Connecting Students to the World of Work

Evaluation Mini-Guide #2: MEASUREMENT METHODS AND TOOLS

This mini-guide is one of a series designed to provide recipients of the OAC's Connecting Students to the World of Work grants with guidance for conducting evaluations of their funded programs. Each mini-guide offers basic information on a specific topic to help grantee organizations build capacity to implement evaluation activities.

This mini-guide focuses on measurement methods and tools. Measurement methods refer to the data collection activities and procedures (e.g., conducting interviews, administering surveys, reviewing documents), and measurement tools are the data collection instruments (e.g., the actual interview, survey, coding schema). Together, these two elements determine how and which data are gathered.

Evaluation research commonly employs a "mixed-method" approach, which includes collecting both qualitative and quantitative data. This method provides a balance of gathering the stories (qualitative) and the numbers (quantitative) jointly used to monitor the program's progress, guide its improvement, and understand its impact.

There are several different types of measurement methods, including:

- ***Surveys*** can be administered in person (paper and pencil), on line, or by mail; can vary in length; and are generally cost-effective.
- ***Focus groups*** are conducted in person with 5-10 attendees. They typically follow a semi-structured format, whereby a list of discussion questions/topics structures the conversation, but participants can bring up additional topics.
- ***Interviews*** are conducted with one participant at a time, and can be done in person or over the phone. They can follow either semi-structured or structured formats.
- ***Observations*** can be conducted with a variety of events and occur when the evaluator observes an event, either in person or via video, and rates specific elements of the event. For example, evaluators can observe program delivery (e.g., curriculum sessions, presentations) and staff can observe participants' presentations or performances. Collecting observational data requires that a protocol and a rating scheme to be developed beforehand, and that observers are trained on both.
- ***Document review*** can apply to various types of materials, including participant applications, reports, or portfolios. Similar to collecting observational data, evaluators must develop a review protocol and rating scheme beforehand and train all reviewers to ensure consistency. Then, materials are reviewed and rated for the specific, relevant elements.
- ***Archival data*** can also be obtained and analyzed, such as school records of grades and graduation status.

Surveys are a common approach, since they lend themselves to gathering both qualitative and quantitative data. For more in-depth information, focus groups and/or interviews can provide rich qualitative data. They are generally more expensive than surveys and are best conducted in person when possible. Observations and document reviews require staff to be adequately trained. Quantitative data can lend itself easily to measuring change over time and progress toward program goals, and qualitative data can provide insight into program functioning and participant perceptions.

<u>How do we determine what measurement methods and tools to use?</u>

Many factors can influence measurement selection. Two key considerations are:

1. <u>Budget:</u> Available funding and resources to implement measurement methods and to develop new or identify established measurement tools is a consideration. Typically, program evaluation activities are allocated approximately 10 percent of a program's operating budget. This is a widely accepted estimate in the evaluation community and can vary depending upon the program characteristics, evaluation goals, and measurement methods chosen.

2. Capacity: Staff availability and skills to collect and analyze data is another consideration. A common pitfall in evaluation is committing to collecting more data than what is actually feasible to analyze or manage. For example, if an organization wants to survey students, artists, and parents, an initial consideration would be to determine which staff person(s) has the skills and time to collect, manage, and analyze all the resultant data. Organizations may consider outsourcing evaluation activities to ensure that all data collected are also analyzed and reported in a way that is both accurate and meaningful for the program stakeholders.

How do we determine appropriate means of measurement? (How do we know what to ask?)

This is one of the most commonly asked questions regarding measurement design. Program outcomes, as defined in the program logic model (see Evaluation Mini-Guide #1), serve as essential guideposts for this determination. When possible, each defined outcome should have at least one associated measure, regardless of which measurement method is used. Identifying measures for each program outcome can include finding existing measures used in other studies and/or creating new measures customized to your program.

Example: An arts-related education program targets an outcome of "students will increase their interest in arts-related careers." The program would like to assess its progress in achieving this outcome, so it considers various measurement options. This outcome could be addressed using multiple measurement methods, for example:

- ***Focus Group:*** A small group(s) of student participants would be asked about their interest in arts-related careers and discussion would be facilitated to elucidate their perceptions. Ideally, this group would be held after their program participation has ended so that their retrospective account of program impact could be gathered. Sample questions could be:
  - How do you feel about careers in the arts?
  - What arts-related careers, if any, have you considered?
  - What role, if any, did participating in this program play in considering arts-related careers?
- ***Survey:*** Students would complete a baseline survey prior to program participation and a follow-up survey afterward. In the survey, they would rate their interest in an arts-related career on a numeric scale. During analysis, their baseline and follow-up ratings would be compared to examine change over time, or whether the program impacted their perceptions. Sample survey items could be:

A career in the arts…	**Strongly Disagree**	**Disagree**	**Neither**	**Agree**	**Strongly Agree**
…is interesting to me.	1	2	3	4	5
…is exciting to me.	1	2	3	4	5
…would make my life more meaningful.	1	2	3	4	5
...would allow me to use my education.	1	2	3	4	5
...is my professional goal.	1	2	3	4	5
…would not make me happy.	1	2	3	4	5

This is just an example of how this outcome could be addressed using two different measurement methods. Whichever method(s) is used, it's important to align the measure with program outcomes. It is equally important to ensure that measures are both feasible and meaningful. They should be obtainable within budget and program capacity to ensure that protocols will be implemented fully and that the data collected will be analyzed and used. They should also provide data that are meaningful and relevant so that organizations can easily use findings to illustrate program impact and support continuous program improvement. For more information and additional examples, please refer to the W.K. Kellogg Foundation Evaluation Handbook.[1]

[1] https://www.wkkf.org/resource-directory/resource/2010/w-k-kellogg-foundation-evaluation-handbook