

Program Evaluation Toolkit

Module 3, Chapter 3: Requirements for Evidence of Effectiveness and Design Standards

Regional Educational
Laboratory
Central

From the National Center for Education Evaluation at IES

Speaker 1:

Welcome to the third chapter of module 3. In this chapter, you will learn about requirements for evidence of effectiveness and design standards. Specifically, the chapter discusses the requirements for evidence of effectiveness under the U.S. Department of Education’s 2016 *Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments*, available on the resources page of the website, for implementation of the Every Student Succeeds Act (ESSA). The requirements for evidence of effectiveness under ESSA are referred to as the ESSA tiers of evidence. The chapter also discusses the What Works Clearinghouse (or WWC) design standards for evaluating the strength of evidence. Both ESSA tiers of evidence as operationalized in the 2016 non-regulatory guidance and WWC standards are linked to guidelines for conducting rigorous research and can be used to guide the design of an evaluation to meet specific evidence goals. ESSA tiers of evidence and WWC standards can also be used to assess the evidence behind existing programs and interventions.

The strength of evidence you require for your program may influence your selection of an appropriate evaluation design. If you want to provide stronger evidence of the program’s effectiveness, you must adhere to more rigorous standards of evaluation design.

First let’s look at the tiers of evidence under the Every Student Succeeds Act, or ESSA. ESSA encourages state and local education agencies to use evidence-based programs. Evidence-based programs have evidence of their effectiveness in producing results and improving outcomes when implemented.

There are four ESSA tiers of evidence: “strong,” “moderate,” “promising,” and “demonstrates a rationale.” You will learn about these tiers in more detail later in this chapter.

You may have experience with the ESSA tiers of evidence from using or selecting evidence-based programs. This toolkit focuses on generating new evidence. As a program is being implemented and evaluated, it is important to consider how the desired or required tier of evidence connects to the evaluation design.

Closely related to ESSA tiers of evidence are the What Works Clearinghouse, or WWC, design standards. Part of the Institute of Education Sciences within the U.S. Department of Education, the WWC reviews existing research on education programs, summarizes the findings of these reviews, and assigns evidence ratings to individual studies. These reviews help educators understand whether a given program has been found to be effective. Educators can then use this

information to select programs that are evidence based and, in turn, more likely to work for them.

According to WWC guidelines, a study may be found to meet WWC standards without reservations, meet WWC standards with reservations, or not meet WWC standards. Roughly speaking, studies that meet WWC standards without reservations may be eligible for strong evidence under ESSA, and studies that meet WWC standards with reservations may be eligible for moderate evidence.

For more information on the WWC standards, visit the website: <https://ies.ed.gov/ncee/wwc/>.

The evidence tiers in ESSA correspond to the WWC standards, as illustrated in this table. Now let's discuss each level individually.

To be supported by strong evidence, a program must have at least one study that “meets WWC group design standards without reservations.” This rating is limited to well-implemented randomized controlled trials, or RCTs, with low attrition (the number of participants who leave a study before it is completed). In addition, the study must have a large, multisite sample—that is, the sample must include at least 350 individuals spread out over more than one state, county, city, school district, or postsecondary campus—and at least one outcome that is statistically significant and positive.

To be supported by moderate evidence, a program must have at least one study that “meets WWC group design standards with reservations.” This rating includes well-implemented RCTs with high attrition or well-implemented quasi-experimental designs, or QEDs, in which baseline equivalence is met. Baseline equivalence involves reviewing the treatment and comparison groups before implementing a program, baseline, to ensure that the groups are similar enough, equivalence, to make accurate comparisons. The study must also have a large, multisite sample and at least one outcome that is statistically significant and positive.

To be supported by promising evidence, a program must have at least one study that used a correlational design comparing outcomes for a treatment group and a comparison group, and that used statistical controls for selection bias. The statistical controls can help account for preexisting differences between the treatment and comparison groups. Studies that provide promising evidence must show a statistically significant and positive effect. However, these studies do not meet WWC standards.

To demonstrate a rationale, a program must have a well-specified logic model with one intended outcome of interest that aligns with a stakeholder need. The program must be supported by existing or ongoing research demonstrating how it is likely to improve the outcomes identified in the logic model.

A number of considerations related to evidence and standards can influence an evaluation design. See the handout *Guiding Questions: Evidence and Standards*, available on the resources page of the website, for a series of questions to help you align your evaluation design to the desired tiers of evidence and standards.

Now, let's try to piece together what you have learned from this module. See the *Evaluation Design Selection Worksheet*, available on the resources page of the website, for a series of questions to help you determine which evaluation design category makes the most sense for you. Use the evaluation questions that you developed during Module 2.

You have now completed chapter 3 of module 3. This module is now complete. In module 4, you will explore sampling and develop a sampling plan for your evaluation.

This handout was prepared under Contract ED-IES-17-C-0005 by Regional Educational Laboratory Central, administered by Marzano Research. The content does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.