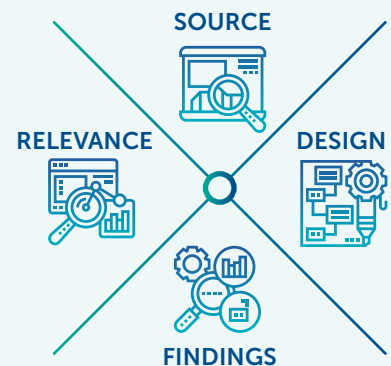


The Basics of Reviewing A Research Study

Purpose

This quick-reference guide will provide you with some basic elements of high-quality research to inform your selection of evidence-based interventions. It provides considerations under four main topics when you are reviewing an individual study that tests the effectiveness of an intervention:

- 👉 Study Source: Where the research comes from.
- 👉 Study Design: How the study was designed and carried out.
- 👉 Study Findings: The measures of the intervention's effectiveness and strength.
- 👉 Study Relevance: How the research may apply to your context.



Intended Use

This guide is intended to help state education agency, school district, and school staff review research studies to identify high-quality evidence-based interventions that meet their needs. These are some basic considerations when reviewing research studies, but this is not an exhaustive list.

The guide provides, for each topic area, the elements to look for in a study, questions to ask about each element, why it matters to consider each element, implications, and resources.

Study Source: Where the research comes from.



Element	Questions to Ask	Why It Matters
Credibility	<p>Where did you find the research?</p> <p>Was it from a credible and reputable source?</p>	<p>Ensuring the research is from a reputable and credible source gives you more confidence in its findings. Examples of credible sources include the Education Resources Information Center (ERIC), What Works Clearinghouse (WWC), Evidence for ESSA, and peer-reviewed journals. If you read a summary or second-hand description of a study, it is best to review the original article or white paper.</p>
Peer Review	<p>Were the results of the study subject to peer review (i.e., formally reviewed by other researchers for its quality)?</p>	<p>Articles published in peer-reviewed journals are usually considered less biased than those published from other organizations, especially partisan ones. The Institute of Education Sciences also has a peer review process for its research publications.</p>
Independent Researchers	<p>Was the study conducted by independent third-party researchers?</p> <p>Who funded the study?</p> <p>How was objectivity attended to in the study?</p>	<p>Even if published in peer-reviewed journals, articles and/or white papers authored by intervention developers, researchers associated with the intervention being evaluated, or whose authors received funding from intervention developers may be more biased towards positive findings than those authored by independent or third-party researchers or evaluators. Look for information that demonstrates the study authors ensured objectivity in the research.</p>

CONFIDENCE IN STUDY'S CREDIBILITY

Less confidence

More confidence



- Study not from a credible source
- Not peer reviewed
- Researcher developed the intervention

- Study from a credible source
- Peer reviewed
- Researcher was independent from intervention development

Study Design: How the study was designed and carried out.



Element	Questions to Ask	Why It Matters
---------	------------------	----------------

Study Design

What study design did the researchers use?

Was the design experimental (that is, used a randomized controlled trial), quasi-experimental, correlational, or descriptive?

Study design is considered when assigning an Every Student Succeeds Act (ESSA) evidence tier.¹ Studies that use a group that did not receive the intervention (i.e., a comparison group) are usually considered stronger evidence than those that do not use any comparison group. The designs below are listed in order of their strength and ESSA evidence tier. The stronger the research design, the more confidence you can have that the intervention itself was responsible for the outcomes.

- A randomized controlled trial (RCT) assigns people, schools, or districts to receive the intervention or not based on a lottery, not on any characteristics (typically Tier I). It is considered the strongest research design.
- A quasi-experimental design (typically Tier II) assigns people, schools, or districts to receive the intervention or not based on key characteristics (e.g., age, race, pre-test scores) to match the groups closely to one another. No lottery is used in the process.
- A correlational study is a nonexperimental study that looks at the association between an intervention and outcomes. Typical Tier III correlational studies include statistical controls to account for characteristics that differ between the group that received the intervention and the group that did not.
- A descriptive study examines data at one point or over time and is one of the weakest research designs. Studies that compare data from before and after the intervention from one group are usually descriptive studies. Studies without a comparison group of people, districts, or schools that did not receive the intervention do not tell us if any of the observed changes would have happened anyway, without the intervention. These types of studies are typically Tier IV, when they include a theory of action and ongoing evaluation.

¹ See non-regulatory guidance on ESSA evidence provisions, available at <https://www2.ed.gov/policy/elsec/leg/essa/guidanceuseinvestment.pdf>, for more descriptions of these research designs and how they relate to ESSA evidence tiers.

