

While the previous chapters deal with conceptual aspects of an evaluation, this section discusses the more technical issues related to implementing an evaluation. Instruments are the “hammers and saws” used to implement your evaluation. They measure the services of your program and collect information needed to determine your outcomes. See “Instrument Development” in the reference section for more information.

Chapter 3 Instrument Development- The Hammers and Saws

Background Concept

Existing Information

Existing information is information collected by a school or a service agency staff as a part of their operation. (Examples include grades, attendance, test scores, homework completion records, crime statistics, and shelter records.)

Program-Generated Data

Program-generated data are collected using instruments created and administered by your program. It takes more time in the development stage than using existing data. However, program-generated instruments provide a greater degree of control over the type of data collected.

Sampling Issues

A sample is when you deliberately administer an instrument to fewer people than you serve. The following are some simple rules for samples:

- If a group is small enough, use all group members in the sample.
- If the group is large, draw a sample by making a list of the individuals and choosing a respondent at evenly spaced intervals. For example, with new clients, you could choose every fourth person who walks in the door. For existing clients, you could create a list and pick every fifth person, using the entire list. This technique is called random sample selection. Following these rules should prevent a “biased” sample, which means choosing too many respondents with certain characteristics (for example, too many people who did well on a test or too many people who failed).



Tip: Keep it simple.
Make sure your instrument is easy to administer and respond to, with clear directions, specific questions, and (for written instruments) sufficient spacing to allow responses. Keep your formatting and wording simple and consistent.



Tip: Keep it relevant.



Keep the content and type of data requested relevant to the activities' expected results and standards of success. Focus on changes in competencies and behaviors, as well as knowledge. Make sure it is the most rigorous, but appropriate, test of what your program is trying to achieve, given the focus and constraints of your program. Make sure it is acceptable to your program's stakeholders in focus, format, and method of application.





- *A good sample size is larger than thirty, but you can include as many as you can afford. More than 1,000 are too many in most circumstances. For some instruments, such as logs or tally sheets, sampling is not appropriate.*
- *For meaningful results, a high proportion of your respondents needs to complete the instrument. For example, a 90 percent response rate (i.e., ninety out of 100 people given a survey returned a completed form) is very good and 30 percent is not enough.*

Choosing Methods

There are many methods for collecting data for an evaluation. Each method is useful for certain measurement tasks (or goals) and less appropriate for others. Each method also has its own set of advantages and constraints. You don't have to avoid a method because of its constraints. However, when you plan your evaluation, you should consider constraints and these constraints will be addressed. There may be a number of data collection instruments for any data collection method. The goal is to select the instrument that is most appropriate for your data. The table on pages 13 and 14 describes several common data collection methods.

Common Data Collection Methods

<p>Pre/Post Standardized Tests</p> 	<p>Characteristics: Pre-existing tests with a large group of respondents. Tests are administered at two points in time (i.e., the beginning and end of activities).</p> <p>Advantages: They offer a rigorous, ready-made context for documenting improvement. They are widely accepted as credible evidence if appropriate for the activity. They may allow for comparison across programs or schools.</p> <p>Constraints: The tests may not be designed to measure outcomes the program expects. They lose validity if changes in content, administration, or context occur.</p>
<p>Pre/Post Program- Based Tests</p> 	<p>Characteristics: An alternative to standardized tests. AmeriCorps programs can create such tests to document specific knowledge or performance but they capture gains directly related to the consequences of AmeriCorps program activities. These tests are administered at two or more points in time (e.g., the beginning, quarterly, and the end of activities).</p> <p>Advantages: The tests are widely accepted as credible evidence of accomplishments, if they are directly related to the services provided. They must be administered to respondents both before their participation (a “pre upon the conclusion of their participation (a “post</p> <p>Constraints: It is difficult to verify the degree to which the responses to test questions are an accurate representation of changes in knowledge or skills because of the program. They may not show changes in a consistent manner.</p>
<p>Logs or Tally Sheets</p> 	<p>Characteristics: A log documents a participant’s attendance or achievement such as “acquisition of skills.” It is especially appropriate for programs where it is difficult to identify exactly what will be learned at any point in time.</p> <p>Advantages: Logs are performance-based. They accommodate a range of starting and ending points and are easy to complete.</p> <p>Constraints: Data are unreliable and invalid if observation/recording is not systematic. Logs should include specific questions or categories directly tied to the results and indicators to prompt the user.</p>
<p>Rubrics</p> 	<p>Characteristics: Rubrics provide a detailed scale that can be used to measure performance. Rubrics are used either with other records, such as portfolios or written work, or with direct performance, such as conversation.</p> <p>Advantages: Rubrics can be used to measure a variety of abilities and behaviors. When well constructed, they are relatively easy to administer. (See “Instrument Development” in the reference section for more information.)</p> <p>Constraints: Developing a good rubric takes time. Off-the-shelf rubrics may be useful, but you need to match the rubric to the services you provide. The people administering the rubric must be thoroughly trained in its use.</p>
<p>Performance Ratings</p> 	<p>Characteristics: Set of questions regarding the manner in which AmeriCorps members carry out their activities. The focus is on issues such as attitude and ability to carry out specific tasks.</p> <p>Advantages: Data collection can be integrated with regularly scheduled meetings with the supervisor, or accomplished through a supervisor questionnaire.</p> <p>Constraints: Rating for performance standards must be explicit and consistent. The rating process must be short and focused. Supervisors are unlikely to be able to assess the persistence of any traits observed outside the job site. It may be difficult to link to outcomes of member development activities.</p>

