Non-Regulatory Guidance:
Using Evidence to Strengthen Education Investments

Revised September 28, 2023

Purpose of the Guidance

The U.S. Department of Education (Department) is issuing this guidance to provide information to State educational agencies, local educational agencies, institutions of higher education, schools, educators, partner organizations, and other partners to assist them in selecting, using, and building evidence-based activities, strategies, and interventions, as defined in Title VIII of the Elementary and Secondary Education Act of 1965, as amended. If you are interested in commenting on this guidance, please email us your comment at Evidence@ed.gov *or write to us at the following address:*

U.S. Department of Education
Office of Planning, Evaluation and Policy Development
Grants Policy Office
400 Maryland Avenue, S.W.
Washington, D.C. 20202

The Department does not mandate or prescribe practices, models, or other activities in this non-regulatory guidance document. This guidance contains examples of, adaptations of, and links to resources created and maintained by other public and private organizations. This information, informed by research and gathered in part from practitioners, is provided for the reader's convenience and is included here to offer examples of the many resources that educators, parents, advocates, administrators, and other concerned parties may find helpful and use at their discretion. The Department does not control or guarantee the accuracy, relevance, timeliness, or completeness of this outside information. Further, the inclusion of links to items and examples do not reflect their importance, nor are they intended to represent or be an endorsement by the Department of any views expressed, or materials provided.

Introduction

This version of the Non-Regulatory Guidance is a revision to the Non-Regulatory Guidance first issued in 2016, after the reauthorization of the Elementary and Secondary Education Act, as amended (ESEA). This update retains the structure of the Non-Regulatory Guidance and reframes it to clarify that the cycle of continuous improvement and evidence definitions have applicability beyond programs authorized by ESEA, including career and technical education, postsecondary education, and special education. In addition, this version of the Non-Regulatory Guidance provides current information about the evidence provisions in the Education Department General Administrative Regulations (EDGAR).

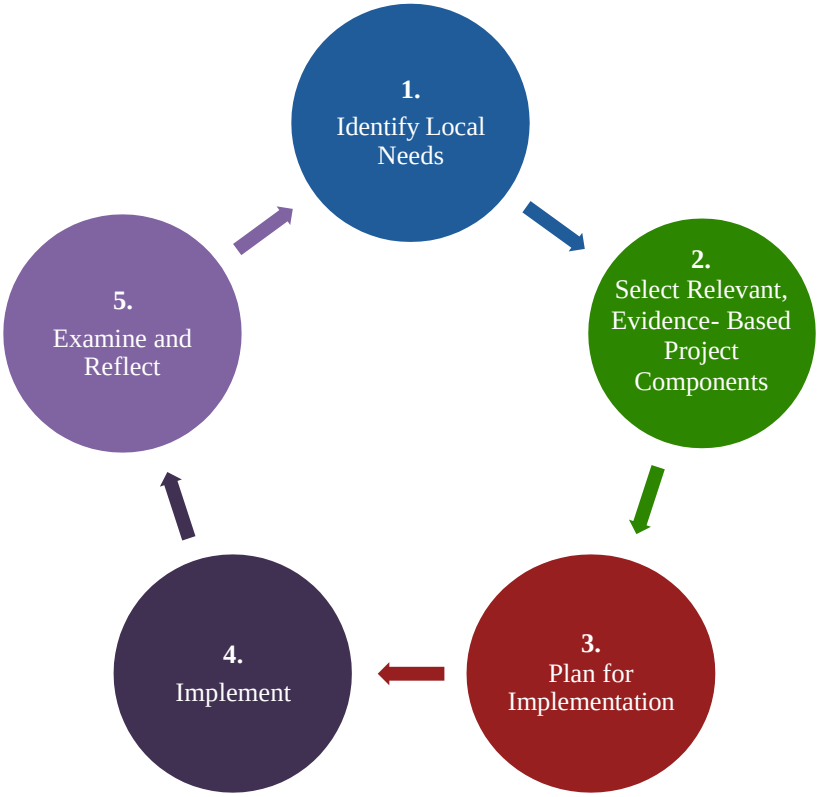
Using, generating, and sharing evidence about effective strategies to support learners of all ages gives stakeholders an important tool to improve their academic and non-academic outcomes. The Department emphasizes the use of evidence-based activities, strategies, and interventions (collectively referred to as "*project components*"[i]) in the design of education programs from pre-kindergarten through adult education.

