

Program Evaluation Toolkit

Module 2, Chapter 2: How to Write Quality Evaluation Questions

Regional Educational
Laboratory
Central

From the National Center for Education Evaluation at IES

Speaker 1:

Welcome to the second chapter of Module 2. In this chapter, you will create quality evaluation questions, using the PARSEC framework.

In program evaluation, asking the right questions is crucial. One of the leading thinkers in the business world, Peter Drucker, has stressed the consequences of asking the wrong question. This applies to the education world as well. As you begin to draft evaluation questions, it is important to remember that you will not obtain actionable answers if you don't ask critical questions. The PARSEC framework, drawn from best practices in evaluation, will help you to draft quality evaluation questions.

PARSEC is an acronym reminding us that our questions should be pertinent, answerable, reasonable, specific, evaluative, and complete. Let's take a closer look at each of these criteria.

Pertinent questions are strongly related to the information that program stakeholders and participants want to obtain from an evaluation. If your questions are derived from your logic model, they will most likely be pertinent, particularly if you codeveloped the logic model with stakeholders.

Answerable refers to whether the data needed to answer questions are available or attainable. Good questions are useless if they cannot be answered.

Reasonable relates to what you can expect a program to achieve. Questions should be linked to what the program can practically and realistically achieve or influence.

Specific refers to clearly worded questions that avoid broad generalizations such as variations of "Is this good?" A specific question directly addresses a single component of the logic model.

Evaluative applies to how actionable the answers to questions will be. Actionable means that answers to a particular question can inform changes to a program, policy, or initiative if needed—or that the answers can confirm that a program, policy, or initiative is working as anticipated and no action is needed. The answer to a question should inform next steps to make the question worth the effort.

Complete refers to the entire set of evaluation questions. Again, the logic model is helpful here. To be complete, the set of questions should address all the logic model components.

The resource *Assessing Evaluation Questions*, available on the resources page of the website, includes a checklist you can use to revise evaluation questions against the PARSEC framework.

In the remainder of this module, you will see the process of revising evaluation questions using the PARSEC criteria.

Module 1 introduced a fictitious after-school program called AMMP!.

A middle school has been experiencing low rates of math homework completion among its students. In addition, students have been unsupervised after school. To address these issues, the middle school has started AMMP!.

AMMP! offers math tutoring, math extension, homework completion support, recreational activities, and field trips during after-school hours.

The grant that the school district received requires that the middle school evaluate the implementation of AMMP!. The school has assembled a team of stakeholders to design and implement the evaluation. The team needs to develop evaluation questions.

Here is the AMMP! logic model, introduced in Module 1. The AMMP! evaluation team uses the logic model to decide what questions to ask about the program. Evaluation questions should come directly from the logic model.

The AMMP! logic model is available, including citations, on the resources page of the website so that you can download it for further examination. Take a few minutes to review the logic model so that you can understand how the PARSEC criteria apply to the AMMP! and logic model.

The AMMP! evaluation team decides that, because AMMP! is new, the team needs to ask process evaluation questions to learn more about how the program is being implemented and to inform potential changes to the program. However, the team is also interested in the outcomes of the program and decides to ask outcome evaluation questions as well. The teams' two draft process evaluation questions are "How many students attended AMMP!?" and "Did tutors enjoy receiving professional development?" The team's two draft outcome evaluation questions are "Did students report having more math homework?" and "How much did students' test scores increase three years after the program was implemented?"

After reviewing these draft questions, the AMMP! evaluation team quickly realizes that the questions needed a lot of revision to truly capture what the team wants to know. The team decides to revise the questions by applying the PARSEC framework. Doing this will result in higher quality questions that allow the team to gather actionable and meaningful data about the program.

In the remainder of Module 2, you will see how the AMMP! evaluation team applies the PARSEC framework and how the questions evolve into the final evaluation questions. You will have a chance to create evaluation questions related to your own program after the examples.

Pertinent questions are strongly related to the information that program stakeholders and participants want to obtain from an evaluation. Here are the questions that the AMMP! evaluation team has drafted. Referring to the AMMP! logic model, what changes would you

make to improve these questions so that they meet the pertinent criterion? For example, you might change question 3 to “Did students complete more math homework?” because simply asking students to self-report whether they received more homework might not be an accurate measure of homework and therefore not be as meaningful to stakeholders. Think about the other questions. Then look at how the AMMP! evaluation team makes their questions more pertinent.

After reviewing the draft evaluation questions against the pertinent criterion, the AMMP! evaluation team revises the questions. This activity requires them to really focus on what they want to learn about AMMP!. First, the team looks at question 1 and realize that the total number of attendees really won’t tell the team what it wants to know. The team feels that it will be more important to know the attendance for each month of the program. For question 2, the team decides that the tutors’ knowledge of effective techniques is more important than the tutors’ enjoyment of the professional development. For question 3, the team decides that it cares more about whether students completed more homework than about whether they had more homework. Finally, the team decides that question 4 is already pertinent and does not revise it.

The answerable criterion refers to whether the data needed to answer questions are available or attainable. Good questions are useless if they cannot be answered. Before you look at what the AMMP! evaluation team decides, refer to the logic model and consider what changes you could make to improve the evaluation questions so that they meet the answerable criterion.

