Mathematica
Progress Together

Comment on Request for Information on

Potential New Program, "From Seedlings to Scale" (S2S)

RFI Number: ED-2023-IES-0182
November 13, 2023

Submitted to:

Institute of Education Sciences
Department of Education
400 Maryland Avenue SW
Washington, DC 20202–7240
Attention: Erin Higgins,
Education Research Analyst

Submitted by:

Mathematica
1100 First Street, NE, 12th Floor
Washington, DC 20002-4221
Phone: (202) 484-9220
Fax: (609) 228-4958

This page has been left blank for double-sided copying.

Comment on Request for Information on Potential New Program, "From Seedlings to Scale" (S2S)

The proposed new program from the Institute of Education Sciences (IES), From Seedlings to Scale (S2S), within its Accelerate, Transform, and Scale initiative, offers a critical opportunity to develop bold, innovative ideas to address some of the nation's most pressing educational challenges. The program will promote the development and scaling of breakthrough solutions for these challenges, including improving education outcomes for all learners; eliminating persistent achievement and attainment gaps; supporting the diverse needs of learners; introducing today's most technologically advanced tools into educators' toolboxes; and making all schools and classrooms into environments that uplift students' behavioral, emotional, and mental health.

Mathematica's response to this request for information (RFI) provides recommendations to strengthen the development of the S2S program. For more than 50 years, Mathematica has worked closely with federal agencies, grantees, foundations, state and local education agencies, community organizations, and other community members—such as educators, students, and families—to evaluate the effectiveness of innovations in education and support scaling for the most effective. Our response draws on our robust experience in educational innovations, rapid-cycle evaluation, and scaling highly effective programs to provide insights and recommendations for how to build a successful program designed to support the development and scaling of breakthrough educational solutions.

Response

(2) To successfully develop products and ecosystems that make a major impact on learners' education outcomes, teams will need a variety of supports. IES may require support from private industry in areas such as providing consultation and coaching to teams, convening potential partners for research and scaling.

(a) What would an ideal team look like to maximize the likelihood of success? For example, what role would researchers, education agencies (at the state or local level), and private companies play in the team?

We recommend that teams include the following roles to maximize the likelihood of success:

- **Solution developers** bring technical knowledge of solution design (that is, design of new products or services), including knowledge of feasibility constraints on design. This technical knowledge can encompass a broad range of domains, for example knowledge of specific technologies or different subject-based pedagogical expertise..
- **Researcher partners** ensure the use of high-quality study designs, measures, and data collection and analysis practices, as well as subject matter expertise in education, to generate evidence that informs iterative development throughout all phases of the work. If the research organization has the capability to effectively convene other team members to facilitate co-creation of learning plans, then it can also play a convening and facilitating role. The research organization might also be best positioned to facilitate communication between individual solution design teams and other teams within the same focus area, ensuring that information on challenges, successes, and findings from individual solution design teams continuously informs the broader effort for that focus area.

– **Educators, community partners, and state and local education agencies** ensure that the proposed solution aligns with a pressing need and identifies potential blind spots or inaccuracies in the assumptions underlying the solution's design. For example, if a solution requires students to use it proactively (as is the case for on-demand tutoring), educators and community partners might be well positioned to identify supports needed for students who are less engaged to benefit from the solution. Or they might recommend features needed for a professional learning service focused on classroom discourse to benefit students who are English learners. State and local education agencies are also indispensable partners in testing solutions at scale, where developers may encounter unique challenges arising from the diversity of implementation contexts that they need to address to ensure the solution's effectiveness at scale.

To further maximize success of the S2S program, we recommend it include an **overarching learning and evaluation partner**—a role that could be played by a research organization with the required capabilities. This overarching partner would convene the teams of evaluators; solution developers; and educator, community partner, and agency collaborators to provide ongoing guidance and foster learning within and across teams. Within each focus area, the overarching partner would guide the efforts of all distinct teams working within that focus area. This would include, for example, coordinating the use of consistent measures and processes across the teams in each focus area to raise the quality of evidence generated and support synthesis and sense-making across the teams.