This guidance is designed for state educational agencies; local educational agencies; institutions of higher education; schools; classroom educators; partner organizations, including nonprofit organizations; and other stakeholders (collectively referred to as "organizations"). Its purpose is to help those organizations successfully choose and implement evidence-based *project components* that are designed to improve outcomes for learners.

Part I of this guidance reviews steps for effective decision-making about evidence use and evidence building. Part II of this guidance reviews criteria for identifying the strength of evidence that support evidence-based *project components*. In Part II, we describe in more detail four evidence levels introduced in Section 8101(21)(A) of the ESEA, and used by the Department in its discretionary grant programs, including postsecondary programs, through EDGAR (34 CFR 77.1). This guidance also supports the goals of the Foundations for Evidence-Based Policymaking Act by clarifying how a variety of evidence can be used to inform decision-making. *Italicized* words found throughout the document are defined in its endnotes.

Part I: Strengthening the Effectiveness of Education Investments

Educators and education policymakers can strengthen the effectiveness of education investments when practice and policy are rooted in *project components* with demonstrated effectiveness for the learners they serve and the context in which they serve them. To do so, the Department recommends a process of learning and improvement that includes: (1) identifying local needs, including the needs of the learner population being served; (2) selecting evidence-based *project components* that organizations have the capacity to implement, including implementing with partners; (3) planning for and then (4) supporting and refining the implementation of the *project components;* and (5) examining and reflecting upon how the *project components* are working. These steps, when taken together, promote a cycle of continuous improvement designed to support better outcomes for learners.



Step 1. IDENTIFY LOCAL NEEDS

Organizations should engage in timely and meaningful consultation with a broad range of stakeholders—including families, students, educators, State and local government agencies, and community partners—and examine relevant data to understand the most pressing needs of learners, schools, and/or educators and the potential root causes of those needs.

Interviews, focus groups, and surveys, as well as additional information on learners (e.g., assessment results, graduation rates, college completion rates), schools (e.g., resources, climate, mobility among schools) and educators (e.g., effectiveness, retention rates) provide insights into local needs, including the needs of the learner population being served.

Questions to consider throughout the local needs assessment process include:

- How do you reach all stakeholders, particularly from underserved groups, to contribute to the work of identifying local needs or their root causes and potential ways to address them? How are you ensuring access for a diverse group of stakeholders to contribute?
- What data do you already have that can be used to better understand local needs or their root causes? What questions do you need to collect data on to better understand the local needs or their root causes?
- What assumptions have you made in drafting these questions or in the analysis of your data? Are there any areas of potential bias in those assumptions or inaccuracies that need to be corrected?
- How do learner outcomes compare to identified state, local, school, or institutional performance goals?
- Are there inequities in learners' access to resources, participation in programs, or academic and non-academic outcomes within the organization?
- What are the potential root causes of performance gaps or observed inequities, including those causes that may immediately and directly affect learners (e.g., peers, educators, family) as well as those that may operate more indirectly (e.g., policies, laws, current and historical events).
- How should needs be prioritized when several are identified?

Step 2. SELECT RELEVANT, EVIDENCE-BASED PROJECT COMPONENTS

Once needs have been identified, an organization should identify evidence-based *project component(s)* that are most likely to address them. Generally, organizations should:

- Explore the broadest possible range of relevant evidence;
- Select *project components* that are supported by rigorous and relevant evidence, emphasizing *project components* that have been demonstrated to be effective for the learners they plan to serve and the context in which educators will serve them; and
- Select *project components* that are feasible to implement given the organization's capacity (e.g., funding, human resources, expertise, and stakeholder support).

