Appendix: Issues in Developing Data Collection Instruments and Procedures

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

Record Extraction Forms

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

Instrument Development

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

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

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

Exhibit 1

Example of a Record Extraction Form

Identification number:	___ ___ ___ ___ ___ ___ ___	
Date of birth:	___ ___ / ___ ___ / ___ ___ month day year	
Sex (circle one):	1 Male 2 Female	
Race or ethnicity (circle one):	1 African American	5 Asian, Pacific-Islander
	2 White, non-Hispanic	6 American Indian, Alaska Native
	3 Black, non-Hispanic	7 Other ____________________
	4 Hispanic	
Intake date:	___ ___ / ___ ___ / ___ ___ month day year	
GED pre-test score:	___ ___ ___	
Date of first GED exam:	___ ___ / ___ ___ / ___ ___ month day year	
Result of first exam (enter score OR circle "9"):	___ ___ ___	
	9 Not applicable. Youth dropped out before first exam.	
Date of second GED exam:	___ ___ / ___ ___ / ___ ___ month day year	
Result of second exam (enter score OR circle "9"):	___ ___ ___	
	9 Not applicable. Youth completed GED with first exam or dropped out before second exam.	

With data from the form, the program can measure:

- The number of youth who completed the program (all those who passed the GED exam the first or second time).

- The percent of youth who completed the program (the number of youths who obtained a GED divided by the number who entered the program).

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

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

Procedural Issues Specific to Record Extraction

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

- Are there any restrictions on the information they will release? If you were counting on obtaining data they cannot release, you will need to go back to your outcome measurement framework to identify an alternative source, or perhaps to modify the indicator or even the outcome you will measure.

- Do your definitions of key terms coincide with theirs? If not, be sure you know their terms or data categories for the information you want.

- Are their records computerized? If so, what are the procedures for obtaining an electronic file with the needed information? You may want to develop an inter-agency agreement with another program regarding sharing of confidential data.

- If the extraction must be done manually, who will be allowed to review the records and copy the information? It may be that a member of their staff will have to do this. Or perhaps a member of your staff can have access, but not a volunteer or student intern.

- What information is available to confirm that someone in their records with the same name as your participant is in fact the same person?

Questionnaires

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

Instrument Development

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

Content

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

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

- The amount and type of service received. Although this information is probably available from program case records, the program may want to obtain such information from the participant. The participant's perception of the services received may provide useful information to the program and will avoid requiring the program to link two different sources of information.

- Why respondents gave particular answers or ratings. Ask especially for explanations about poor ratings. Keep in mind, however, that questions requiring more than a check mark or a few words to answer take more time to collect and more time to analyze.

- Demographic and other information, such as age, sex, income, race/ethnicity, and location, needed to break out the outcome findings. Some respondents are uncomfortable answering some demographic questions. To increase the likelihood that they will answer the other questions, put demographic questions at the end of the form and explain briefly why the data are needed.

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

- Awareness of program services
- Timeliness of service
- Convenience/accessibility of hours of operation
- Accuracy of help
- Pleasantness/friendliness of staff
- Condition of facilities participants see
- Convenience/accessibility of location
- Ease of reaching someone to talk to
- Ratings of specific service characteristics
- Overall satisfaction with program
- Suggestions for improving service

Participants also can be asked about their needs, to help with needs assessment, and about what additional help they would like. Such information may not be directly related to outcome measurement, but it is a side benefit of participant surveys. Be careful, however, that you do not lengthen the questionnaire so much that you discourage participants from responding.

Exhibit 2 illustrates questions that can be included in participant questionnaires. These questions are from questionnaires that the Family Service Association of America used to assess the outcomes of family counseling services.

The first question will provide information on an *outcome indicator*: number and percent of participants who report improved outcomes after receiving services. The second question illustrates that participant surveys also can be used to obtain information on the participant's perception of the extent to which the services *affected the outcomes* that the participant reports. A program can combine the data from responses to the first and second questions to calculate an additional outcome indicator: number and percent of participants that indicate improved outcomes *and* report that the services they received contributed to that improvement. As noted in the Introduction and Step 1 of this manual, outcome information generally does not tell *why* the outcomes occurred. If such questions as this one are included in participant surveys, the program can obtain some information on the "whys."

