Evaluation Use in Philanthropy

Evaluation Data Collection

Anita Baker, Ed.D.
Beth Bruner, Bruner Foundation

Bruner Foundation
Rochester, New York

Remember, Evaluation Questions . . .

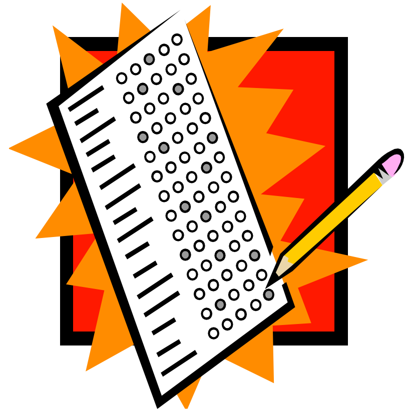
- ✓ Focus and drive the evaluation
- ✓ Should be carefully specified and agreed upon in advance of other evaluation work
- ✓ Generally represent a critical subset of information that is desired (i.e., you can't ask about everything, so you must choose what is most needed)

So you have your evaluation questions.... What's next?

- Evaluation questions = ***what*** the evaluation is trying to answer
- Data collection methods = ***how*** you're going to get the information you need to answer
- Data analysis planning = ***how*** you are going to summarize and use the information, what procedures will data be subjected to

Surveys:

- Have a series of questions (items) with pre-determined response choices.
- Can include all independent items or groups of items (scales) that can be summarized.



- Can also include open-ended items for write-in or clarification;
- Can be completed by respondents or survey administrators;
- Can be conducted via mail, with a captive audience, on the phone or using the internet, or through a variety of alternative strategies.

Instruments are called surveys, questionnaires, assessment forms

Use Surveys:

- To study attitudes and perceptions
- To collect self-reported assessment of changes in response to program
- To collect assessments or ratings of programs
- To collect some behavioral reports
- To test knowledge
- To determine changes over time

Surveys are best used with big or distant groups, for sensitive information

Interviews

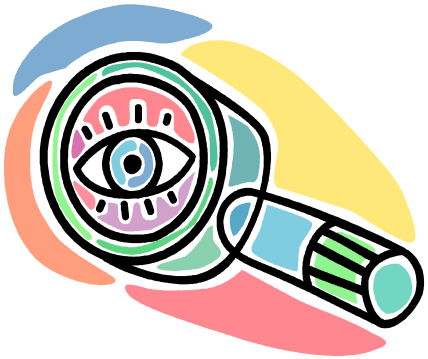
- An interview is a one-sided conversation between an interviewer and a respondent.
- Questions are pre-determined, but open-ended, and can be structured or semi-structured.
- Respondents are expected to answer using their own terms.
- Interviews can be conducted in person, via phone, one-on-one or in groups. Focus groups are specialized group interviews.

Instruments are called protocols, interview schedules or guides

Use Interviews:

- To study attitudes and perceptions using respondent's own language
- To collect self-reported assessment of changes in response to program
- To collect assessments or ratings of programs
- To document program implementation
- To determine changes over time

Interviews are best used with smaller groups with whom you have more/regular contact, when depth and detail are needed.



Observations

Observations are conducted to view and hear actual program activities so that they can be described thoroughly and carefully.

- Observations can be focused on programs overall or participants in programs.
- Users of observation reports will know what has occurred and how it has occurred.
- Observation data are collected in the field, where the action is, as it happens.

Instruments are called protocols, guides, sometimes checklists

Use Observations:

- To document program implementation
- To witness levels of skill/ability, program practices, behaviors
- To determine changes over time

Observations are best used in conjunction with other strategies and must be conducted more than once for any purpose other than general description.

Record Review

Review of program records involves accessing existing internal information or information that was collected for other purposes. Data are obtained from:

- a program's own records (e.g., intake forms, program attendance)
- records used by other agencies (e.g., report cards; drug screening results; hospital birth data)
- adding questions to standard record-keeping strategies (e.g., a question for parents about program value can be added to an enrollment form)

Instruments are called protocols.

Use requires identification of and access to available information.