Study Design (continued)



Element	Questions to Ask	Why It Matters
Baseline Equivalency	<p>Did the study provide a comparison of the groups before one received the intervention?</p>	<p>In quasi-experimental design studies, you can be more confident of a study's findings if the researchers establish that the types of people (or schools or districts) in the treatment and comparison group were not different in meaningful ways—also called “baseline equivalency.”</p>
Sample	<p>Did the study rely on a large enough sample to draw its conclusions?</p> <p>How many people dropped out of the study (referred to as attrition)?</p>	<p>A sample is a set of people, districts, or schools selected from the larger population for measurement purposes. The sample size should be large enough to detect meaningful differences. A large sample, based on the ESSA non-regulatory guidance, is at least 350 participants from more than one school or district.² In addition to sample size, you want to look at how many people dropped out of the study (i.e., attrition) because the more attrition from a study, the weaker the evidence of effectiveness becomes. Generally speaking, an overall attrition rate below 20% is acceptable. The WWC has developed standards to define high and low study attrition.³</p>
Outcome Measures	<p>Are the outcomes measured in the study relevant to the intervention?</p> <p>Were the study's outcome measures designed by the intervention's developers, or did the researchers rely on established measure(s) from outside sources?</p>	<p>More confidence can be placed in a study's findings when established and verified measures from outside sources were used to measure effects that are relevant to the intervention. For example, a study of an algebra intervention should use an established algebra assessment that measures relevant algebra skills. Outcomes developed by the intervention developer are sometimes too closely tailored to that specific intervention being tested. This may create bias towards positive findings for that intervention. This can be the case even if a comparison or control group was used in the study.</p>


2 The non-regulatory guidance on ESSA evidence provisions references Part 77.1 of EDGAR which defines a large sample as 350 or more students (or other single analysis units), or 50 or more groups (such as classrooms or schools) that contain 10 or more students (or other single analysis units). <https://www2.ed.gov/policy/elsec/leg/essa/guidanceusseinvestment.pdf>

3 https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_attrition_080715.pdf

CONFIDENCE IN STUDY'S DESIGN

Less confidence

More confidence

- 
- Weaker research design
 - No comparison group
 - Smaller sample
 - Many study participants drop out of the study
 - Use of an outcomes measure designed by the developer
 - None of the outcomes measured are relevant to the intervention

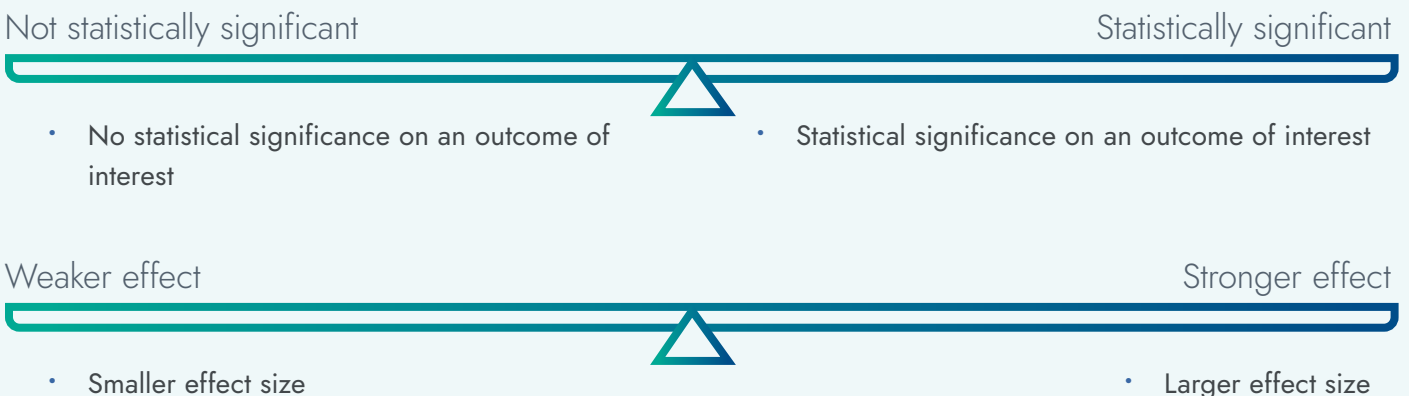
- Stronger research design
- Use of a comparison group
- Larger sample
- Few study participants drop out of the study
- Use of an established outcome measure not designed by the developer
- At least one of the outcomes measured is relevant to the intervention

Study Findings: The measures of the intervention's effectiveness and strength.