Exhibit 2

Survey Questions on Family Counseling

1. Since you started at the agency, has there been any change for better or worse in the way the members of your family get along with each other? Would you say you now get along:

   __ Much better
   __ Somewhat better
   __ The same
   __ Somewhat worse*
   __ Much worse*

* Please explain: ______________________________

2. How do you feel the service provided by the agency influenced the changes you have reported?

   __ Helped a great deal
   __ Helped somewhat
   __ Made no difference
   __ Made things somewhat worse*
   __ Made things much worse*

* Please explain: ______________________________

3. How satisfied were you with the way you and your counselor got along with each other?

   __ Very satisfied
   __ Somewhat satisfied
   __ No particular feelings one way or the other
   __ Somewhat dissatisfied*
   __ Very dissatisfied*

* Please tell us why you felt this way: ______________________________

The third question will provide information on one aspect of service that may affect whether the outcome occurred: satisfaction with the participant's counselor.

Note that if participants give unfavorable ratings in response to a question, they are asked to explain "why." Summaries of participants' responses should be made available to program personnel, keeping individual responses anonymous. Such information can provide guidance on how the program can be improved, especially to the extent that particular problems are reported by many participants.

You also might ask "why" of respondents who give favorable ratings to questions. This would be particularly useful for programs where the specific service activity varies among participants, such as camping, mentoring, or scouting.

Wording

Questions must be clear, unambiguous, and neutral in their presentation. It is surprisingly easy to use biased or unclear wording. For instance, questions that provide four response categories, of which three are favorable and one is unfavorable, bias the response in a favorable direction. The question, "*When* do you experience this event?" may mean "at what time" or "under what circumstances." Be specific.

Most questions should ask respondents to check or circle the appropriate response from a fixed set of answers (i.e., *fixed-choice* questions), rather than asking them to compose an answer (i.e., *open-ended* questions). A self-administered questionnaire might include two or three open-ended questions. Interviews can include a few more. In either case, open-ended questions are harder to analyze. They also take longer to complete. This may discourage respondents from finishing the survey, although it also may encourage responses from people who have strong reactions about the program. Exhibit 3 gives examples of fixed-choice response sets.

To write effective questions:

- Keep each question short. Use simple sentence structures.
- Use basic vocabulary.
- Ask only one question in each question. For example, "In disciplining your child, do you state the rule clearly and explain the consequences of breaking the rule?" is two questions.

Exhibit 3

Examples of Fixed-choice Response Sets

Question Type	Examples
Yes or No	Have you used this new skill in the past week? ❑ Yes ❑ No
Agree/disagree scale	How much do you agree or disagree with the following statement: "The public service announcements about the effect of second-hand smoke gave a very powerful argument against smoking around small children." ❑ Strongly agree ❑ Disagree ❑ Agree ❑ Strongly disagree
Frequency scale	How often do you read to your child? ❑ At least once a week ❑ Once a month or less ❑ At least twice a month ❑ Never
	Last week, how many evenings did you read a bedtime story to your child(ren)? ❑ 1 ❑ 2 ❑ 3 ❑ 4 ❑ 5 ❑ 6 ❑ 7 ❑ None
	What percentage of your income do you put into a savings account? ❑ None ❑ 6 to 10 percent ❑ 1 to 5 percent ❑ More than 10 percent
Comparative response scale	Compared to other programs that your child(ren) take part in, how important is the after-school tutoring program? ❑ Very important ❑ Somewhat important ❑ Unimportant
Identification response	Below is a list of services provided by our program. Please check the ones your child participated in since the beginning of the year: ❑ After-school program ❑ Sports program ❑ Tutoring program ❑ Summer camp

Adapted from: Irwin Epstein and Tony Tripodi. *Research Techniques for Program Planning.* © 1977 by Columbia University Press. Used with permission of the publisher.

- Think carefully about words and phrases that may have double meanings. This is especially true if the question writer is of a different age, ethnic or cultural background, or educational level, or is from a different part of the country than the respondents. (Exhibit 4-D in Step 4 describes some cultural considerations in questionnaire development.)

- Make sure nothing in the question, including the response choices, suggests a "right" answer.

- Be specific about time frames. For instance, say "within the past three months" or "since March 1" instead of "recently."

- Make response categories as specific as possible. Choices like "excellent; good; fair; poor" and "frequently; infrequently; rarely" are subject to wide interpretation. An alternative frequency scale might be "every day, once a month, fewer than 12 times a year." An alternative qualitative scale might be "have used the skill already; haven't used it but am confident I can; not sure I understand how to do this; don't understand how to do this."

- Read questions aloud. This practice will pinpoint some awkward phrases and unclear meanings.