Use Record Reviews:

To collect some behavioral reports

To test knowledge

To verify self-reported data

To determine changes over time

Key Tasks to Complete Before Collecting Data

1. Identify data sources and data collectors
2. Select methods
3. Develop and/or test instruments and procedures as appropriate (paying careful attention to validity – potential to accurately collect information; and reliability – potential to be consistently used under similar conditions; as appropriate)
4. Plan for analysis

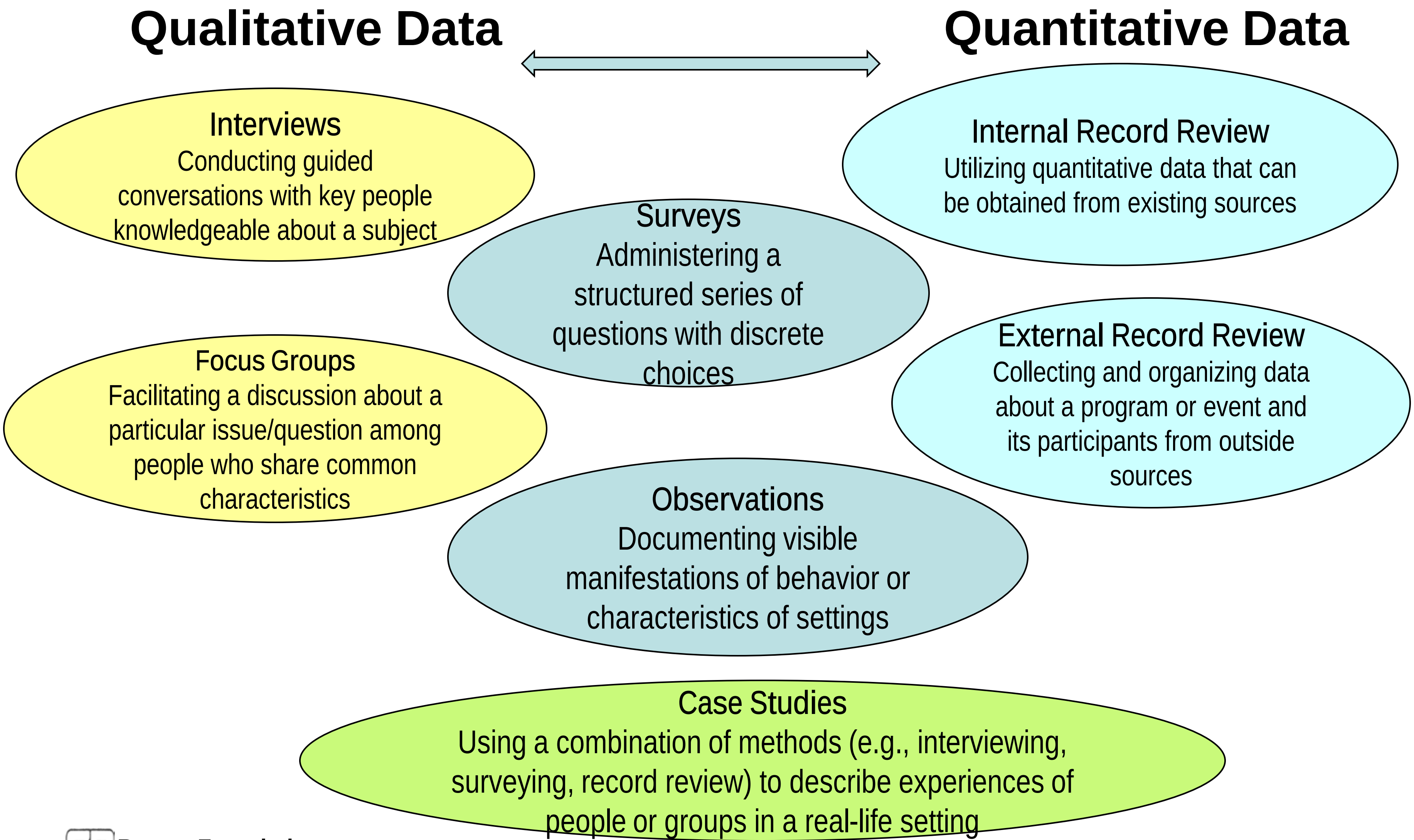
Practical Considerations for Data Collection

- Timeliness….Collect data that is timely enough to be useful

- Feasibility….Keep in mind resource constraints, ethical issues

- Low burden…. Consider what are you already collecting; determine if you can gather data from automated systems rather than direct contact with respondents.

