

Guidance for Validation Grant Applicants Proposing Impact Evaluations

The following guidance is designed to provide clarity to prospective Ascendium grantees about the information we look for in applications for [validation grants](#) that include either quasi-experimental designs (QEDs) or randomized controlled trials (RCTs). **Our main goal in funding such impact evaluations is to grow the number of postsecondary education and workforce training reforms and innovations backed by strong causal evidence of effectiveness in improving outcomes that matter for learners from low-income backgrounds, students of color, rural students, first generation students, and incarcerated learners.**

To that end, we fund independent impact evaluations featuring QEDs and RCTs that:

- Target mature, well-defined reforms or innovations for which there is a compelling reason to believe they could meaningfully improve important outcomes.
- Use study methodologies that maximize the chances of producing causal evidence of impact on identified outcomes, and that document implementation quality and contextual factors affecting a reform's effectiveness.
- Consider equity at each step of the study's development, execution and dissemination.
- Have the potential to generate evidence that is relevant and useful to policymakers, practitioners, and learners.

Below we discuss in more detail key factors we consider when gauging how well a given proposal will achieve the above goals. Unless otherwise noted, these considerations apply to impact studies involving both QEDs and RCTs.

If you have any questions or comments about this guide, please feel free to contact jvalentine@ascendiumeducation.org.

Consideration #1: Is the reform or innovation (i.e. program, practice, policy) being evaluated well-defined?¹

In the early stages of a reform or innovation's development, the details of its implementation inevitably evolve. As a result, it can change substantially many times before reaching a relatively steady state. Conducting a QED or an RCT of a reform or innovation before its key elements are well-defined will often lead to an undesirable situation where a study produces a finding regarding a reform's effectiveness but cannot clearly tell us what the reform actually was. This leaves the field in the dark as to (i) how to replicate the reform (in the case it is found effective) or (ii) what to avoid doing in the future (in the case the reform is found ineffective or harmful). We, therefore, look for the applicant to clearly and specifically describe the key elements of the reform they seek to evaluate, including the specific problem(s) the reform is meant to address and how it will theoretically address them.

¹ Throughout this document, we use the terms 'reform' and 'innovation' interchangeably to refer to programs, practices, and policies intended to improve learner outcomes or other outcomes that are relevant to learner or worker success (e.g., faculty or staff outcomes, institutional outcomes).

Consideration #2: Has the reform or innovation been implemented as intended in the real world?

In some cases, a reform or innovation may be reasonably well-defined, but, for whatever reason, cannot be implemented as intended. For example, a university might develop a thoughtful and well-defined system for improving academic advising to incoming students from historically underrepresented groups. However, when the university launches the innovation, they find students almost never attend advising sessions. If a QED or an RCT were to be used to evaluate such a reform, the study would almost certainly find the program produces little-to-no effect on its desired outcomes, but there would be no way of distinguishing between two possible explanations for this lack of impact: (i) the advising program was so fundamentally flawed in its conception that it would not have worked no matter what; or (ii) it was a well-conceived and potentially effective reform that was simply poorly implemented. Rather than conduct a QED or RCT of such a reform, its developers should explore reasons for the lack of attendance at advising sessions (e.g., poor communication about sessions). Only once the reform or innovation is being implemented as intended, including reaching its intended groups, should its effectiveness be evaluated in a QED or RCT.

Consideration #3: Is there a compelling reason for doing the study – i.e., (i) the reform or innovation being evaluated might be effective and/or (ii) the proposed study could generate evidence that is relevant and useful to policymakers or practitioners.

In gauging this, we look for applicants to make a compelling case for why the reform being evaluated has the potential to positively impact important outcomes, like persistence and credential completion, and/or outcomes of particular value to a study's intended users. Applicants are encouraged to describe findings from prior evaluations of the reform to be studied, or other reforms like it, which suggest the proposed study could find positive effects. We are particularly interested in understanding the extent to which prior evidence suggests effectiveness of a reform or innovation for groups who have historically been underrepresented in higher education. We are also interested in understanding the extent to which the proposed study will extend prior studies and address any important gaps in the evidence base. Evidence can come from a variety of sources (e.g., prior randomized, quasi-experimental or pre-post evaluations of the targeted intervention or interventions with similar features); applicants should distinguish between prior studies that provide promising descriptive evidence of effectiveness versus prior studies that provide strong causal evidence. This information from prior studies helps us gauge whether a proposed study is likely to help advance our goal of growing the number of higher education and workforce training reforms and innovations backed by stronger evidence of effectiveness.

