

What Educators and Administrators Really Need to Know When It Comes to Evaluating ESSA Evidence Standards

The Elementary and Secondary Education Act was passed in 1965 and created a longstanding commitment to equal opportunity for all students. To build upon this important work, the bipartisan Every Student Succeeds Act (ESSA) was signed on December 10, 2015. The intent of ESSA is to ensure that educators are implementing high standards and policies that prepare all students for success in post-secondary education and their future careers. ESSA supports our nation's schools and our dedicated educators in their continuing vision for excellence, emphasizing equity and quality for every student. ESSA outlines specific provisions to ensure success for students and schools, including the adoption of evidence-based educational solutions, especially those purchased with federal funding.¹

To assist stakeholders in successfully selecting evidence-based educational solutions that improve outcomes for learners, the U.S. Department of Education (DOE) issued guidance in 2016, and updated it in 2023, to clarify that a "cycle of continuous improvement" should include selecting relevant, evidence-based educational solutions that are most likely to address local needs.²



¹ <https://www.ed.gov/essa>

² U.S. Department of Education (2023, Sep 28). Non-regulatory guidance: Using evidence to strengthen education investments. <https://www2.ed.gov/fund/grant/about/discretionary/2023-non-regulatory-guidance-evidence.pdf>

ESSA has established study evidence levels that are determined by study design, results, findings from related studies, sample size, setting, and organizational match.

The ESSA Tiers of Evidence provide districts and schools with a framework for determining which educational solutions will work for their specific student population and unique setting.³ ESSA has established study evidence levels that are determined by study design, results, findings from related studies, sample size, setting, and organizational match (i.e., how well the study sample matches the specific setting and population of the districts/schools reviewing the research). The following diagram describes the ESSA evidence levels.



³ Regional Educational Laboratory (2019). ESSA tiers of evidence: What you need to know. American Institute for Research. <https://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/blogs/RELMW-ESSA-Tiers-Video-Handout-508.pdf>

Common Misconceptions About ESSA

The complexity of ESSA and the associated evidence levels can leave even the savviest educators and decision makers wondering how to navigate existing research. Misconceptions around ESSA levels can lead to challenges applying available research in a positive, meaningful way. The following common misconceptions surrounding ESSA's evidence levels can pose a barrier when it comes to making informed decisions on future educational solutions.

Misconception #1:

Level 1 is always better than Level 2 and Level 3.

ESSA evidence levels were designed to categorize the research conducted, but do not necessarily indicate how well a learning solution will do in a variety of settings. ESSA levels can tell us what study design was used, the results of the study, findings from related studies, sample size and setting, and how well the study population matches the population at the district/school that is reviewing the research.

When evaluating potential educational solutions using ESSA, it's important to remember that ESSA Levels 1, 2, and 3 share more similarities than they do differences. Specifically, all Level 1, 2, or 3 ratings require positive effects with no overriding negative effects from related studies.⁴ ESSA Levels 1 and Level 2 are also identical in terms of sample size requirements and setting. In fact, the only difference between Level 1 and Level 2 is the type of study that was conducted and how it was designed (randomized control trial/RCT vs. quasi-experimental design/QED).

ESSA evidence levels were designed to categorize the research conducted, but do not necessarily indicate how well a learning solution will do in a variety of settings...it's important to remember that ESSA Levels 1, 2, and 3 share more similarities than they do differences.

	Level 1	Level 2	Level 3
Study Design	Must be a well-designed and well-implemented study		
Group Assignment	Random assignment of treatment and control groups	Follows groups as they naturally occur; however, these groups are "matched" to ensure students in both groups are as similar as possible	Follows groups as they naturally occur; however, these groups are not matched
Results of the Study	Statistically significant and positive effects		
Findings	Positive effects with no overriding negative effects		
Sample and Setting	At least 350 participants, conducted in more than one district or school		No sample size requirements

⁴ U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. *What Works Clearinghouse Procedures Handbook* (Version 5.0, 2020). https://ies.ed.gov/ncee/WWC/Docs/referenceresources/Final_WWC-HandbookVer5_0-0-508.pdf

The U.S. Department of Education (DOE) clearly indicates that schools/districts/states should consider the entire body of relevant evidence when making informed decisions on learning solution selection.

When it comes to educational research, randomized control trials (RCTs) and quasi-experimental design (QEDs) studies are both very likely to compare data from schools that share similar demographics and performance factors. While RCTs prevail as the current gold standard of research, they can come with disadvantages for participating teachers and students. Specifically, some teachers and students may have to continue using a learning solution that needs to be replaced, denying some students beneficial educational interventions.

In comparison to RCTs, QEDs can be more generalizable to the real-world, since in the real-world teachers and administrators typically take an active role in curriculum selection which can ultimately affect implementation and buy-in.⁵

Misconception #2:

The level of the study is the only criteria when evaluating the evidence of a learning solution:

As the selection of educational solutions becomes more evidence-based, there is a growing need to scrutinize the results from rigorous educational evaluations. This is especially true when it comes to understanding how the characteristics of schools or districts that participate in rigorous research compare to the actual population, demographics, and other unique characteristics of the districts and schools reviewing the research. A single study rated Level 1 with a sample of study participants that DO NOT reflect the population and setting for which the learning solution is being considered, can be less relevant than several studies at various levels that DO reflect the characteristics of the district/school the program is intended for.⁶ Regardless of ESSA evidence level, a larger body of research (many studies) across multiple sites that match the setting and population of the organization where the educational solution under review will be applied can be stronger than a single Level 1 study.