The AMMP! evaluation team reviews their revised questions against the answerable criterion. The team decides that data on monthly AMMP! attendance, tutors’ knowledge of effective tutoring techniques, and student homework completion rates are attainable, so the team does not change evaluation questions 1, 2, or 3. However, the team decides that looking at state test scores three years after the program is not feasible, given the delays in state test results, the scope of the program, the available funding, and the need to report results quickly. As a result, the team modifies question 4 to focus on end-of-year state test data so that the team can get immediate feedback about the outcomes associated with the program.

The reasonable criterion relates to the extent to which questions are linked to what a program can practically and realistically achieve or influence. For example, is it reasonable to expect AMMP! to influence state test scores in only one year? This question might be more reasonable in relation to the goals of the evaluation if its focus was shifted to a more immediate outcome.

When the AMMP! evaluation team considers the reasonable criterion, the team realizes that measuring improved knowledge of effective tutoring techniques as a result of the tutors’ professional development would be very hard. It is more reasonable to answer how many tutors received the professional development and how many professional development sessions they attended. In addition, after further discussion, the team realizes that even one-year state test results are not reasonable, given the delays in reporting and the many other factors that influence these results. Instead, the team decides that another achievement outcome from the logic model—preparedness for high school math—is more reasonable right now, so the team changes the question. The homework and attendance questions seem reasonable, so the team leaves them as they are.

The specific criterion refers to clearly worded questions that avoid broad generalizations such as variations of “Is this good?” Are the sample questions specifically aligned to components of the logic model? Is the wording too general? How can the questions be improved?

Although the questions are looking a lot better, the AMMP! evaluation team realizes that the questions are not specific enough according to the specific criterion. For example, the team feels that the professional development question needs to specify the topic of the professional development to ensure that the question measures what they want to see. After all, professional development that is not aligned to the goals of AMMP! should not count as data about program implementation.

Also, determining the amount of homework that students complete is important (remember that the students were completing only 25 percent of their homework before the intervention), but the team wants to know how well students did on homework as well. So, the team makes the question more specific by adding an accuracy threshold. The team also realizes that the question is not about the homework completion rate for all students but about the rate for only AMMP! participants. The team decides that comparing the homework completion rate for AMMP! participants with the rate for nonparticipants would be more specific.

Finally, the team looks at the student achievement question and decides that a comparison of student scores on high school math placement tests would be more specific.

The evaluative criterion applies to how actionable the answers to the questions will be. Answers should inform next steps in order to make them worth the effort. Can you envision actions that would result from the answers to the four AMMP! evaluation questions? Is it enough to know that students are accurately completing their homework? Or would it be more actionable to know why students are or are not completing homework?

The AMMP! evaluation team examines the questions and asks if the answers will inform changes to AMMP!. First, the team looks at professional development and realizes that the team needs to know why tutors might not be able to take advantage of professional development opportunities. This information will help the team consider changes to the delivery of professional development.

Next, the team decides that it is necessary to not only look at completion and accuracy of homework but also dig deeper to discern why AMMP! participants are not completing homework. Doing so might provide information to help the team make programmatic changes to AMMP!.

Finally, the team decides that the question about high school placement tests is the summative outcome the team wants to achieve. That question doesn’t need to inform program improvement, so the team leaves it as it is. In addition, the attendance question is evaluative as written.

The complete criterion refers to the entire set of evaluation questions. Again, the logic model is helpful here. To be complete, the set of questions should address all the logic model components

that are of critical interest. Looking at the AMMP! logic model, can you think of an additional question that would make the set of evaluation questions more complete?

The AMMP! evaluation team examines the developed questions and the logic model and decides that a few additional questions are needed. For example, the team realizes that the grant requires the team to report the total number of students who participate in the program. The team also decides to gather data about the number of field trips taken and about teacher promotion of AMMP!.

After applying the PARSEC framework and using the logic model, the AMMP! evaluation team comes up with a revised set of evaluation questions to use moving forward. Because AMMP! is a new program, the team knows that it needs to focus on process questions early in the evaluation to inform program improvements. Chapter 3 will discuss prioritizing the questions.

Now that you have looked at an example, it is time for you to draft and revise evaluation questions for your program. At the end of chapter 1, you drafted initial evaluation questions for your program, identifying the aligned logic model components and categorizing the questions as process or outcome. If you have not done this yet, you can find the *Identifying Evaluation Questions Worksheet* on the resources page of the website. Now use the PARSEC framework to examine the questions you drafted at the end of chapter 1. Keep these criteria in mind as you revise your questions and draft additional questions. A checklist for the PARSEC framework, called *Assessing Evaluation Questions*, is available on the resources page of the website. This checklist may help you as you revise your questions.

For now, focus on drafting a complete set of questions. You may notice that your complete set of questions does not address every element of the logic model. This is fine! Although it is critical to develop a logic model that completely describes the program, it is not always necessary to develop a set of evaluation questions that addresses every element. The reality of conducting an evaluation will require that you prioritize the questions. You can point out the elements that you did not address in your evaluation report. This will allow readers to more easily identify the limitations of your evaluation as well as the areas in which further research is needed. Prioritizing your evaluation questions will be addressed in chapter 3.

Complete this activity before moving on to the next chapter. If possible, allow some time to pass between chapters 2 and 3 so that you can reexamine your evaluation questions with a fresh perspective.

Once you have completed the activity for drafting evaluation questions, you are ready to move on to module 2, chapter 3: “How to Prioritize Evaluation Questions.”