

XPRIZE DIGITAL LEARNING CHALLENGE

Sponsored by the Institute of Education Sciences (IES)

RULES & REGULATIONS 2.0: DEMONSTRATION PHASE

September 2022

The Demonstration Phase (final phase) of the XPRIZE Digital Learning Challenge | IES is governed by these **Rules & Regulations**. **These Rules** provide expectations for this phase and how to execute against the requirements. Teams may continue to use the **Challenge Guidelines** originally published March 22, 2021, but should use these Rules as the final source of information for this round (in addition to clarifying questions to XPRIZE staff). Failure to adhere to these Rules may result in consequences as detailed in the Competitor Agreement.

Unanticipated issues, including restrictions to travel, may also necessitate modifications to these documents. XPRIZE reserves the right to revise these Rules and Regulations and all registered Teams will be notified of any revisions in a timely manner.

Please send any questions about this challenge to DigitalLearning@xprize.org.

NOTE: **Bolded** items are defined in Section 08: Glossary.

This version supersedes other versions.

1: OVERVIEW

The purpose of this release of the Rules and Regulations (“Rules”) is to provide the details and technical specifications required for Teams to provide submissions for the Demonstration Phase of the Digital Learning Challenge.

2: ELIGIBILITY

The only teams eligible to compete in the Demonstration Phase of the challenge and considered for the prize purse are the five (5) Finalists selected by the Judging Panel in August 2022. All requirements detailed in the following sections must be completed in order for a finalist to be eligible for the prize.

3: KEY PHASE DATES

The key dates of the Demonstration Phase are below (exact dates subject to change).

September 1, 2022	Demonstration Phase Opens
January 2023	Submission Instructions for Demo Phase Released
March 1, 2023	Demonstration Phase Deadline; Final Submission Due
Early-Mid March 2023	Demonstration Phase Judging
Mid-Late March 2023	Grand Prize Winners Announced

Note 1: All dates subject to change

Note 2: No virtual Team Summit will take place this round. Instead the XPRIZE staff will continue to host a series of bimonthly (twice a month) Office Hours to help teams with instructions around this Phase and answer any questions.

4: DEMONSTRATION PHASE REQUIREMENTS

Teams that have advanced to the Demonstration Phase (aka Finalists) will have approximately six months to conduct a new set of **Pilot and Replication Studies** in accredited educational institutions settings to further demonstrate the capabilities of their systems. The Finalists must conduct at least one experiment and at least five replications with at least three learner demographics along with the subsequent reporting and data by the Demonstration Phase submission deadline (March 1, 2023). Finalists are permitted to run concurrent studies (further info below).

Finalists will also be asked to update their previous information and details around their plans for further deployment. Such plans should include but are not limited to: a business plan showing how the winner will use its tool or method to generate revenue; an open sourcing strategy for the codes, algorithms, and models of the winning solution to be adopted; and deployment plans through partnerships or joint ventures with research institutions.

A full summary of the requirements/details for the DemonstrationPhase studies::

- Finalists must conduct one experiment (“Pilot” study) and at least five replication studies
 - These five replication studies must be with at least three learner demographics and each study must run a minimum of 30 days in length
- Studies must be conducted in a formal education setting. Formal education settings include public, private, or charter nonprofit or for-profit pre-kindergarten, primary schools, secondary schools, colleges, universities, vocational schools, adult education programs, and career and technical education programs. These education settings should reflect, to the extent possible, the diversity of students and characteristics of the American education system.
- The intervention used in the studies will not be evaluated in terms of actual learner outcomes. Rather, the study must demonstrate the ability of the platform to collect and analyze data that would indicate what learning outcomes were achieved and what factors led to the intervention’s effectiveness.
- The study sample size and allocation to condition should be such that the minimum true impact detectable size with 80 percent power and a 95 percent confidence interval is no larger than the minimum relevant size impact for policy or practice.
- The study design should be driven by a hypothesis that is based in established educational theory or practice that will be advanced through the findings of the study.
- Studies should be pre-registered using an open science platform (e.g. [Open Science Foundation](#))
- The study must collect at least the following data points, but is not limited to collecting just them. Teams may collect any other data they deem relevant:
 - Individual student identification (can be assigned and MUST be anonymized) with means to identify each student
 - Individual student demographic data and information on student characteristics that impact education outcomes
 - Student baseline measures
 - Student outcomes measures
 - The baseline and outcome measures must be measured using the same units
 - Student attrition
 - Process (e.g. clickstream) data describing how students interact with educational materials and activities