<p>Interviews</p> 	<p>Characteristics: Data are collected orally. The interviewer asks clearly defined, systematic questions. Usually questions are predetermined and limited to a specific topic. Sometimes there are additional questions asked to elicit a more detailed response.</p> <p>Advantages: The data demonstrate specific examples of the observed outcome of AmeriCorps programs. Interviews allow for flexibility.</p> <p>Constraints: The interviewer must be skilled in the process of interviewing and conduct the interviews in a systematic manner to ensure unbiased results.</p>
<p>Surveys</p> 	<p>Characteristics: The data are collected in a written format. Each respondent provides data on a set of clearly defined questions.</p> <p>Advantages: The data can be used to demonstrate specific examples of the observed outcome of AmeriCorps programs.</p> <p>Constraints: It is difficult to balance specific and general questions and ensure that larger or unexpected issues are not missed. Survey instruments must be completed consistently to avoid biased results.</p>
<p>Focus Groups</p> 	<p>Characteristics: A moderator guides a group discussion involving six to ten individuals representing specific stakeholders.</p> <p>Advantages: Focus groups provide specific, pertinent data. Group interaction can produce more information than individual interviews.</p> <p>Constraints: A specific set of skills is required of the focus group moderator. Data are difficult to summarize succinctly.</p>
<p>Plugging Into Existing Information</p> 	<p>Characteristics: Other sources have collected the existing data, often statistical in nature. This may range from student grade point averages to neighborhood crime statistics.</p> <p>Advantages: It is often perceived as being more reliable and less subject to bias than other kinds of data. It can be less burdensome than other methods and prevents duplicating data collection.</p> <p>Constraints: The usefulness depends on whether the program being evaluated can reasonably be expected to influence the data directly. A variety of factors typically influence these indicators, and they may change very slowly even if an AmeriCorps program does have a great deal of impact on the problem being studied (e.g., crime statistics).</p>

Creating vs. Borrowing Your Instrument

As with carpentry, there is no one all-purpose tool; particular tasks need special tools. The more specific the instrument is to your measurement task, the more likely it is that you will achieve a high-quality result. There are two ways to acquire instruments – create your own or adopt an existing one. Consider these factors:

- **Availability:** Check the availability terms before committing to the use of an instrument. Some instruments require copyright payments.
- **Ease of use/low burden:** Select instruments that are easy to use for both the person administering the instrument and for the respondent.
- **Appropriateness:** Choose an instrument appropriate for the data to be collected. For example, an interview protocol is a better way to collect detailed data about attitudes than a written survey because of the interviewer’s ability to probe for more information.
- **Resources for analysis:** Administering a survey to a large number of individuals is worthwhile only if you are sure that you will be able to collect the data and then analyze a large data set. Otherwise, it may make more sense to get more detailed information from fewer people. Take stock of your available resources –both technical and human.

Tip: Pilot test.

Pilot test your method for effectiveness. (See the “Instrument Development” in the reference section for further assistance.)

Tip: Make it easy.

Limit the burden on those administering and responding, as well as the partner organization or its representatives.