We also value studies that seek to evaluate a program, practice, or policy in widespread use, on which little is known about its effectiveness, given such studies' potential to substantially influence higher education policy.

Consideration #4: How will the research team ensure the findings are accessible to and used by policymakers, practitioners, and/or learners?

Fundamentally, we want to understand the value proposition of any evaluation request. Who may benefit from the evaluation, and how? How can research teams ensure that research questions and ultimate findings are relevant, accessible, and useful to policymakers or practitioners? In validation grant applications, we look for applicants to discuss specifically which stakeholders will benefit from and use the research, and what engagement and dissemination strategies will be used to reach these stakeholders (i.e. a dissemination plan). Importantly, we believe that locally focused projects can in some cases provide insights that are applicable far beyond these projects' geographic boundaries. To maximize the reach and impact of the studies we fund, we want to make sure their findings influence not just the research community, but also practitioners, policymakers, and learners in higher education. To that end, in all communication products (e.g., reports, briefs, presentations) we encourage

grantees to communicate their study methodology and findings in a manner that enables a non-research audience to readily understand (i) how the study was conducted, (ii) why its methodology was credible, (iii) the magnitude of any impacts observed, (iv) any conclusions or hypotheses as to why and how the reform being evaluated did or didn't work, and (v) an argument for why the conclusions from the study likely apply (or not) beyond the study's geographic or contextual boundaries.

Consideration #5: How similar are the study's proposed treatment and comparison groups likely to be on both observable and unobservable characteristics; and thus, how confident can we be the study will provide valid causal estimates of the program's impact?

In reviewing both QED and RCT proposals, we want to be confident that the study's treatment and comparison group will be highly similar on observable characteristics (e.g., education history, demographic characteristics, geographic location), and, preferably, also on unobservable characteristics (e.g., motivation, psychological resilience). Having equivalent treatment and comparison groups – and/or methods to address differences between groups - is essential for these studies to rule out alternative explanations for observed differences in outcomes between the groups.

For QEDs, when gauging how likely it is that study groups will be highly similar on observable characteristics, we look for a discussion of what characteristics the groups will be matched on and how the research team will obtain the necessary data for matching. This discussion should provide a clear explanation for why the selected matching characteristics are likely to be important predictors of the study's main outcomes of interest. Importantly, in reviewing a proposal's matching strategy, we also look for evidence that the pool from which comparison group members are drawn is already reasonably similar to the treatment group on key observable characteristics before any statistical methods, such as propensity score matching or regression adjustment, are used to equate the groups. This is meant to avoid asking statistical methods to do more than they can reasonably be expected to do (i.e., equate two fundamentally different groups of people or institutions).

Observable characteristics like those described above often serve as valuable predictors of outcomes, but we also cannot ignore the influence of unobservable characteristics on outcomes. Controlling for such unobservable characteristics is often challenging in QEDs since, in many such studies, treatment groups volunteer for (or self-select into) a program, practice, or policy being evaluated, whereas comparison group members do not, potentially introducing an inherent difference in motivation between the groups. Given this common dynamic, we look for QED proposals with designs that (i) minimize self-selection as much as possible, (ii) at minimum, provide a detailed discussion of the extent to which self-selection is an issue and how it will influence interpretation of any observed difference in outcomes between treatment and comparison groups, and (iii) ideally, explore ways to establish, as best as possible, equivalence on unobservable characteristics (e.g., by conducting surveys that seek to measure baseline levels of resilience, motivation, or persistence for both treatment and comparison group members) or include methods to address unobserved differences between groups (e.g., difference-in-difference).

For RCTs, we look for a description of (i) when random assignment will happen in the sample recruitment process, (ii) who will carry out random assignment and how, (iii) how sample attrition will be minimized, and (iv) procedures to monitor random assignment to ensure that it is conducted objectively and consistently for all sample members. Randomly assigning a sufficient number of individuals or clusters of individuals to treatment and control groups should result in two groups that are highly similar on both observable and unobservable characteristics.

Consideration #6: Are the proposed study's research questions and outcomes well-articulated, and the evaluation design and analytical approach well-equipped to answer them?