Evaluation Data Collection Options



Evaluation Data Collection: Common Mistakes

- Collecting but not analyzing data
- Planning for analysis AFTER data are collected
- Limiting data collection from clients/participants to satisfaction only
- Basing data collection decisions on what stakeholders want rather than what will help inform decision-making

What Happens After Data are Collected?

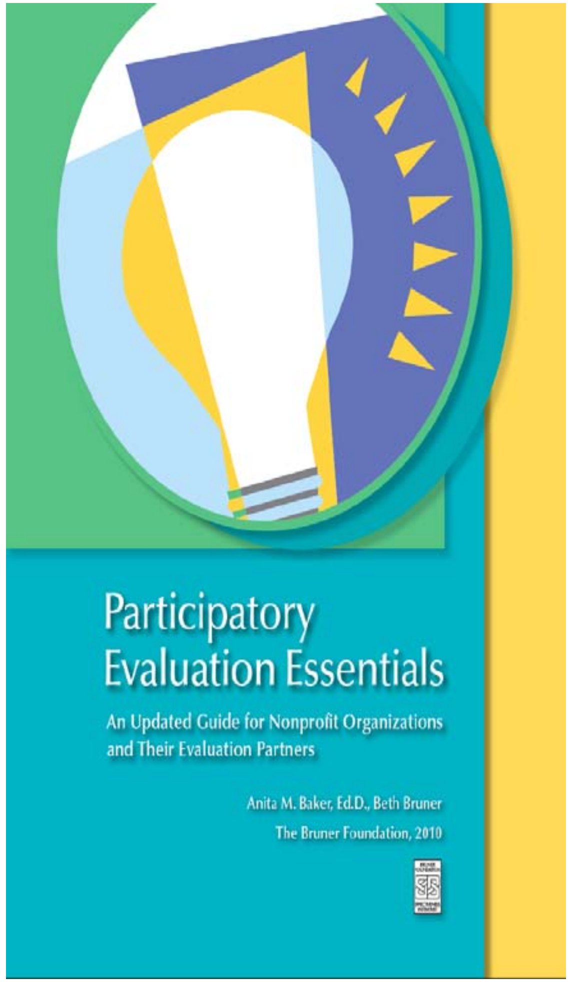
1. Data are analyzed, results are summarized.
2. Findings must be converted into a format that can be shared with others.
3. Action steps are developed from findings

“*Now that we know _____ we will do _____.*”

Step 3 moves evaluation from perfunctory compliance into the realm of usefulness.

List of Attachments

- Making Data Collection Decisions
- Observation Protocol Example – Checklist
- Observation Protocol Example – Comprehensive
- Record Review Protocol



Be sure to see *Participatory Evaluation Essentials: An Updated Guide for Nonprofit Organizations and Their Evaluation Partners, 2010, on the Bruner Foundation Website for additional details about collecting and analyzing survey, interview, observation and record review data.*
www.brunerfoundation.org

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, overinterpretation 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 interobserver 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?

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.

AFTERSCHOOL LITERACY DEVELOPMENT PROJECT: OBSERVATION SUMMARY (CHECKLIST PROTOCOL EXAMPLE)

Site ______________________________ Location _________________________ Date: __________
Observer _________________________ Purpose for Visit:
Activity Description:

PROGRAM FACILITIES/EQUIPMENT

In the space below, provide a brief description of the site including size of the space (sq. footage, # of rooms), arrangement of furniture, distinguishing features.

Is the program site accessible to all potential participants?...................................... ☐ No ☐ Yes
Does the site have an area where participants can casually interact (hang out)?........... ☐ No ☐ Yes
Is participant work displayed at the site? .... ☐ No ☐ Yes
Are there any unmet maintenance needs? .... ☐ No ☐ Yes

Please rate the following features of the physical environment at the site.

