

SEPTEMBER 23, 2026

Zachary Rogers
U.S. Department of Education
400 Maryland Ave., SW
Washington, DC 20202

Docket ID # ED-2026-OPEPD-2542

Dear Mr. Rogers,

Thank you for the opportunity to comment on the Department of Education's proposed revisions to the "Education Department General Administrative Regulations" (EDGAR) (ID # ED-2026-OPEPD-2542). Arnold Ventures is a philanthropy dedicated to tackling some of the most pressing challenges in the United States. We invest in research to better understand the root causes of problems and build evidence about what works to inform policy solutions, particularly through rigorous, causal evaluations. Additionally, we work to identify, evaluate, and scale exceptional social programs that produce important improvements in people's lives, including in education.

Our comments focus on the parts of the Notice of Proposed Rulemaking (NPRM) that would most affect the quality of the evidence used to select or prioritize grantees and the rigor of federally funded research. We appreciate the Department's efforts to update and simplify various evidence definitions in the regulations and to increase the timeliness and accessibility of evidence reviews. We are also concerned that some of the proposed changes to the Department's grantmaking procedures would create instability and undermine the credibility of federally funded research. In our comment below, we suggest improvements to the regulatory text to strengthen the evidence used in grant selection and prioritization and to ensure that federally funded education research is sufficiently rigorous and objective.

Section 75.210 on general selection criteria – § 75.210(c)(2)(ix)

- Feedback #1: Add one item to the list of factors used to determine the quality of the design of the proposed project: "(ix) The proposed project's commitment to the registration of planned research design and methods, including primary outcomes, on public registry prior to outcome data collection." Renumber subsequent items.
- Rationale: Advance sharing of pre-analysis research plans greatly enhances the credibility and transparency of research by documenting the intended methodology and analysis plans before the study begins and helps to prevent selective reporting and data-driven results upon research completion. We recommend this additional factor be included to help reviewers determine the quality of a proposed project's design and strengthen research integrity.

Section 75.226 on procedures used to give special consideration – § 75.226(d)

- Feedback #2: Add “with projects supported by strong, moderate, or promising evidence given greatest consideration, followed by projects that are not supported by strong evidence, moderate evidence, or promising evidence but do demonstrate a rationale” after item (d).
- Rationale: While we agree that it is critical to incentivize innovation when solving systemic problems in education, we believe applications supported by rigorous impact evidence, as demonstrated by strong, moderate, or promising evidence, should be prioritized over projects that solely demonstrate a rationale and do not have any evidence of impact. We suggest the Department adopt a tiered approach to the procedures used to give special considerations, where applications supported by strong, moderate, or promising evidence are given greatest consideration, followed by applications that do not have this prior evidence but do demonstrate a rationale.

Section 75.253 on continuation of a multiyear project after the first budget period

- Feedback #3: Withdraw the proposed revisions to section 75.253 that would expand the Secretary’s discretion to make continuation determinations and set continuation award amounts.
- Rationale: Rigorous research studies often unfold over multiple years and depend on predictable funding over their term. The proposed revisions would broaden the Secretary’s discretion to decline or reduce continuation awards for multiyear research projects for reasons unrelated to research progress or the grantee’s performance. Exercising that discretion could compromise the integrity of multiyear research projects and leave the field without reliable long-term findings on key questions. Stable continuation funding protects existing federal investments in education research and ensures that Department-funded studies produce credible and useful evidence.

Section 75.901 on the suspension and termination of discretionary grants – § 75.901(a)

- Feedback #4: Withdraw the proposed addition to section 75.901 that would permit the Department to terminate a discretionary grant “for convenience of the Secretary or a pass-through entity.”
- Rationale: Rigorous evaluations can take years to design, implement, and complete. Their ultimate value depends on researchers reporting transparently what they find, including when a program does not work. Authority to terminate an award “for convenience” would let the Department end a study midstream for reasons unrelated to the grantee’s performance or compliance, both of which the Department can already address under existing termination provisions. Even if the authority is rarely used, its existence changes incentives. When continued funding may depend on producing results a funder wants to see, researchers face pressure to frame findings favorably, emphasize positive outcomes, and soften null or negative results. This dynamic is well documented. In education, evaluations commissioned by program developers report substantially larger effects than independent evaluations. In medicine, industry-sponsored trials are more likely to report favorable results than independently funded trials. The consequences are overstated claims of effectiveness, distorted

guidance about "what works," and taxpayer dollars spent scaling programs that do not improve student outcomes. Preserving research independence protects the credibility of federally funded education research and helps the Department direct resources to programs proven to work.