A quality QED or RCT proposal should clearly state the research questions and define the study's primary outcomes of interest, including how the necessary data will be accessed to measure those outcome(s), and over what length of time they will be measured. In evaluating proposals, we hope to see concrete and specific descriptions of the study's primary outcomes. For example, if you are most interested in persistence, does that mean continuous enrollment over two years, three years, or some other length of time? If you are most interested in degree completion, does that mean two-year degrees, four-year degrees, industry-recognized certifications, or all of the above?

Applicants should describe why the research design they have chosen is the most rigorous one possible for measuring impacts on primary outcomes. We ask that applicants provide a thorough discussion of the study's targeted sample size, why it is realistic that the study will hit that target, and what is the minimum effect the targeted sample will enable the study to detect on its primary outcome(s) (as demonstrated by a power analysis). Importantly, applicants should provide a clear justification, based on existing higher education or workforce training research, for why their evaluated reform or innovation could realistically produce the study's minimum detectable effect.

Consideration #7: How will the study use mixed methods to ensure we understand why, how and under what conditions the estimated effects occurred?

We ideally want to find out not just "what works," but also "for whom does it work (and not work)," "why," "how," "under what conditions" and "at what cost". Therefore, we prioritize mixed methods evaluations which include an impact study (i.e., a QED or RCT) paired with implementation studies that leverage quantitative and qualitative methods (e.g. observations of programs, in-depth interviews or focus groups with students and providers, etc.). For proposed implementation studies, we'll want to understand how such studies will add to lessons learned or address questions raised by prior field testing and evaluation of the reform. We value (i) description of reform implementation including assessment of service contrast between the treatment and control groups, (ii) inclusion of participant voice in understanding implementation successes and challenges as well as the value of reforms, and (iii) measures of implementation quality, quantity, and variation (e.g., average number of reform elements received, percentage of participants receiving the full reform). As warranted, we also support cost benefit or cost effectiveness studies.

Consideration #8: What steps will the research team take to pre-specify its study design?

QEDs and RCTs typically require researchers to make many different choices when deciding a study's parameters. Who will the treatment and comparison groups be? How long will outcome data be tracked before and after the evaluated reform or innovation is implemented? What baseline characteristics will be measured to gauge the equivalence of treatment and comparison groups? What statistical model will be used to analyze the study's results and what variables will it control for? What outcomes are of primary interest?

We recognize that applicants may not know all of a study's key parameters at the time they apply, but we strongly encourage applicants to commit to pre-registering these parameters (e.g., with [OSF](#) and/or [REES](#)) before the study's launch, or at least before the research team views outcome data broken out by treatment and comparison groups. Doing so helps ensure that researchers do not unconsciously let outcome data influence them to make design or analytical choices that would lend themselves to producing a desired result.

We also encourage applicants to limit their pre-specified primary analyses to answering a small number of discrete research questions; this is meant to avoid a scenario where a study runs a large number of analyses on many outcomes and highlights the few statistically significant findings that inevitably emerge (even if it is likely they only appeared by chance). That said, we encourage studies that disaggregate impacts by key subgroups (e.g., race, gender, socioeconomic status), ideally anchored in an equity-informed theory of change. If a study proposes to pre-specify a subgroup analysis as one of its primary analyses, applicants should make a case for why the impact on that subgroup is particularly important (e.g., a discussion of why impacts might be different for that subgroup relative to others and/or why that subgroup is of particular relevance to the field).

A study also may measure additional, secondary or exploratory outcomes as long as they are clearly differentiated from the study's primary outcomes. We believe that such an approach increases the transparency and thus the credibility of both RCT and QED research.

Consideration #9: How will the study advance equity?

We are open to the many permutations of what advancing equity, along with related concepts like diversity, inclusion, and justice, means within your research project. We expect and encourage research teams to clearly define what advancing equity means within their work and include specific plans or examples of how they will center equity in their study. One approach—more tied to the concepts of diversity or representation—emphasizes broadening who does the research. Others may focus on ensuring results are disaggregated by important subgroups or examining whether a reform or innovation reduces race or class-based gaps in key outcomes. Others may call for a more fundamental rethinking of research practices, like who sets research questions and how the questions are determined. Below, we link to several resources on these topics. Some resources as you consider responses include:

- We All Count: <https://weallcount.com/>
- Chicago Beyond's Why Am I Always Being Researched?: <https://chicagobeyond.org/researchequity/>
- Democratizing Evidence In Education: <https://www.democratizingevidence4.us/>
- MDRC's Equity Resources: <https://www.mdrc.org/pub/equity-resources>
- Equitable Evaluation Initiative: <https://www.equitableeval.org>