- Have someone experienced in questionnaire development review the questionnaire before it is finalized.

Format and Style

Once the questions are written, you are ready to arrange and format your questionnaire. Make it clear and easy to answer. Here are some suggestions on format and style, especially for self-administered questionnaires:

- Group related questions together. Start with the least-personal and most obviously relevant set of questions.

- Be sure instructions are short and explicit.

- If you must use *skip patterns* (e.g., if respondents answer "no" to a question about whether they have children and you want them to skip the next three questions because they concern children), be sure the instructions are easy to follow. Consider using arrows to lead visually to the appropriate next question. As an alternative to skipping questions, consider adding a "not applicable" response. For example, in the three questions about children, add a response category that says "not applicable; no children."

- Make the questionnaire easy to read. Use a clear typeface. Leave plenty of "white space." Use left justification only. If interviewers or respondents are older persons, use a larger typeface.

- If you plan to use colored paper, pick a color that photocopies well and is easy on the eyes.

If substantial numbers of persons with limited English proficiency will be respondents, translate the questionnaire into their native language. Then, test the accuracy of the translation by having a second person translate the foreign-language version back into English to see if the questions still say the same thing.

Procedural Issues Specific to Questionnaires

Step 4 of this manual summarized several general issues related to data collection procedures. Collection of data via questionnaires involves a number of additional issues, discussed below.

Administration

The primary methods of administering questionnaires are interviews and self-administration. Generally, self-administered questionnaires are either given to respondents at the program site or mailed to their homes. Interviews generally are administered either by telephone or in person. In-person interviews can be conducted at the respondent's home, at the program site, or at another accessible facility. Considerations related to these options are as follows.

- **Mail.** Mail surveys usually are inexpensive; however, they require special procedures to obtain acceptable response rates. Mailed questionnaires need to be short and simple. They are not useful for respondents with very low levels of literacy. For those with difficulty understanding English, the questionnaire can be translated into other languages. Note that many communities with non-English speaking residents have a diverse ethnic make-up and translation into only one additional language may not suffice. Getting adequate response rates can be a major problem with mailed questionnaires. You will need multiple mailings, and possibly telephone follow-ups, to encourage a response or to conduct a phone interview. Additional suggestions are given later in this appendix.

- **Telephone.** Telephone surveys are a less expensive alternative than in-person interviews and can get good response rates. However, they require considerable interviewer time, and interviewers need training to conduct telephone interviews. If the agency has staff or volunteers that have time to do telephone interviews, this option may be a very good one. Note, however, that to avoid biasing the answers and to ensure treating participants ethically, the interviewers should not be the same persons that provided services to the participants. If a substantial percentage of program participants do not have a telephone, then clearly this is not a good data collection mode for the program.

  Surveys focused on low-income populations do need to seek responses from adequate numbers of those populations. One possibility is to conduct a telephone survey and then conduct in-person interviews in those neighborhoods known to have few telephones. If substantial numbers of persons speaking limited English are expected, bilingual interviewers will be needed.

- **In person at the person's home.** Generally, this is quite expensive and not likely to be feasible for regular data collection. However, it is sometimes the best option for elderly homebound participants and other special populations. For instance, if you are interested in obtaining information about experiences with attendant care provided to elderly persons and persons with disabilities, you might consider in-home interviews with a sample of attendant-care recipients.

- **In person at the agency or other readily accessible facility.** Usually, this is quite inexpensive but not useful for obtaining post-service information unless the participant is coming for other services (for instance, parents who have completed a parenting class but still need to see their case worker on a regular basis). The survey can be administered through an interview or by asking the respondent to check off the answers to a written questionnaire.

- **Combinations of the above.** For example, a meals-on-wheels program could have drivers distribute post card questionnaires to new recipients, then use interns or volunteers to tabulate responses and do telephone follow-up with recipients who do not respond.

Key concerns in selecting a method of administration include response rates, accuracy of responses, and cost. Choosing a method involves trade-offs among these concerns. Exhibit 4 provides a summary comparison of the various methods of questionnaire administration.

Most programs will find that mailing questionnaires will be the most practical approach because of its low cost. Programs that undertake second mailings and possibly telephone reminders to achieve return rates of 50 percent or more should achieve sufficiently accurate data for outcome measurement. However, if the program has a cadre of staff, volunteers, or students it can depend on to undertake regular telephone interviews at a low cost to the program, telephone interviewing would be the preferred method.