Finally, to encourage participation from a broad range of qualified contributors, we recommend the S2S program use a funding mechanism that allows a broad range of organizations to engage in the work. Specifically, we recommend the program use a funding mechanism that allows a wide range of organizations that specialize in developing and evaluating innovative solutions to engage in the research and development. Specifically, funding that prohibits fees or requires cost sharing could exclude many types of commercial developers and research organizations.

(b) How can we ensure community engagement and input?

We recommend engaging community members as early as possible in the process of designing S2S, beginning with partnering with educators, students, and parents to provide input on IES' proposed focus areas to determine whether they address the most critical needs. We have developed a guide on meaningfully engaging community collaborators in research that could be useful when specifying how projects under this effort can ensure community engagement. Although it focuses mainly on research and evaluation, this guide describes best practices that would also apply to developing, refining, and scaling solutions.

Early engagement with community collaborators—including educators, parents, students, and community organizations—can yield the following benefits:

– **Empower educators and learners during the identification of promising focus areas and potentially promising solutions,** to increase the likelihood that the U.S. Department of Education (ED) investments target the problems the educators and learners see as most pressing.
– **Promote the design of solutions that meet specific learner needs,** including, for example, the needs of students with disabilities or those who require support on learning standards from prior grades.

– **Draw on the lived experience of the community collaborators** to identify key drivers and inhibitors of successful implementation.

(c) What kind of experience does your organization have with supporting ARPA-style R&D efforts, especially those related to the education sciences? What case studies can you share from your experience?

When considering how to support evidence-based development of innovative solutions in education, Mathematica draws on its own experience developing a new solution to support learning about education products' effectiveness. Specifically, Mathematica developed an innovative web-based platform, the Evidence to Insights Coach (e2i Coach), to put the tools of rigorous evaluation in the hands of education solution developers and their partners in state and local education agencies. Initiated with support from the U.S. Department of Education, this effort included the full suite of product design and user testing activities: needs sensing and the development of user profiles, creating an initial and refined prototype, and user testing of a prototype with dozens of district and developer partners. The e2i Coach is a user-friendly platform that guides a solution developer or district through designing and completing a rigorous comparison group evaluation. This includes describing the solution to be tested and the evaluation context; specifying research questions and thresholds defining meaningful impacts; data collection and cleaning; and analysis and reporting in metrics that support concrete decision-making on future use of the solution. After releasing the initial version of the Coach in 2017, we further refined it with support from the Chan-Zuckerberg Initiative to make it usable in applications beyond education.

Since its initial release, this tool has been publicly available without charge, and its use has scaled to over 450 users who have completed comparison group evaluations with it. Our experience iteratively developing this tool through user testing and piloting has provided invaluable first-hand experience on how to conduct evidence-based solution development centered on users' needs, iteratively refining the tool to improve product-market fit and usability.

Mathematica also draws on extensive experience supporting rigorous learning by education solution developers seeking to refine and scale products through cycles of learning and solution improvement. Over the past five years, Mathematica has partnered with the Bill & Melinda Gates Foundation and more than two dozen education solution developers to conduct research to inform development of new education solutions in middle school math and high school writing. This work has yielded several insights into how to structure phased investments in learning, starting with needs-sensing and continuing through user engagement to inform solution design, user testing, and evaluation of efficacy and effectiveness.

Mathematica developed a toolkit to guide the implementation of staged investments in generating evidence on breakthrough solutions that can provide insights to guide the development of the S2S program. Exhibit 1 presents the learning objectives of each of four phases in the staged learning approach we developed, accompanied by guidance on how to

Exhibit 1. Guidance on implementing a four-phase approach to developing education solutions