Element	Questions to Ask	Why It Matters
Statistical Significance	<p>Was there a positive and statistically significant effect on a relevant outcome?</p> <p>Were the results positive across all relevant outcomes?</p>	<p>Quantitative research articles almost always report statistical significance⁴ and p-values. The p-value is the probability of observing a difference at least as large as the one between the two sample groups even if the true difference were zero. A p-value of $<.05$ is usually considered statistically significant. Sometimes a study summary or abstract will report positive intervention findings for one or a few outcomes. However, closer examination of the article may reveal that there were no effects on other outcomes, or even negative effects.</p>
Effect Size	<p>What was the effect size or magnitude of the positive impact?</p>	<p>Effect size is a measure of the strength or magnitude of the intervention's effect. The larger the effect size, the more robust the intervention. Effect size is most helpful when it is converted to a real-world metric (e.g., days of instruction; increase in a standardized test's scale score). Statistical significance and effect size are independent of each other. Effect size is typically reported as a proportion of a standard deviation. For example, an effect size of $.20$ is 20% of a standard deviation of an outcome. ESSA evidence tiers do not consider or include effect size.</p>

STATISTICAL SIGNIFICANCE OF THE FINDINGS AND MAGNITUDE OF THE EFFECTS



⁴ Statistical significance is based on statistical tests that consider both the magnitude of difference between group average outcomes (e.g., test scores) and the variability of those statistics.

Study Relevance: How the research may apply to your context.



Element	Questions to Ask	Why It Matters
Match to Your Population and Setting	<p>How well does the sample examined in the study match your population of interest?</p> <p>How well does the study setting match your setting?</p>	<p>If an intervention was successful with a group of students or educators or in a setting that was very different than yours, it may not be as successful with your population or in your setting. Alternatively, you may find you have to adapt the intervention to make it more suitable to your population or setting and less similar to how it was carried out in the study. However, these adaptations could weaken the intervention's effectiveness.</p>
Match to Your Needs	<p>How well does the intervention match your needs?</p> <p>How closely do the outcomes examined in the study match the outcomes you are interested in?</p>	<p>To the extent possible, you want to look for interventions that improve the exact outcomes you have targeted to improve. If the outcomes are not exact, they should be close approximations but the interventions may not be as effective on these other outcomes.</p>

RELEVANCE OF THE STUDY'S FINDINGS

Less relevant

More relevant



- Population in the study does not match your population
- Setting of the study does not match your setting
- Outcomes in the study do not match your outcomes of interest

- Population in the study matches your population
- Setting of the study matches your setting
- Outcomes in the study match your outcomes of interest

Additional Resources

Study Source

Education Resources Information Center (ERIC). This searchable database of education research and information includes publicly available research studies, including peer-reviewed studies.

Available at: <https://eric.ed.gov/>

What Works Clearinghouse (WWC). This clearinghouse reviews existing research on interventions, programs, practices, and policies using rigorous standards for research quality. Available at: <https://ies.ed.gov/ncee/wwc/>

Evidence for ESSA. This resource reviews research on programs against ESSA evidence standards. Available at: <https://www.evidenceforessa.org/>

Study Design

Key Components of a Research Paper. This reference outlines the components to help determine the credibility of a study. Available at: https://ies.ed.gov/ncee/edlabs/infographics/pdf/REL_SE_Key_Components_of_a_Research_Paper.pdf

ESSA Tiers of Evidence: What You Need to Know. This handout provides an overview of the ESSA tiers. Available at: <https://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/blogs/RELMW-ESSA-Tiers-Video-Handout-508.pdf>

Going Public: Writing About Research in Everyday Language. This brief includes a glossary that defines complex research concepts for practitioners, policymakers, and other nontechnical readers. Available at: https://ies.ed.gov/ncee/pubs/REL2014051/pdf/REL_2014051.pdf

Study Findings

Understanding Education Statistics. This set of presentation slides provides an overview of statistical language and tests. Available at: https://education.ky.gov/school/evidence/Documents/Understanding_Education_Statistics.pdf

Interpreting Effect Sizes of Education Interventions. This working paper offers guidelines for interpreting effect sizes in education. Available at: <https://annenbergbrown.edu/publications/interpreting-effect-sizes-education-interventions>

Study Relevance

Applicability of Evidence-Based Interventions. This tool provides considerations for the fit of an evidence-based intervention to your context. Available at: https://ies.ed.gov/ncee/edlabs/infographics/pdf/REL_WE_Applicability_of_Evidence_Based_Interventions.pdf

Tool 6: Comparing Evidence-Based Interventions Found in Evidence-Based Improvement: A Guide for States to Strengthen Their Frameworks and Supports Aligned to the Evidence Requirements of ESSA. This tool helps to identify the feasibility of implementing an intervention in a local context. Available at: <https://www.wested.org/resources/evidence-based-improvement-essa-guide-for-states/>