Organizations should explore the broadest possible range of relevant evidence. Potential sources of relevant evidence include:

- The Department's What Works Clearinghouse™ (WWC).[ii] The WWC uses rigorous standards to review education research, offering evidence of effectiveness on a wide range of *project components*. WWC Practice Guides, which typically focus on discrete topics (e.g., reading interventions, career pathways) in particular student populations and contexts (e.g., early elementary school, community colleges), may be particularly useful in identifying useful sets of *project components*. The WWC also reviews specific branded and non-branded interventions and individual studies. (WWC reviews also summarize the settings and populations in the studies when that information is available.)
- Other federal evidence clearinghouses, including the Department of Labor's Clearinghouse for Labor Evaluation and Research (CLEAR), the Department of Justice's CrimeSolutions clearinghouse, the Department of Health and Human Services' Evidence-Based Practices Resource Center, or AmeriCorps's Evidence Exchange.
- Other non-federal entities have their own clearinghouses that review and curate education research using their own standards, summarizing their findings for educators or policymakers. As referenced in the first bullet, the WWC conducts its own reviews and posts the findings of those reviews in the WWC™ Clearinghouse.

Organizations should select *project components* that are supported by the most rigorous evidence available, consider the needs of the learner population being served, and consider the ability and capacity of the organization to implement.

Organizations should look for *project components* supported by evidence that was built in a similar setting and/or population to the ones being served.[iii] When evidence supporting a *project component* is aligned to an organization's setting (e.g., elementary school), capacities, or learner population (e.g., students with disabilities, English learners, students experiencing homelessness, migratory students), it may increase the likelihood that prior positive outcomes will be replicated. In practice, finding evidence that was built in a context exactly like the one in which it is going to be applied can be difficult. (This is one reason why the Department encourages evidence building whenever possible: it increases the amount of evidence that helps educators understand what "works," for whom, and under what conditions.) *Project components* supported by evidence for a specific setting and/or population that are not the same population and/or setting of an organization might meet a lower evidence level for the organization's setting and/or population. The levels of evidence are further discussed later in this document. When relevant evidence is limited, organizations may seek to innovate by leveraging *project components* that are based on positive findings from high-quality research and evaluation activities conducted in contexts, settings, or populations that differ from their own. In those cases, those *project components*—when paired with a *logic model*[iv] that demonstrates how the *project component* is meant to improve an important learner outcome and with a plan to build high-quality evidence—might meet a less rigorous tier of evidence. The evidence tiers are discussed in Part II.

Organizational capacity to effectively implement a *project component* is also important to consider. Prior to selecting a *project component*, organizations should consider whether they have the resources needed to effectively implement the *project component*. This includes available funding, staff resources, staff skills, and support for sustaining a *project component*, including implementation partners, as needed. Organizations can work with technical assistance providers, including those supported by the Department, to improve their capacity to implement evidence-based *project components*. A list of some Department-funded providers can be found at https://osepideasthatwork.org/find-center-or-grant/find-a-center.

Some questions to consider when selecting evidence-based *project components* include:

- What is the highest level of evidence available for the issue seeking to be addressed?
- What do most studies—particularly those meeting higher tiers of evidence (tiers of evidence are discussed in Part II)—on this *project component* find? Does the *project component* have positive and statistically significant effects on important learner or other *relevant outcomes*,[v] or are there null, negative, or not statistically significant findings?

- Were studies conducted in settings and with populations relevant to the local context (e.g., students with disabilities, English learners, students experiencing homelessness, migratory students)?

- Is the *project component* based on a *logic model* and supported by research?

- How can outcomes associated with the use of the *project component* be measured?

- How does this *project component* align to feedback from stakeholders on identified local needs, their root causes, and potential ways to address them?

Some questions to consider about capacity:

- What resources are required to implement this *project component*?

- Will the potential outcomes associated with this *project component* justify the costs, or are there more cost-effective *project components* that will accomplish the same outcomes?