- Primary outcome measures should include student outcomes sensitive to the performance change the intervention is intended to bring about. Consistent with SEER principles, widely used “common measures” must be included in addition to any researcher developed measures used. Outcome measures should be pre-specified, have been demonstrated as reliable and valid for the intended purposes, and based on data-collection methods that have been shown to yield reliable data.
- Study must last no less than four weeks from the date of baseline data collection and outcome data collection. The study must be conducted within the timeframe of the phase, which begins September 1, 2022 and ends March 1, 2023.
 - Concurrent Studies are permitted as long as:
 - Replication studies cannot completely overlap with the pilot study
 - Finalists use data from the pilot/initial study in this phase to inform their replication studies
 - Finalists can illustrate/explain how their pilot/initial study and replication studies are distinct and different
- Study must have received IRB approval (if required).

4.1: SUBMISSION DETAILS

Finalists will submit their Demonstration Phase Submission online at pop.xprize.org by March 1, 2023 for the Judging Panel to review. The Demonstration Phase Submission will consist of:

- A full technical report of the experiments (a template will be provided by XPRIZE by early 2023)
- The raw data generated by the studies (with appropriate safeguards protecting privacy and in compliance with the team’s IRB requirements)
- Reports summarizing the data,
- A set analyses using the raw data

To be eligible for the final **Prize Purse**, finalists must submit a Demonstration Phase Submission that is complete and meets all the requirements laid out in Section 4 of this document.

XPRIZE plans on providing more detailed submission instructions including any applicable reporting templates in January 2023 (if not earlier).

5: JUDGING CRITERIA AND SELECTION OF WINNERS

Submissions will be reviewed by the Judging Panel to select a Grand Prize Winner and a Runner-Up Team for the challenge who will, respectively, be awarded a \$500,000USD Grand Prize and a \$250,000USD Runner-up Prize. The *winning **Solutions** will be those that meet the minimum requirements and are best able to demonstrate the robustness of their system to host a variety of experiments of education interventions. This might include randomly assigning students, teachers/classrooms/schools to groups, collecting relevant high quality data, and conducting reproducible analyses based on those data that demonstrate the capabilities of the system.* Ideally,

the winning team will be able to provide comprehensive measures, multi-dimensional representation of learner engagement, robustness of measures in relation to the constructs that are attempted to be measured, and will include contextual and granular data as well.

Teams will NOT be assessed on whether the interventions used in their experiments produce the desired impact, but rather will be evaluated by the ability of their systems to conduct experiments and *measure* learning processes and outcomes.

The Judging Panel will evaluate submissions for this phase based on the proposed judging criteria described in the table below. Judges will use ranked scoring.

Table 1. DIGITAL LEARNING CHALLENGE SCORING CRITERIA

Criteria	Rubric			
<p>Team conducted the study using rigorous research methods and techniques Sampling Condition Assignment</p>	<p>Pilot and replication study elements DO NOT employ random sampling and conditional assignment techniques.</p>		<p>Pilot and replication study elements reflects some elements of rigorous research methods and techniques.</p>	<p>Pilot and replication studies employ random sampling and conditional assignment.</p>
<p>Pilot and replication are designed such that they have distinct data collection processes Note: due to timing constraints the studies can overlap, but must be distinct.</p>	<p>Pilot and replication study are not staggered or distinct, and do not use pilot data to inform replication.</p>		<p>Pilot and replication study are staggered, distinct, but does not use pilot data to inform replication.</p>	<p>Pilot and replication study are staggered, distinct, and use pilot data to inform replication.</p>
<p>Studies are based on established educational theory or practice that will be advanced through the findings of the study.</p>	<p>The study lacks a valid hypothesis based in established educational theory or practice, or the hypothesis is flawed and in conflict with established learning sciences findings.</p>		<p>The team's hypothesis is based in established educational theory or practice. However, it is unclear whether the team and platform will modernize, accelerate, and improve the ways in which we identify effective learning tools and</p>	<p>The team's hypothesis is firmly based in established educational theory or practice. There is high confidence that the team and platform will modernize, accelerate, and improve the ways in which we identify effective learning tools and processes that improve learning outcomes.</p>

