Program Evaluation Toolkit
Module 5, Chapter 2: Data Quality
Considerations

Regional Educational
Laboratory
Central

From the National Center for Education Evaluation at IES

Speaker 1:

Welcome to the second chapter of module 5. In this chapter, you will review data quality considerations.

As you begin to address your evaluation questions, it is important to consider the quality of the data you will use. If the quality of the data are poor, your findings may not accurately represent the program resources, activities, outputs, or outcomes.

Data quality involves the extent to which data accurately and precisely capture the concepts you intend to measure. In program evaluation, these data relate to your resources, activities, outputs, and outcomes. With high-quality data, you can be more confident in the findings from your evaluation. In the AMMP! example, the school leaders hope to

- "increase the number of students who complete their homework with better than 80 percent accuracy," and
- "identify existing barriers that prevent students from completing homework."

If the information related to student homework completion or accuracy is inaccurate or incomplete, the AMMP! evaluation team will likely not be able to address the evaluation question related to homework completion.

For a full list of the AMMP! evaluation questions, see module 2.

There are six elements to consider in relation to data quality: *validity*, *reliability*, *timeliness*, *comprehensiveness*, *trustworthiness*, and *completeness*. You can use these elements to assess the quality of the data associated with your evaluation.

Let's start by looking at *validity*. Module 3 covered the validity of an evaluation design. In the context of data quality, validity is the extent to which a data source really measures what it is intended to measure. In the AMMP! example, if the measure of students' math knowledge is their performance on a math placement test, the evaluation team should make sure that the test assesses math knowledge as opposed to knowledge in another subject, like reading comprehension. The data source should be designed in a way that allows the team to feel confident in assessing students' math knowledge.

Now let's look at *reliability*. Reliability is the extent to which a data source yields consistent results. You can examine reliability in different ways. For assessments, look at whether the items consistently measure the same topic. This is called *internal consistency*. For example, if one

section of a math exam is focused on ratios and proportional reasoning, you expect a student's performance on questions in that section to be similar. Also look at whether the results would be similar if the same student took the assessment twice, across a period of time when you might expect scores to be stable. For example, if you use a diagnostic screening test to identify students with math learning difficulties, you expect that the test results would be similar if you tested the same student more than once. This is called *test–retest reliability*.

For interviews or observations, look at the extent to which multiple raters or observers are consistent in coding or scoring. If, for example, the AMMP! evaluation team conducts observations of tutors' instructional practices, the team can expect that each rater or observer will similarly score the quality of an individual tutor's instructional practices. This is called *inter-rater reliability*.

Now let's turn to *timeliness*. Timeliness is the extent to which data are current and the results of data analysis and interpretation are available when needed. For example, if the goal of AMMP! is to improve students' rates of homework completion, the evaluation team should continually collect data on homework completion while the program is being implemented in order to assess whether the program is having an effect on students' rates of homework completion. If the team does not collect data within an appropriate time frame, the data will not be helpful in addressing the evaluation questions. In another example, if the evaluation team wants to assess and adjust AMMP! implementation, the team should not examine high school graduation rates because it will be too late to change program implementation four or more years later, given that AMMP! participants are middle school students.

Now, let's consider *comprehensiveness*. Comprehensiveness means that the data collected in an evaluation include sufficient details or contextual information and can therefore be meaningfully interpreted. For example, if the AMMP! evaluation team wants to understand barriers that prevent students from completing their math homework, the team would likely be missing an important barrier if the team does not have access to or if the data does not include students' socioeconomic status.

Trustworthiness is the extent to which data are free from manipulation and entry error. Trustworthiness is often addressed by training data collectors, which will be discussed in module 6. Sometimes, data can be manipulated to produce a desired result—for example, if survey participants answer questions in a biased way to influence the results or a data collection team wants to emphasize a particular result. When examining the trustworthiness of your data, consider whether there was an opportunity or an incentive to manipulate the data during data collection. You will learn more about strategies to identify errors in data entry in module 7.

Completeness means that the data are collected from all participants in the sample and are sufficient to answer the evaluation questions. Completeness also relates to the degree of missing data and the generalizability of the dataset to other schools, districts, or state education agencies that may want to implement the program. In the AMMP! example, the evaluation team wants to analyze responses from a survey of math tutors, but less than 85 percent of the tutors responded to a particular question. As a result, the team might worry that the data are incomplete.

For more information on missing data and generalizability, review module 3 ("Evaluation Design") and module 4 ("Evaluation Samples"). For more examples and information regarding data quality, see the handout *Data Quality Dimensions*, available on the resources page of the website.

You will likely need to think a little differently about validity and reliability for qualitative data. Because data from interviews, focus groups, or, in some cases, observations may vary considerably due to participants' unique perspectives, behaviors, or actions, do not expect consistency. Instead, investigate any data that are very different from other related data.

Consider using triangulation to support the validity and reliability of your qualitative data. Triangulation involves reviewing multiple sources of data to look for similarities and differences. Do the results from multiple data sources point in the same direction? Do the results make sense in the context of your evaluation? If the answer to either of these questions are no, you might consider completing another technique for ensuring validity and reliability for qualitative data.

You may also do a member check, or ask participants involved in data collection to review your preliminary results to ensure that you have accurately captured their perspectives. As you attempt to code and summarize qualitative data, it is possible that you will not accurately capture participants' perspectives. Member checking provides participants with the opportunity to confirm whether you have or have not accurately done so.

Another way to support the validity and reliability of your qualitative data is to develop an audit trail. An audit trail is a documented history of your qualitative data collection and analysis. By carefully documenting your data collection procedures, training of data collectors, and notes, you will be able to cross-reference your findings with the conditions under which the data were collected. For instance, if multiple participants suggest that you have missed a key perspective during a member check, you may want to review the codes you created to ensure that you did not miss a key code. If you were not keeping an audit trail, this would not be an easy task.

Data quality is both objective and subjective. So far, you have learned about objective considerations of data quality. However, if stakeholders subjectively perceive that the quality of your evaluation data is poor, they will likely not trust the findings. For example, if the AMMP! evaluation team is collecting data to address the evaluation question "How many of each type of recreational activities were offered to AMMP! participants?" and stakeholders believe that school staff may have inflated counts of these activities to create a favorable outcome, the stakeholders may distrust the findings. Careful documentation of data collection procedures can help address these concerns. For example, the evaluation team might collect detailed logs of program activities related to recreational activities to share with skeptical stakeholders.

To help you evaluate the quality of your data, review the handout *Data Quality Checklist*, available on the resources page of the website. This checklist includes a set of guiding questions you can ask to check the quality of each data source you plan to include in your evaluation. In the real world, data quality will not be perfect. Nevertheless, you should feel that the data you use to address your evaluation questions provide an accurate picture of the program resources, activities, outputs, and outcomes. If you have doubts about the quality of your data, take steps to

improve the quality. In the AMMP! example, the evaluation team might change the test they are using to measure students' math knowledge, or the team might offer additional training to observers who are having difficulty in assessing tutors' use of instructional strategies. Or, if the evaluation team finds large amounts of missing data, the team might use a different data source to answer the evaluation questions. You will learn about ways to address some of these concerns in modules 6 and 7.

This concludes chapter 2 of module 5. In the next chapter, you will learn how to align your data sources to the evaluation questions you drafted in module 2.