- What is the capacity to implement this *project component* or larger intervention? Are there available funds? Do staff have the needed skills? Is there buy-in for the intervention? Is there an implementation partner, as needed, to help?

- How does this *project component* or larger intervention fit into the organization's larger strategic goals and other existing efforts?

- How will this *project component* or larger intervention be sustained over time?

Step 3. PLAN FOR IMPLEMENTATION

A thorough and thoughtfully designed implementation plan, developed with input from stakeholders, can increase the likelihood that organizations can successfully deploy *project components* (and larger interventions) as planned.

Implementation plans should include the following components:

- A *logic model*, informed by high-quality evidence, that suggests how the *project component* is likely to improve *relevant outcomes*;
- SMART (Specific, Measurable, Achievable, Relevant, and Time-Bound) goals;
- Clearly outlined roles and responsibilities for people and organizations involved, including those implementing the *project component*, those with a deep understanding of the *project component*, and those ultimately responsible for its success;
- An examination of the organization's capacity to build evidence about the *project component's* performance or effectiveness that can inform decision-making and next steps, including, if support is need, what options are in, near, or available to an organization's local community to assist in building evidence.
- Implementation timelines for successful execution;
- Resources required to support the *project component*; and
- Strategies to monitor performance and ensure continuous improvement, including plans for collecting and analyzing data about a *project component's* implementation and student outcomes (see Step 5 in this guidance). These plans may include strategies to gather ongoing feedback from participants and stakeholders on the implementation of the *project component.*

Step 4. IMPLEMENT

Implementation will impact the ultimate success of a *project component*, so organizations should have ways to collect information about how the implementation is working and make necessary changes along the way.

Some questions to consider:

- What information will be collected to ensure the quality of implementation? Is additional information needed to understand how the implementation is working?
- Is the implementation plan being followed, including partner roles in implementation? If not, why not? Are changes to *project components* needed?
- Are more or fewer resources required than originally intended? Do resources need to be realigned or timelines adjusted?
- Are stakeholders being engaged in implementation and any revisions of implementation based on continuous improvement efforts?
- What are unforeseen barriers or potential biases in assumptions or inaccuracies to successful implementation?
- How is implementation working with other existing efforts?
- What does the collected information/data suggest about the success of the implementation, including implementation as planned and changes to strengthen implementation?
- Are changes needed to improve the implementation?
- If applicable, is the *project component* or larger intervention ready to be scaled to more learners or educators, taking into consideration unmet demand, addressing barriers to scaling, the efficiency in scaling, and sustainability? How have personnel (including partners), financial resources, and management capacity been considered?

Step 5. EXAMINE AND REFLECT

Performance monitoring and evaluation are different—but complementary—approaches to building evidence about a *project component* and the project as a whole. The evidence built by both approaches are most valuable when shared with key stakeholders for decision-making.

Performance monitoring involves regularly collecting and analyzing data to track progress against targets and goals, including data on program outputs (e.g., students served) and, when available, about program outcomes (e.g., graduation rates). It can help identify whether key elements of a *logic model* are being implemented as planned, identify whether the *project component* is meeting interim goals and milestones, and suggest ways the *project component* could be changed for continuous improvement. Performance information can also provide insight into whether the expected outcomes are being achieved, but in most circumstances cannot demonstrate that it was the *project component* (or the project as a whole) that caused the observed outcomes.

In contrast, certain rigorous evaluations can demonstrate whether a *project component* (or the project as a whole) is causing an observed outcome. Well-designed evaluations build evidence of the effectiveness of a *project component*, answering questions about the impact of a specific *project component* on *relevant outcomes*.[vi]

Some questions to consider:

- What are the expectations of success and how can success be measured?
- What are interim progress and performance milestones that can be tracked?
- What have participants (i.e., learners and educators) shared about their experience and how *project components* have been implemented?
- What forms of evidence—performance monitoring, evaluation, or both—are needed for decision-making? Are evidence needs expected to change over time (e.g., performance monitoring for short-term decision-making versus rigorous evaluation for long-term investment in a specific *project component*)?
- What is the extent to which the methods of evaluation will produce evidence about the effectiveness of the *project component* on *relevant outcomes*?
- Based on information, should this *project component* continue as is, be modified, or be discontinued?
- How could knowledge about this *project component* be shared with others and incorporated into decision-making going forward?