			processes that improve learning outcomes.	
Setting: Formal education settings include public, private, or charter nonprofit or for-profit pre-kindergarten, primary schools, secondary schools, colleges, universities, vocational schools, adult education programs, and career and technical education programs. These education settings should reflect, to the extent possible, the diversity of students and characteristics of the American education system.	The research setting IS NOT a formal education setting, and DOES NOT reflect , to the extent possible, the diversity of students and characteristics of the American education system.		The research setting is formal, however, the institution may possess unique characteristics, niche student populations that impact how instruction and learning occurs.	The research setting IS a Formal education setting, and, to the extent possible, reflects the diversity of students and characteristics of the American education system.
Compliance with competition criteria: The team's proposed study meets the basic specifications laid out in the Guidelines and Rules and Regulations (i.e., platform/technology, RCT/QED, minimum 30 days, at least one replication with one learner demographic).	There is nothing in the study report that shows RCT or QED study lasting a minimum 30 days, and at least one replication with one learner demographic		The study report and data don't fully show an RCT or QED study lasting minimum 30 days, and at least one replication with one learner demographic.	The study report fully addresses the requirements for conducting RCT or QED study lasting a minimum 30 days, and at least one replication with one learner demographic and confident that the team fully met the requirements.
Effect Size Parameters: The study sample size and allocation to condition should be such that the minimum true impact detectable size with 80 percent power and a 95 percent confidence interval is no larger than the minimum relevant size impact for policy or practice. What is the substantive relevant effect? Note: Samples must be large enough to be meaningful;	The study is not adequately powered to estimate the minimum detectable effect size.		The minimum true impact detectable size with 80 percent power and a 95 percent confidence interval is not consistent with (too small, negative effects) the minimum relevant size	The study sample size and allocation to condition should be such that the minimum true impact detectable size with 80 percent power and a 95 percent confidence interval is no larger than the minimum relevant size impact for policy or practice.

<p>Minimum effect size a standard of ERIC or What Works Clearinghouse</p>		<p>impact for policy or practice.</p>	
<p>Platform Performance: The study must collect at least the following data points, but is not limited to collecting just them.</p> <p>Individual student identification (can be assigned and MUST be anonymized) with means to identify each student; Individual student demographic data and information on student characteristics that impact education outcomes; Student baseline measures; Student outcomes measures; The baseline and outcome measures must be measured using the same units; Student attrition; Process (e.g. clickstream) data describing how students interact with educational materials and activities; Primary outcome measures should include student outcomes sensitive to the performance change the intervention is intended to bring about (and information about the power to detect effect).</p>	<p>The team DOES NOT demonstrate the platform's ability to collect and analyze any of the required data points.</p>	<p>The team demonstrated the platform's ability to collect and analyze SOME of the required data points.</p>	<p>Team demonstrated the platform's ability to collect and analyze all of the required data points.</p>
<p>Outcome measures: are pre-specified, have been demonstrated as reliable and valid for the intended purposes, and based on data-collection methods that have been shown to yield reliable data.</p>	<p>The study lacks pre-specified, valid, and reliable outcome measures.</p>	<p>The outcome measures are pre-specified, however, the data collection method has not been shown to yield reliable data .</p>	<p>outcomes measures are pre-specified, have been demonstrated as reliable and valid, team's data collection methods yield reliable data.</p>