1 = Poor 2 = Fair 3 = Good 4 = Excellent	1	2	3	4	NA
Attractiveness of physical facility (freshly painted, good lighting etc.)	☐	☐	☐	☐	☐
Condition, appropriateness, and quantity of furniture	☐	☐	☐	☐	☐
Ability to accommodate large and small group activities at same time	☐	☐	☐	☐	☐
Availability of independent study areas	☐	☐	☐	☐	☐
Availability of reading areas with comfortable seating	☐	☐	☐	☐	☐
Availability of computers for participant use	☐	☐	☐	☐	☐
Participant access to the internet	☐	☐	☐	☐	☐
Availability of books or other reading materials	☐	☐	☐	☐	☐
Availability of art supplies for participant projects	☐	☐	☐	☐	☐
Attractiveness to children	☐	☐	☐	☐	☐
Overall facility rating	☐	☐	☐	☐	☐

Engagement in Effective Practice

Practice	# of Participants Engaged (circle one)	Description/Examples
Independent Reading	None Some Most NA	
Paired Reading/Discussion Groups	None Some Most NA	
Dramatic Interpretation	None Some Most NA	
Independent Writing	None Some Most NA	
Reading Instruction	None Some Most NA	
Theme-Based Project	None Some Most NA	
Other:	None Some Most NA	

Overall session rating (circle one): Poor Fair Good Excellent

Comprehensive Program/Session Observation Protocol (EXAMPLE)

Program Name:____Observer's Name: _____________

Date: ___________ Time Observation Began: __________ Time Ended: __________________

Before the observation begins, briefly describe in #1 below, what you expect to be observing and why you have selected it.

1. Subject of the Observation. ____

At the very beginning of the observation, describe the setting. Be sure to note any changes in setting as the observation proceeds. Also note how the session begins.

2. Describe the **program setting** (color, size, shape, number of desks/tables, number of windows, furniture or equipment in the space room, temperature, noise level)

3. Describe **how the session begins.** (who is present, what exactly was said at the beginning)

4. Describe the **chronology of events i**n 15 minute intervals.

15 Min.	
30 Min.	
45 Min.	
60+	

5. By answering the following questions, describe the **interactions that take place during the observation.**

5A. Who is interacting?

Youth with Adults Youth with Youth Adults with Adults
Girls with Girls Boys with Girls Boys with Boys
Youth of different racial/ethnic backgrounds

5B. How do they interact? Describe 1 or 2 examples.

5C. Are there any changes in interaction during the observations.

6. Describe how decisions are made during the observation period (by answering the following questions)

6A. Who makes decisions?

Only Adults	Mostly Adults	Only Youth
Youth and Adults	Mostly Youth	

6B. How are decisions communicated? (e.g., written, verbal,).

6C. Document examples of decisions that are made during the observation. *(Be sure to record who is making the decision.)*

7. Describe **Nonverbal communication** (How do participants get attention? How much do they fidget, move around? How do participants: dress, express affection, physically place themselves in the setting?)

8. Describe program activities and participant behaviors (i.e., what's happening during the session and how participants respond).

9. How did participants respond or react to what was happening with the program during the observation? **Roughly** what proportion (some, most, all) are actively engaged?

10**. How does the program end?** (What are the signals that the activity is ending? Who is present, what is said, how do participants react, how is the completion of this activity related to other activities?)

Example of a Record Review "Protocol" NCSC Data Collection Summary

Data Category/Form	Data Elements	Collected By	When	Uses
Registration Form	Contact Information, Demographics (age, gender, race/ethnicity) **Includes signature of parent/guardian	Site	At enrollment	FOR SITE USE ONLY
Family Fact Sheet	Language, race/ethnicity, school, household composition, lunch eligibility, sibling participation, special needs, insurance, interests	Site (copies submitted to evaluator)	At enrollment	Profile of participating youth Target population assessment
Participant List	Identification numbers	Site – submitted to evaluator	Quarterly	Update files
Attendance/Activity Summary	Daily attendance by staff member and activity summary by category	Staff members at site – copies submitted to evaluator	Daily by staff, quarterly submission to evaluator	Determine mean attendance, by grade and gender. Calculate program intensity.
Outcome Report - Participants (random sample)	Staff report of activity, ratings of interest, behavior, interaction with youth, interaction with adults, literacy skills, study habits.	Staff	Quarterly	Baseline profile of participants on outcome measures; assessment of change on outcome measures.
Quarterly Report - Site	Service delivery – events, workshops, counseling, school relationship ratings	Staff	Quarterly	Service delivery and contextual data