Part II: Guidance on the Meaning of "Evidence-Based"

Prior evidence of a policy, program, or practice's effectiveness is a powerful tool to identify ways to address education problems and build knowledge on what works. Several federal education laws, including the ESEA, emphasize the use of evidence-based *project components*. Evidence-based *project components* are also encouraged—and sometimes required—in competitive grant programs administered by the Department. Section 8101(21)(A) of the ESEA defines an evidence-based *project component* as being supported by *strong evidence, moderate evidence, promising evidence,* or evidence that use of the component *demonstrates a rationale* (see text box below), and some ESEA programs encourage the use of evidence-based *project components* while others require it.

This guidance does not address the specific role of evidence in each program the Department administers and therefore should be used in conjunction with program-specific guidance.

WHAT IS AN "EVIDENCE-BASED" INTERVENTION?
(from section 8101(21)(A) of the ESEA)

"…the term 'evidence-based,' when used with respect to a State, local educational agency, or school activity, means an activity, strategy, or intervention that –

(i) demonstrates a statistically significant effect on improving student outcomes or other *relevant outcomes* based on –

(I) *strong evidence* from at least one well-designed and well-implemented *experimental study*;

(II) *moderate evidence* from at least one well-designed and well-implemented *quasi- experimental study*; or

(III) *promising evidence* from at least one well-designed and well-implemented correlational study with statistical controls for selection bias; or

(ii)(I) d*emonstrates a rationale* based on high-quality research findings or positive evaluation that such activity, strategy, or intervention is likely to improve student outcomes or other *relevant outcomes*; and

(II) includes ongoing efforts to examine the effects of such activity, strategy, or intervention."

Criteria for EDGAR Evidence Levels

To complement and enable the use of the definitions in the ESEA in any of the Department's competitive grant programs, including programs authorized under other program statutes like the Carl D. Perkins Career and Technical Education Act; the Higher Education Act, as amended; the Individuals with Disabilities Education Act; and the Rehabilitation Act, as amended by the Workforce Innovation and Opportunity Act, EDGAR (34 CFR 77.1) outlines the Department's evidence framework. It includes definitions for *evidence-based, strong evidence*, *moderate evidence*, *promising evidence*, and evidence that *demonstrates a rationale*, as well as other relevant definitions, based on the ESEA definitions. The additional detail in EDGAR is necessary for inclusion and use in the Department's competitive grant programs to be able to differentiate between different evidence levels among applicants when a limited amount of funding is available. The Department makes determinations about where and how to include evidence-based on an examination of the existing evidence base and how the Department can encourage evidence-building. This framework can serve as a helpful, optional, guide for the education community in using evidence in policy decisions.

The criteria below represent how the Department identifies evidence at each of the four levels in its discretionary grants, using EDGAR (34 CFR 77.1). These criteria are summarized in Table 1 on page 17.

Strong Evidence. Strong evidence means that there is evidence of the effectiveness of a key *project component* in improving a *relevant outcome* for a sample that overlaps with the populations and settings proposed to receive that component, based on a relevant finding from one of the following:

(i) A practice guide prepared by the WWC using version 2.1, 3.0, 4.0, or 4.1 of the WWC Handbooks reporting a "strong evidence base" for the corresponding practice guide recommendation;

(ii) An intervention report prepared by the WWC using version 2.1, 3.0, 4.0, or 4.1 of the WWC Handbooks reporting a "positive effect" on a *relevant outcome* based on a "medium to large" extent of evidence, with no reporting of a "negative effect" or "potentially negative effect" on a *relevant outcome*; or