A large body of research supporting an educational solution, including ESSA Levels 2 and 3, is much more informative, when it comes to learning solution selection, than a single study showing effectiveness rated Level 1. The U.S. Department of Education (DOE) clearly indicates that schools/districts/states should consider the entire body of relevant evidence when making informed decisions on learning solution selection, specifically:⁷

- **Research evidence that was conducted in a similar setting and/or population.** When evidence supporting a learning solution is aligned

⁵ Steven, M. R. (2022). Evidence Viewpoint: Choosing the Optimal Design for an Efficacy Study: Gold Rushes vs. Silver Linings. A blog from the Center for Research and Reform in Education at Johns Hopkins University. <https://education.jhu.edu/news/evidence-viewpoint-choosing-the-optimal-design-for-an-efficacy-study-gold-rushes-vs-silver-linings/>

⁶ Elizabeth A. Stuart, Stephen H. Bell, Cyrus Ebnesajjad, Robert B. Olsen & Larry L. Orr (2017) Characteristics of School Districts That Participate in Rigorous National Educational Evaluations, Journal of Research on Educational Effectiveness, 10:1, 168–206, DOI: 10.1080/19345747.2016.1205160

⁷ U.S. Department of Education (2023, Sep 28). Non-regulatory guidance: Using evidence to strengthen education investments. <https://www2.ed.gov/fund/grant/about/discretionary/2023-non-regulatory-guidance-evidence.pdf>

to a school/district's setting (e.g., elementary school), capacities (e.g., available technology, financial means, etc.), or learner population (e.g., students with disabilities, English learners, students experiencing homelessness, migratory students, etc.), it may increase the likelihood that the reviewing school/district will see positive outcomes similar to those reported in studies they are reviewing. If a school/district's setting and/or population is different from that of the study participants, it could mean a lower evidence level for that particular school/district even if the learning solution used in the study is supported by evidence.

- **A body of research showing student success across multiple student populations and settings.** Multiple studies across various levels are preferable to a single Level I study showing student success in a specific student population and setting.
- **Note the relevancy of research studies that have been conducted for a given learning solution.** This includes when the research was conducted as it is an important factor when determining what programs work for whom and under what conditions.

It is essential that educators reviewing research for a potential program go to great efforts to make sure it was tested with a population and setting similar to their own and that it reflects the educational climate of today (i.e., beware of outdated research).

Misconception #3:

All third-party reviewers are endorsed and aligned with the standards put forth by the U.S. Department of Education.

Not all third-party reviewers are equal when it comes to endorsing educational research. The U.S. Department of Education (DOE) bases their evidence definitions and determinations on three major sources, including the What Works Clearinghouse (WWC). As part of the DOE, the WWC's mission is to review educational research on the effectiveness of education policies, programs, products, and practices to see what works to improve student outcomes and other outcomes relevant to education.

Following recently updated DOE guidelines, organizations should explore the broadest possible range of relevant evidence, and this should include research reviewed by the WWC. Because the WWC is endorsed by the DOE and addresses research design by incorporating ESSA evidence levels into their rigorous standards, they are able to offer evidence of effectiveness on a wide range of programs. WWC reviews include specific educational solutions and individual studies, as well as a summary of the setting/s in which research was conducted and details about the demographics of participating districts/schools/students when available.

Not all third-party reviewers are equal when it comes to endorsing educational research.

When reviewing
learning solutions,
look for:

-  1. Evidence that reflects local school setting and/or population
-  2. Evidence from multiple studies
-  3. DOE endorsed third-party reviews

Conclusion

In the coming years, it is likely the emphasis on evidence-building will continue as districts and states determine how to best spend federal funding when it comes to purchasing research-based educational solutions. Based on DOE recommendations, educators and administrators should focus on programs that are supported by evidence that reflects their local school setting and/or population. When the existing body of research for a potential educational solution reflects multiple student populations and settings, it's easier for decision-makers to more accurately determine if: 1) the educational solution meets their needs; and 2) the positive outcomes seen in research will be replicated in their own setting. This means keeping in mind that a larger body of evidence, including ESSA Levels 2-4, can provide a greater breadth of insight than a single Level 1 study and that ESSA Levels 1, 2, and 3 share more similarities than they do differences. Educators and administrators should also look to DOE endorsed third-party reviewers like the WWC as they navigate existing research and make sure what they are reviewing includes research that is current and relevant in today's educational climate.

Savvas Learning Company is dedicated to conducting research that is aligned with WWC standards and uses a variety of methods to ensure its studies produce high-quality data that conscientiously reflects the unique experiences of students and teachers across the nation.

Explore the ESSA Levels of Evidence for Savvas learning solutions [here](#).



Savvas.com
800-848-9500

Copyright © 2025 Savvas Learning Company LLC. All Rights Reserved. Savvas® and Savvas Learning Company® are the registered trademarks of Savvas Learning Company LLC in the US and in other countries.
09-25v2

Join the Conversation
@SavvasLearning