Section 77.1 on the definition of “Evidence framework” – § 77.1(c)

- Feedback #5: Ensure sufficient guardrails are in place to maintain consistency across evidence reviews and the rigor of the Department’s evidence standards.
- Rationale: We appreciate the Department’s efforts to make evidence reviews more timely and accessible through the proposed definition of “evidence framework.” However, we are concerned that the definition could affect the quality and rigor of the evidence review process. Under the new definition, the NPRM notes that, “the burden of evidence reviews could, at the discretion of the Department, shift from Department-funded evidence reviewers to reviewers funded by third-parties (e.g., philanthropy, State educational agencies).” Because the proposed definition would condense the detailed What Works Clearinghouse (WWC) standards into five broad criteria, the same study could receive different evidence ratings depending on the reviewer. This inconsistency could create confusion and weaken the Department’s evidence standards. In the final rule, we urge the Department to establish clear guardrails to ensure that evidence reviews remain consistent, objective, and rigorous across reviewers.

Section 77.1 on the definition of “Experimental study” – § 77.1(c)

- Feedback #6: Further revise the definition of “experimental study” by removing “single case design studies,” adding “lottery-based designs,” and including definitions of “randomized controlled trial” and “lottery-based design,” so the definition reads:

Experimental study means a study that is designed to compare outcomes between two groups (such as students) that are otherwise equivalent except for their assignment to a treatment group receiving a project component as compared with a control group that does not. Experimental studies can support claims of strong evidence. Randomized controlled trials and lottery-based designs ~~and single case design studies~~ are specific types of experimental studies that meet this definition:

(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 lottery-based design employs a random lottery to assign a group, such as students, teachers, classrooms, or schools, to receive the project component being evaluated (the treatment group).

- Rationale: We appreciate the proposed simplification to the definition of “experimental study.” In the final rule, we recommend refining the list of example methodologies to maintain an emphasis on the most rigorous available designs. Lottery-based designs are a well-established form of random assignment and are frequently used in education research when program demand exceeds available

capacity. Because they evaluate impacts across larger populations of participants and mirror real-world program allocation decisions, lottery-based designs often generate findings that are more useful for informing education policy and program funding decisions. By contrast, single-case design studies typically focus on a small number of individuals and are less commonly used to assess the effectiveness of education interventions at the program or system level. Accordingly, we urge the Department to limit the definition's illustrative examples to the methodologies most widely used to generate rigorous, causal evidence about program effects at scale.

Section 77.1 on the definitions of “Strong evidence,” “Moderate evidence,” and “Promising evidence” – § 77.1(c)

- Feedback #7: Further revise the definitions of “strong evidence,” “moderate evidence,” and “promising evidence” to retain key safeguards that ensure rigor in the application of each evidence standard, including requirements for minimum sample size and adherence to established design standards.
- Rationale: We appreciate the proposed simplification to the definitions of “strong evidence,” “moderate evidence,” and “promising evidence.” However, we are concerned that some of the requirements the Department proposes to eliminate from each definition will weaken the evidence standards. Specifically, for the definitions of “strong evidence” and “moderate evidence,” the NPRM proposes to eliminate the requirement that qualifying studies 1) include multiple sites and at least 350 total individuals, and 2) meet WWC design standards. Without these guardrails, lower-quality studies with limited sample sizes or methodological weaknesses could qualify for the same evidence designation as more rigorous evaluations. In the final rule, we urge the Department to maintain the requirement that qualifying studies are based on a sample from more than one site, include at least 350 total individuals, and meet WWC standards, particularly for attrition and randomization integrity.

Section 77.1 on the definition of “Quasi-experimental design study” – § 77.1(c)

- Feedback #8: Further revise the definition of “quasi-experimental design study” by removing “cross-sectional group designs,” and “growth curve designs,” and adding “natural experiments,” “regression discontinuity,” and “instrumental variable designs,” so that the definition reads:

Quasi-experimental design study 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. ~~Cross-sectional group designs, comparative interrupted time series, differences in difference, and growth curve designs~~ Natural experiments, such as regression discontinuity, comparative interrupted time series, difference-in-differences, and instrumental variable designs, are specific types of quasi-experimental studies that meet this definition. This type of study can meet the definition of moderate evidence but not strong evidence.

- Rationale: We appreciate the proposed simplification to the definition of “quasi-experimental design study.” In the final rule, we recommend refining the list of example methodologies that qualify as

quasi-experimental designs to focus on the most rigorous and widely accepted approaches for estimating causal effects when random assignment is not possible. The designs should rely on a natural experiment (e.g., an exogenously chosen threshold, a policy shift, etc.) that cannot be manipulated or selected into at the individual level, and include regression discontinuity, instrumental variables, difference-in-differences, and comparative interrupted time series.

Thank you again for the opportunity to share our comments. Please let us know how we might be helpful or answer any questions.

Sincerely,

Justin Milner

*Executive Vice President of Evidence and
Evaluation
Arnold Ventures
jmilner@arnoldventures.org*

Kelly McManus

*Executive Vice President of Education
Arnold Ventures
kmcmanus@arnoldventures.org*

[Arnold Ventures](#) is a philanthropy that supports research to understand the root causes of America's most persistent and pressing problems, as well as evidence-based solutions to address them. By focusing on systemic change and bipartisan policy reforms, AV works to improve the lives of American families, strengthen communities, and promote economic opportunity.