<p>Data & Analysis: Presents data and analysis/analytic models on a specified topic in a clear, highly-organized, and effective manner.</p>	<p>Outputs in the report and data are not useful for drawing conclusions about the effectiveness of independent variables in the experiments.</p>	<p>Outputs in the report and data include the presentation of data from the tool are useful for drawing conclusions about the effectiveness of independent variables in the experiments.</p>	<p>Outputs in the report and data include the presentation of data from the tool are useful for drawing conclusions about the effectiveness of independent variables in the experiments. And the outputs include detection of the effect of intervention are presented in a way that is useful to practitioners.</p>
<p>Impact/Outcomes: The team demonstrated the platform's capacity to collect and analyze data that indicate what learning outcomes were achieved and what factors led to the intervention's effectiveness</p>	<p>The team did not demonstrate the platform's capacity to collect and analyze data that indicate what learning outcomes were achieved and what factors led to the intervention's effectiveness</p>	<p>The team demonstrated the platform's capacity to collect and analyze data, however, it is unclear the data points indicate whether the learning outcomes were achieved and what factors led to the intervention's effectiveness</p>	<p>The team demonstrated the platform's capacity to collect and analyze data that indicate what learning outcomes were achieved and what factors led to the intervention's effectiveness.</p>

<p>Innovation/Novelty of Technology in terms of collecting and analyzing data on learning outcomes.</p>	<p>The study DOES NOT demonstrate that the platform addresses collecting and analyzing data on learning outcomes.</p>	<p>The study demonstrates that the platform addresses collecting and analyzing data on learning outcomes, but does NOT introduce innovative or novel techniques or processes for instrumentation and does not address any new data types or analysis approaches.</p>	<p>The study demonstrates that the platform introduces significantly new techniques or processes for instrumentation of learner experiences, new data types, or new analysis approaches.</p>
<p>Open Science Registration: Causal impact studies must be pre-registered in a recognized study registry, documenting their confirmatory research questions and planned analytic activities. Researchers should execute research and analysis activities as proposed in their original study registration. When deviations from pre-registered plans occur, researchers must update their registry entries and provide an explanation for why a change took place.</p>	<p>The team DOES NOT pre-register their study, DOES NOT execute their planned research and analysis activities, nor provide an explanation for deviations (where applicable)</p>	<p>The team pre-registers their study, but it is missing important information, such as research and analytic plans, not fully developed.</p>	<p>The team pre-registers their study, executes their planned research and analysis activities, and provide an explanation for deviations (where applicable).</p>

The Judging Panel shall have sole and absolute discretion to declare the winners. Decisions of the Judging Panel are binding on XPRIZE, sponsors, and each Team Member. Per the Competitor Agreement for this Challenge, all parties will agree not to dispute any decision or ruling of the Judging Panel. While XPRIZE may decide to release scores and/or feedback, Finalists shall have no right to be informed of their calculations or other Teams' calculations, measurements, and results, unless such information is made available by XPRIZE (as stated in the Competitor Agreement).

If no Team meets the criteria for an award, then the Judging Panel will retain sole and absolute discretion to declare /not declare a winner of the Challenge and/or otherwise coordinate with XPRIZE to allocate/choose not to allocate one or more of the Awards and/or any other Award associated with the Challenge.

XPRIZE and the Judging Panel reserve the right to modify the scoring criteria as necessary before the submission date. Any such changes will be promptly communicated.

6: COMPLIANCE WITH LAWS AND REGULATIONS

6.1: LAWS AND REGULATIONS

All Teams and all Team members must adhere to all laws (including but not limited to) local, regional, national, and international laws, orders, directives, ordinances, treaties, rules, and regulations for all aspects of the Challenge. Teams are solely responsible for acquiring any appropriate licenses, waivers, or permits from the applicable regulatory bodies or other applicable third parties.

6.2: DATA STANDARDS AND OPEN DATA REQUIREMENTS

Competing Teams must demonstrate compliance with Federal and Department of Education open data requirements, which can be found at https://ies.ed.gov/funding/datasharing_implementation.asp. Teams must also demonstrate adherence to all necessary and relevant data privacy and confidentiality requirements including federal, state and local law in the locality where the pilots are done.

Competing Teams must use CEDS data standards and governance, as outlined in <https://ceds.ed.gov/dataModelEntities.aspx>. XPRIZE will create a centralized repository of data following CEDS standard models. The system will validate the accuracy of the data in the centralized repository. Teams will connect their solutions to this repository so that XPRIZE can collect and validate the data.

6.3: INSTITUTIONAL REVIEW BOARD (IRB) REQUIREMENTS

The XPRIZE Foundation is a non-profit within the United States and all Teams awarded funds by the XPRIZE Foundation and The Institute of Education Sciences (IES) must adhere to the ethics processes typical of research universities within the United States. Teams performing any experimentation and/or who collect data about people (e.g., having people interact with a chatbot) will likely require IRB review.

