administered by ActKnowledge. ActKnowledge is an action research organization dedicated to working with community organizations, not-for-profits, foundations and governmental agencies to transform traditional institutions and environments for social change. We work with these organizations as partners in a process aimed at creating transformative knowledge through the interplay of learning and action. Be sure to also check out www.theoryofchange.org/documents/TOCOlayoutandprintingexamples.pdf

http://www.wmich.edu/evalctr/
The Evaluation Center, located at Western Michigan University, is a research and development unit that provides national and international leadership for advancing the theory and practice of evaluation, as applied to education and human services.

Useful Evaluation Terms

Assessment - is a synonym for evaluation, but often used to refer to a technique (e.g., practical assessment) or a mini-study.

Benchmarks - performance data used for comparison purposes. They can be identified from your program's own prior data or relative to performance in the field.

Compliance/Monitoring - type of evaluation where evaluation questions are focused on adherence to pre-specified procedures.

Comparison Groups are non-participants who are identified as a reference for comparison (e.g., individuals at different sites).

Control Groups - are non-participants who are usually identified in the use of an experimental design, ideally on an over-subscribed program (i.e., where there are more participants than slots). The treatment or experimental group actually participates in the program and the control group, although eligible and similar, does not receive or participate in the program. Results of treatment and control group outcomes are then compared to determine program contribution to outcomes.

******** **WARNING -- Comparisons must be conducted very carefully.**

Extrapolation - modest speculations on the likely applicability of findings to other situations under similar, but not identical conditions. Extrapolations are logical, thoughtful, and problem-oriented rather than purely empirical, statistical and probabilistic.

Formative Evaluations - focus on ways of improving and enhancing programs, and are conducted in the early or ongoing stages of a program.

Generalize - to assign qualities based upon group membership, or to make inferences about groups or programs based on the outcomes of a sample or subset of members.

Goals - are conditions (usually broad) that programs are working toward (e.g., to promote well-being).

Indicators - observable, measurable characteristics of changes that represent elements of an outcome (e.g., normal birth weight is an indicator of a healthy baby outcome).

Needs Assessments - determine whether existing services are meeting needs, where there are gaps in services and where there are available resources. These are often conducted prior to initiation of an evaluation or in response to evaluation findings.

Objectives - something that is worked for or strived for, which can be observed or measured.

Outcomes - results for participants, during and/or after participation in a program

Outputs - products of a program's activity (e.g., # of sessions held, # of participants served).

Qualitative Data - consist of detailed description of situations, events, people, interactions, and observed behaviors; direct quotations from people about their experiences, attitudes, beliefs and thoughts; and excerpts or entire passages from documents, correspondence, records and case histories. Qualitative data collection methods permit the evaluator to study selected issues in depth and detail and typically produce a wealth of detailed data about a much smaller number of people and cases.

Quantitative Data - come from questionnaires, tests, standardized observation instruments, and program records. Quantitative data collection methods permit the complexities of the world to be broken into parts and assigned numerical values. To obtain quantitative data it is necessary to be able to categorize the object of interest in ways that permit counting.

Random Assignment - a technique which allows program providers to randomly divide participants into treatment (those who get services) and control groups (those who don't).

Reliable Measures - those which can be repeated under similar conditions.

Research - in social science is also a systematic collection of information, but it is undertaken to discover new knowledge, test theories, establish universal truths and generalize across time and space.

Summative Evaluations - are aimed at determining the essential effectiveness of a program. They are especially important in making decisions about terminating, maintaining or extending a program.

Triangulation - multiple streams of information obtained by either collecting different kinds of data about the same subject, using different workers to complete the same tasks; using multiple methods to obtain data; using multiple perspectives to analyze information.

Valid Measures - those which accurately measure what they are intended to measure. (Warning, this is difficult to test. For most social and behavioral variables, no agreed upon testing standards exist).