For programs that are seeking data on outcomes which occur early in the service delivery process or immediately thereafter, and/or where the program does not obtain the names or addresses of participants, then in-person administration at the location where the service is delivered probably would be the appropriate method of questionnaire administration. For example, this method might be appropriate to obtain data on the knowledge gained in a parenting education workshop or skills obtained in a CPR course.

In circumstances where a program can reach participants only in their own environment, in-person interviewing there may be necessary. For example, to reach homeless individuals who do not come into shelters, a program providing homeless services might need to attempt interviews on the street. Another example is that some participants may be home-bound, without a telephone, and unlikely to be able to respond to a mailed questionnaire. At-home interviews might then be the only option for interviewing those participants.

Exhibit 4

Summary Comparison of Questionnaire Administration Options

	Method of Questionnaire Administration			
Characteristic	**Mail**	**Telephone**	**In Person at Home**	**In Person at Facility**
Response rate	Low/Medium*	High	High	High
Amount of information obtainable	Low	Medium	High	Medium
Ability to obtain post-service information	High	High	High	Low
Speed in getting responses	Slow	Fast	Medium	Fast
Cost	Low	Medium	High	Low/Medium

*Low, if only one mailing. Medium, if multiple mailings and possibly telephone follow-ups are done.

Voluntary Completion

Answering questionnaires, regardless of how they are administered, must be voluntary. Persons asked to respond have the right to refuse and must not be pressured to respond. Programs need to be certain that declining to answer in no way penalizes participants in receiving future services. Further, participants should be told explicitly that choosing not to respond will not affect their eligibility for services.

Participant Confidentiality

Depending on the type of program you operate, the fact that the participant obtained services from your program may be confidential. One cannot, for example, call someone's home and leave a message asking the person to call her or his former substance abuse or mental health counselor. For programs in which this is an issue, telephone interviewers must be sure to reach the participant directly; mail questionnaires probably are not appropriate. For other programs, such as child care or youth development, confidentiality may be less of an issue.

Anonymity

If participants are interviewed in person, their responses can be kept confidential, but the participants are not anonymous because the interviewers know who they are. This generally is true of telephone interviews as well. Mail questionnaires, however, can be returned without an identifying name or code, and thus be anonymous. The key advantage to anonymity is that respondents may be more honest with their responses.

There are several disadvantages. One is that the responses cannot be linked with data from program records. Therefore, all information that is needed about date of entry into the program, type and amount of services received, and other outcomes and influencing factors needs to be asked on the questionnaire. Another disadvantage is that the program cannot follow up with non-respondents because it has no way of knowing who has responded and who hasn't. This means either a lower response rate or increased costs because a second mailing needs to go to all respondents.

A third disadvantage of anonymous questionnaires is that there is no way of linking data that are collected from participants at multiple points to track changes over time. Programs sometimes try to overcome this problem by having participants put a personal code, such as their mother's maiden name, on each questionnaire they complete. The effectiveness of such strategies depends on each participant's code being unique and on participants' memory of the code they used on previous questionnaires.

Times for Contacting Potential Interview Respondents

Decisions about days and times that potential respondents are contacted for interviews depend to a great extent on the characteristics of the respondents. Elderly and retired individuals may be called during daytime hours. Employed parents who work day shifts are best contacted in the evening. Avoid religious holidays and those days respondents celebrate as the Sabbath. Set time limits on when contacts may be made on other days (e.g., not before 10:00 a.m., during the dinner hour, or after 9:00 p.m.).

Response Rate

The smaller the response rate, the greater the possibility that the sample is not representative of the population from whom the program is seeking responses. Whatever administration method you use, as long as the respondent selection method provides a representative sample, aim to get completed questionnaires from at least 50 percent of those from whom you are seeking responses. (The federal government seeks a 75 percent rate for federal surveys, but such a target is likely to be too costly for regular surveys by most private, non-profit agencies.) With mail surveys, second and third mailings or telephone reminders to non-respondents usually will be needed to obtain response rates high enough to provide reliable information. For telephone surveys, multiple calls at various times of the week are likely to be needed to reach respondents. Exhibit 5 gives tips for increasing response rates on mail surveys.

Costs

Here are some ideas for saving costs on questionnaire administration.

- Use already-available questionnaires.
- Use volunteer technical consultants to help design the questionnaire and procedures. Some otherwise well-paid consultants will accept a lower rate (or provide free time) when working for a non-profit organization. Although technical help *may* cost more initially, it is likely to save you the cost of collecting unreliable, unusable, or insufficient data.
- Use mail surveys (but plan on doing second and even third mailings to obtain an adequate response rate).
- If your program has several hundred participants, survey a sample rather than everyone.
- Use volunteers to conduct interviews. Ask for help from a local university or college.