Competing Teams must obtain their own Institutional Review Board (IRB) approvals based on the Human Subject Regulations Decision Charts (HHS Tree), if necessary, for conducting human subjects research and submit them to the Judges. Teams will be required to provide details regarding any current and/or planned IRB review status. Teams who declare that they are IRB exempt must provide documentation to that effect for Judges' review. Systems must also comply with all other relevant regulations, such as COPPA.

- Teams will need to select one of the following options: The proposed research program is (1) Self-determined exempted from IRB review, (2) Formally exempted from ethical review by an IRB, (3) Pending ethical review, (4) Granted approval by an IRB, (5) Rejected by an IRB, (6) Pending establishment of international process equivalency, or (7) Currently exploring our legal and ethical responsibilities. Exempt Teams will still need to adhere to all relevant safeguards of privacy, confidentiality, and data protection.
- Option 1 requires submitting a written justification of your exemption qualification to XPRIZE Digital Learning Challenge staff (submitted via POP during a later phase of the Challenge) citing a specific exemption, such as detailing your [traversal of the HHS Tree](#). Options 2 through 4 above may be supported with IRB communications.
- XPRIZE and the Judging Panel will summarily reject any team inappropriately conducting research of and or providing ethically dubious results.

Please note: XPRIZE is not responsible for determining whether a team requires IRB review and it is the sole responsibility of each team to determine their status and requirements for obtaining IRB review. XPRIZE may provide additional educational materials on IRB requirements through workshops and trainings, but Teams should rely on IRB experts to ensure their compliance with all necessary regulations.

7: GLOSSARY

Below are glossary terms and additional definitions for the purposes of this Challenge:

Advisory Board: A select group of prominent advisors who contribute their wisdom, knowledge and guidance to various aspects of the prize.

Challenge Guidelines: Document for the public and for teams that describes the requirements and parameters of the challenge.

Competitor Agreement: A legal and binding document that details the responsibilities of competitors for the prize.

Demonstration Phase Submission: The required submission report for the Demonstration Phase - includes: a full technical report, raw data, data reports, and analyses on the data. A template for the technical report section will be provided in January 2023.

Digital Learning Systems: Digital learning systems (DLSs) are defined as any software that either organizes learning in both formal and informal settings or delivers content and pedagogical tools. DLSs can range from, but are not limited to, Learning Management Systems, online learning tools, curriculum products, school communication tools, data systems, educational operations software, and digital educational content, among others.

Experiments: An experiment is defined as either a randomized controlled trial or a quasi-experimental design that introduces an innovation or a new idea to a subset of a learner population with an intended goal in mind. Experiments must be instrumented to test substantive interventions and collect meaningful learner outcomes, not trivial ones. For example, being able to test the impact of the color of a button on the speed with which learners respond to an answer is a trivial intervention. Conversely, testing an increase in the number of repetitions of a mathematical concept to a subset of learners is substantive.

Experiment Infrastructure: Experiment Infrastructure is defined as an integrated set of features that enables product innovators and education researchers to implement interventions or innovations on a defined population subset and evaluate the outcomes using randomized controlled trials (RCT's) or quasi-experimental designs (QED's). Learner outcomes should include both behavior and learning gains. The best systems will enable a flexible and robust range of experiments, collect rich and insightful learning data, and require the least amount of effort from the researcher.

Judging Panel: The subject matter and technical experts who serve as an impartial and independent evaluation panel for all aspects of this prize. Judges score the team submissions and make the all award determinations throughout the challenge.

Pilot Study: Teams will have approximately six months to conduct Pilot Studies in formal education settings to demonstrate the capabilities of their systems and the Judging Panel will select up to five (5) Finalist Teams to advance to the Demonstration Phase of the challenge and split a Milestone Prize of \$250,000 (up to \$50,000 awarded to each team) based on evaluations of each team's Pilot Study submission.

Prize Operations Platform (POP): The standard XPRIZE portal for teams to input data, documents, and other information for use in the Digital Learning Challenge.

Prize Purse: Money offered, won, or received as a prize from competing in this challenge.