(iii) A single *experimental study*[vii] reviewed and reported by the WWC using version 2.1, 3.0, 4.0, or 4.1 of the WWC Handbooks, or otherwise assessed by the Department using version 4.1 of the WWC Handbooks, as appropriate, and that–

a) Meets WWC standards without reservations;

b) Includes at least one statistically significant and positive (*i.e.*, favorable) effect on a *relevant outcome*;

c) Includes no overriding statistically significant and negative effects on *relevant outcomes* reported in the study or in a corresponding WWC intervention report prepared under version 2.1, 3.0, 4.0, or 4.1 of the WWC Handbooks; and

d) Is based on a sample from more than one site (*e.g.*, State, county, city, school district, or postsecondary campus) and includes at least 350 students or other individuals across sites. Multiple studies of the same *project component* that each meet requirements in paragraphs (iii)(A), (B), and (C) of this definition may together satisfy the requirement in this paragraph (iii)(D).

Moderate Evidence. Moderate evidence means that there is evidence of effectiveness of a key *project component* in improving a *relevant outcome* for a sample that overlaps with the populations or settings proposed to receive that component, based on a relevant finding from one of the following:

(i) A practice guide prepared by the WWC using version 2.1, 3.0, 4.0, or 4.1 of the WWC Handbooks reporting a "strong evidence base" or "moderate evidence base" for the corresponding practice guide recommendation;

(ii) An intervention report prepared by the WWC using version 2.1, 3.0, 4.0, or 4.1 of the WWC Handbooks reporting a "positive effect" or "potentially positive effect" on a *relevant outcome* based on a "medium to large" extent of evidence, with no reporting of a "negative effect" or "potentially negative effect" on a *relevant outcome*; or

(iii) A single *experimental study* or *quasi-experimental design study*[viii] reviewed and reported by the WWC using version 2.1, 3.0, 4.0, or 4.1 of the WWC Handbooks, or otherwise assessed by the Department using version 4.1 of the WWC Handbooks, as appropriate, and that–

a) Meets WWC standards with or without reservations;

b) Includes at least one statistically significant and positive (i.e., favorable) effect on a *relevant outcome*;

c) Includes no overriding statistically significant and negative effects on *relevant outcomes* reported in the study or in a corresponding WWC intervention report prepared under version 2.1, 3.0, 4.0, or 4.1 of the WWC Handbooks; and

d) Is based on a sample from more than one site (e.g., State, county, city, school district, or postsecondary campus) and includes at least 350 students or other individuals across sites. Multiple studies of the same *project component* that each meet requirements in paragraphs (iii)(A), (B), and (C) of this definition may together satisfy the requirement in this paragraph (iii)(D).

Promising Evidence. Promising evidence means that there is evidence of the effectiveness of a key *project component* in improving a *relevant outcome*, based on a relevant finding from one of the following:

(i) A practice guide prepared by WWC reporting a "strong evidence base" or "moderate evidence base" for the corresponding practice guide recommendation;

(ii) An intervention report prepared by the WWC reporting a "positive effect" or "potentially positive effect" on a *relevant outcome* with no reporting of a "negative effect" or "potentially negative effect" on a *relevant outcome*; or

(iii) A single study assessed by the Department, as appropriate, that–

a) Is an *experimental study*, a *quasi-experimental design study*, or a well-designed and well-

implemented correlational study with statistical controls for selection bias (e.g., a study using regression methods to account for differences between a treatment group and a comparison group); and

b) Includes at least one statistically significant and positive (i.e., favorable) effect on a *relevant outcome*.

Demonstrates a Rationale. Demonstrates a rationale means a key *project component* included in the project's *logic model* is informed by research or evaluation findings that suggest the *project component* is likely to improve *relevant outcomes*.

Note: Following the October 2020 update to the Department's evidence definitions for its discretionary grants, the WWC released version 5.0 Handbooks[ix] and updated its website to identify practice guide recommendations and findings in WWC intervention reports and reviews of individual studies as strong, moderate, or promise evidence. These characterizations by the WWC may be helpful for identifying evidence to inform education decisions.

Table 1.
How can I tell using EDGAR what level of evidence a particular project component meets?