Exhibit 5

Tips for Improving Mail Survey Response Rates

1. Establish a procedure to inform new participants about the survey process and encourage them to respond if and when they are surveyed later.
2. Make questionnaire as short and easy to use as possible.
3. Respondents are more likely to complete and return questionnaires that look attractive and simple. Go for a professional, polished look—even if you produce the questionnaire in-house. Use a printer or high-quality copier and good-quality paper. Check for spots, blurs, and crooked copies. If using colored paper, pick a color that is easy on the eyes. Avoid expensive-looking paper, two-color printing, or other features that may give the impressions you "wasted" money on production. If the questionnaire is to be completed by children, make it as much fun as possible to do.
4. To the extent possible, avoid surveying during vacation and holiday seasons.
5. Use an advance post card to notify participants that they will be receiving a mailed questionnaire shortly.
6. Include a carefully worded transmittal letter signed by an agency official, such as the executive director or board chair, or a notable community figure encouraging response. Emphasize the agency's need for the information to improve future services. Explain that each question is important. Guarantee that responses are confidential and will not be attributed to individuals in any way.
7. If possible, personalize the letter and address it to a specific individual.
8. Consider offering inexpensive incentives for responding, such as certificates for a fast-food store drink.
9. Enclose a postage paid, pre-addressed return envelope with the questionnaire at each mailing.
10. Send post card reminders that thank those who already have responded and encourage others to complete and return the questionnaire.
11. Re-mail the questionnaire if the response rate is too low.

Trained Observer Ratings

One of the values of trained observer ratings is that they are more objective than self-reports and can be more consistent than casual observations. Achieving a high degree of accuracy and reliability in observation ratings depends on having clearly defined rating categories used by carefully trained observers. The more observers that are involved, the more critical the rating scales are.

Instrument Development

Rating systems used by trained observers include:

- **Written descriptions.** This is the most basic type of rating system. It relies solely on written descriptions of each grade in the rating scale. These descriptions must be quite specific about what behaviors or characteristics merit each rating. This is necessary so that different raters give approximately the same rating for a particular condition. Exhibit 6 provides an abbreviated example of written guidelines for the condition of a playground connected to a child care program in an inner-city neighborhood.

- **Photographic rating systems.** These systems use pre-selected photos to represent the various grades on the rating scale. Photographic scales are one of the most useful forms of rating systems. They have been used to assess the cleanliness of streets, the condition of parks and playgrounds, and external housing conditions.

- **Other visual scales.** Drawings, sketches, video clips, or other visual alternatives to photographs can be used to depict each grade on a rating scale.

Procedural Issues Specific to Trained Observer Ratings

Achieving a high degree of accuracy and reliability in observation ratings depends on having:

- Clearly defined rating categories.
- Adequate training and supervision of the observers and the process.
- A procedure for checking the quality of the ratings periodically. The goal is for different observers, perhaps at various times, to give the same ratings to comparable conditions.

Observers can be program agency personnel, college or graduate school students, volunteers, or community members. More technical ratings will require persons with more professional training. Avoid having staff or volunteers observe persons or places with whom they work.

Exhibit 6

Sample Rating Scale for Trained Observer Rating of a Playground for Preschool Children

Rating 1: **Clean.** The area is completely, or almost completely, clean; up to two pieces of paper litter are permitted. None of the litter is broken glass, drug paraphernalia, or other dangerous materials.

Rating 2: **Moderately clean.** The area is largely clean; a few pieces of isolated litter are observable. None of the litter is broken glass, drug paraphernalia, or other dangerous materials.

Rating 3: **Heavily littered.** Scattered litter in several areas and/or blown all along the fence. However, none of the litter is broken glass, drug paraphernalia, or other dangerous materials.

Rating 4: **Dangerously littered.** The litter includes broken glass, drug paraphernalia, or other dangerous materials. The area therefore is not safe for children to use.

Training of observers is important. For behavioral ratings, videotaped examples of various levels of the behavior often are effective. Training should include opportunities to practice applying the rating scales and then compare and discuss the ratings with other observers.

Establish procedures for supervising observers, recording the data they collect, and transcribing and processing the data. Also establish procedures for systematically checking the ratings to assure quality control and reliability of the process. For example, the supervisor might check a small sample of the ratings done by each observer.