Rules & Regulations: Document detailing the testing protocols, specific rules, dates/times, and other details that will govern the challenge and will be binding on teams.

Solution: A team's specific submission (including all technical documentation and physical prototypes) that the Judging Panel will evaluate for this challenge.

Systematic Replications: Systematic Replications are defined as those that implement and evaluate the interventions in an original experiment in ways that systematically vary at least one aspect of the prior study, such as the geographical location; the population of learners, educators, and/or schools; and/or the intervention implementation. As Teams may use third-party learning content, we consider revisions to the intervention to include revisions to the systems' ability to conduct a study, not only revisions to the learning content. Considering the time limitations during later phases of the Challenge, competing Teams should also consider developing systems that can run multiple experiments simultaneously with different subgroups. More information about Systematic Replications can be found at https://ies.ed.gov/funding/pdf/2021_84305R.pdf.

TEAM DEFINITIONS

- **Pre-Registered Team:** A team or individual that is interested in participating in the competition and has created a profile in the XPRIZE POP system.
- **Registered Team:** A team that has provided a complete Team Questionnaire submission and has signed the Competitor Agreement will be deemed eligible to submit a Technical Submission for the Judging Panel's review.
- **Finalist Team:** Up to 5 Finalist Teams will be selected by the Judging Panel to proceed to the Demonstration Phase of the challenge based on the strength of their Technical Submission and Pilot Study Submission. Finalist Teams will split a prize purse of \$250,000 (up to \$50,000 per Team).
- **Runner Up Prize Winner:** The second place Team selected by the Judging Panel to receive the \$250,000 Runner Up Prize based on the strength of their submissions throughout the challenge.

- **Grand Prize Winner:** The team that has successfully demonstrated their solution's ability to meet and/or exceed the goals of this challenge and selected by the Judging Panel to receive the \$500,000 Grand Prize for this challenge.

Team Questionnaire: The Team Questionnaire is the initial submission where Teams will provide details about the current capabilities of their system and will be screened by XPRIZE and/or the Judging Panel for both the completeness of the proposals and for meeting minimum requirements outlined in the Rules & Regulations for this challenge.

Technical Submission: Teams will submit a free-to-use version of their systems as well as any technical documentation on the system (such as diagrams, drawings, schematics) and other written explanations of the functionality and architecture of the system in the form of the Technical Submission for this Challenge. All submissions will be screened by XPRIZE for completeness and the Judging Panel will then review the Technical Submissions and select Teams to advance to the Pilot Study phase of the challenge.

Randomized Controlled Trials (RCTs) and Quasi-Experimental Designs (QEDs): While Randomized Controlled Trials (RCTs) rely on random assignment to form intervention and comparison groups, Quasi-Experimental Designs (QEDs) form these groups using methods other than random assignment. Instead of randomly assigning subjects to intervention and control groups, they are split by some other means, with two groups formed through various, non-random processes such as using non-equivalent groups organized through non-random selection, relying on statistical methods to create a comparison group through matching, or relying on before and after time-series. More information on evaluation criteria for RCTs and QEDs can be found in the [What Works Clearinghouse Standards Handbook](#). Consistent with IES' goal to understand the generalizability of interventions, teams will be required to demonstrate their ability to deploy replications across at least three subgroups.

What Works Clearinghouse Standards and Standards for Excellence in Education Research (SEER): Competitors should use rigorous research designs that will meet What Works Clearinghouse standards with or without reservations, as well as IES-wide Standards for Excellence in Education Research (SEER).

The WWC standards can be accessed at

<https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf>.

IES has also laid out principles for conducting rigorous education research that is transparent, actionable, and focused on consequential outcomes, and which has the potential to dramatically improve student achievement. IES's SEER Principles encourage researchers to:

1. Pre-register studies
2. Make findings, methods, and data open
3. Identify interventions' core components
4. Document treatment implementation and contrast
5. Analyze interventions' costs
6. Focus on meaningful outcomes

7. Facilitate generalization of study findings
8. Support scaling of promising results

Competing Teams are highly encouraged to demonstrate their adherence to these principles. You can learn more about IES's SEER Principles at <https://ies.ed.gov/seer/> where you can explore in greater detail information about each of these principles.