Evidence Requirement	Level of Evidence			
	Strong Evidence	*Moderate Evidence*	*Promising Evidence*	*Demonstrates a Rationale*
Outcomes	At least one statistically significant and positive effect on a relevant outcome; no statistically significant and negative effects on a relevant outcome	At least one statistically significant and positive effect on a relevant outcome; no statistically significant and negative effects on a relevant outcome	At least one statistically significant and positive effect on a relevant outcome	Not Applicable
Study Design	Experimental study	Experimental study or quasi-experimental design study	Experimental study, quasi-experimental design study, or correlational study with statistical controls for selection bias	Logic model informed by research or evaluation findings
WWC Evidence Rating	Meets WWC without reservations	Meets WWC with or without reservations	Not Applicable	Not Applicable
Sample Size	A large sample (n = 350+) and a multi-site sample	A large sample (n = 350+) and a multi-site sample	Not Applicable	Not Applicable

[i] Project component, as defined in 34 CFR 77.1, means an activity, strategy, intervention, process, product, practice, or policy included in a project. Evidence may pertain to an individual project component or to a combination of project components (e.g., training teachers on instructional practices for English learners and follow-on coaching for these teachers).

[ii] In addition to the What Works Clearinghouse, evidence resources like the Department's Regional Educational Laboratories (RELs) and other federally funded technical assistance centers may provide summaries of the evidence on various interventions and guidance on how existing research aligns to the ESEA evidence levels discussed in Part II of this guidance.

[iii] See here for an implementation planning and monitoring tool.

[iv] Logic model, as defined in 34 CFR 77.1, (also referred to as a theory of action) means a framework that identifies key project components of the proposed project (i.e., the active "ingredients" that are hypothesized to be critical to achieving the relevant outcomes) and describes the theoretical and operational relationships among the key project

components and relevant outcomes. More information on logic models can be found at http://ies.ed.gov/pubsearch/pubsinfo.asp?pubid=REL2015057.

[v] Relevant outcome, as defined in 34 CFR 77.1, means the student outcome(s) or other outcome(s) the key project component is designed to improve, consistent with the specific goals of the program.

[vi] To ensure these evaluations of effectiveness produce credible results, organizations can leverage Department of Education technical assistance, including working with local RELs to plan, implement, and conduct evaluations and/or by using supporting resources like this free software to simplify analysis and reporting of evaluation results.

[vii] Experimental study, as defined in 34 CFR 77.1, means a study that is designed to compare outcomes between two groups of individuals (such as students) that are otherwise equivalent except for their assignment to either a treatment group receiving a project component or a control group that does not. Randomized controlled trials, regression discontinuity design studies, and single-case design studies are the specific types of experimental studies that, depending on their design and implementation (e.g., sample attrition in randomized controlled trials and regression discontinuity design studies), can meet What Works Clearinghouse (WWC) standards without reservations as described in the WWC Handbooks:

(i) A randomized controlled trial employs random assignment of, for example, students, teachers, classrooms, or schools to receive the project component being evaluated (the treatment group) or not to receive the project component (the control group).

(ii) A regression discontinuity design study assigns the project component being evaluated using a measured variable (e.g., assigning students reading below a cutoff score to tutoring or developmental education classes) and controls for that variable in the analysis of outcomes.

(iii) A single-case design study uses observations of a single case (e.g., a student eligible for a behavioral intervention) over time in the absence and presence of a controlled treatment manipulation to determine whether the outcome is systematically related to the treatment.

[viii] Quasi-experimental design study, as defined in 34 CFR 77.1, means a study using a design that attempts to approximate an experimental study by identifying a comparison group that is similar to the treatment group in important respects. This type of study, depending on design and implementation (e.g., establishment of baseline equivalence of the groups being compared), can meet WWC standards with reservations, but cannot meet WWC standards without reservations, as described in the WWC Handbooks.

[ix] The WWC Handbook can be found at https://ies.ed.gov/ncee/wwc/Handbooks.