	Phase 1 Design the solution	Phase 2 Refine the solution	Phase 3 Assess for early evidence of success	Phase 4 Validate effectiveness
Primary objective by phase	Develop a solution with a well-defined theory of change that documents the links between solution components and desired outcomes.	Refine the solution based on lessons learned and generate evidence that the solution is successfully implemented with the community in focus.	Generate evidence that the solution is, at minimum, **associated with** improved outcomes for students in a single context (for example, one school or district).	Generate evidence that the solution **leads to** improved outcomes for students, across multiple contexts (at least two districts).
How to generate evidence to meet each phase's objectives				
Example research questions	How does the proposed solution address the problem previously defined with the community in focus? Who would benefit from the proposed solution? What would success look like in the short and long terms?	Are students completing modules as assigned? If not, what barriers do students report? To what extent do students in the communities in focus find the solution easy to use?	Do students who participate in the solution perform better in math than similar students who do not participate?	Do students who participate in the solution perform better in math than similar students who do not participate? Is the solution more effective in particular sites or for students who share certain characteristics?
Example study designs	Human-centered design and equitable community engagement	Qualitative user experience (for example, focus groups and interviews), surveys, and quantitative descriptive analysis	Randomized controlled trial (RCT) or quasi-experiment design (QED) preferred; correlational analysis (pre/post) with statistical controls also acceptable	Rigorous QED or RCT required
Sample size guidance	Not applicable	At least one school and at least five students or teachers from the community in focus	If assigning individual students to treatment and comparison groups: 100 to 200 students If using correlational analyses: 50 or more students	If assigning individual students to treatment and comparison groups: 350 to 500 students

pursue those objectives. Although the four phases in this exhibit do not align perfectly with the three phases described in this RFI, we believe the key objectives and guidance can help inform the sequence of goals and activities across the S2S sequence of evidence-based solution development phases. In the section that follows the exhibit, we provide more specific suggestions on activities and benchmarks aligned to the three phases listed in the RFI.

(3) With a focus on developing quick-turn around, high-reward and scalable solutions, what would you propose are the core activities and/or benchmarks for success for a project in each of the phases? What examples can you provide around past successes in social science domains or specifically related to education R&D?

Across all phases

Throughout all phases of S2S, we recommend that project teams consider the roles of transparency and collaboration in their process. Our experience conducting research and design informed by educators, students, parents, and school leaders has underscored the importance of sharing and co-interpreting findings with community participants to ensure they benefit from the learning accomplished. We have developed a guide on how to disseminate study results with a community focus so the knowledge developed in evaluations reaches the participants who made the research possible. To maximize the social benefit more broadly, we also recommend requiring that developers and other team members commit to sharing publicly the knowledge generated by evaluation activities in all three phases.

Before Phase 1

To align the work of all teams during the three phases of learning and development within each focus area, we recommend preparatory activities that an overarching, coordinating partner should facilitate with participating teams, including researchers; solution developers; and educator, community, and education agency collaborators, before engaging in Phase 1 activities.

- **Defining the evidence base** that motivates the investment in a specific focus area or category of solutions and building consensus around the basic attributes that define the solution category. Taking the first proposed focus area in the RFI, this would involve defining the set of broad competencies and domain-specific skills that align with growing areas for international economic competitiveness. It would also involve summarizing the research base on which types of existing interventions or intervention components exhibit evidence of effectiveness in boosting those skills and competencies.

- **Developing a theory of action** that reflects the state of existing evidence from the previous activity and documents evidence-based links between components of existing interventions and priority outcomes. Organizing the existing evidence in this way can clarify how to best position the individual teams to develop distinct and complementary domains of new evidence on specific aspects of the theory of action. For example, in the behavior and emotion regulation focus area, one solution team might focus particularly on strategies to promote positive peer interaction, and another team might focus most on strengthening student–teacher relationships. Specifying a theory of action at this stage can also support efforts to identify consistent measures of key outcome and implementation measurement domains across research teams within a specific focus area. This, in turn, will allow for clearer sense-making during Phases 2 and 3 by synthesizing findings from the different research teams. Clearer sense-making ultimately will accelerate the speed at which new evidence translates into improved solution design and will ensure that breakthroughs from one learning and design team can benefit other solution development teams.

Phase 1

We recommend Phase 1 focus on the following essential core activities: (1) engaging with educators, solution developers, and students and/or families to validate whether the idea responds to an actual need, (2) developing a solution-specific theory of change that grounds the selection of implementation and intermediate outcome measures reflecting distinct pathways to improve primary outcomes, and (3) documenting the base of extant evidence indicating that components of the proposed solution have promise.

We believe Phase 1 should *not* include pilots to assess early evidence of promise before users have tested the prototype. In our experience, there is a high risk of difficult-to-interpret findings if the team has not yet assessed whether users find the prototype usable and useful. This is because prototypes with low perceived usefulness or usability will often have low use, which in turn makes the measures of student outcomes uninformative. Gathering evidence on usability and usefulness—and iteratively designing improvements in response—is not feasible within the one-year time frame of Phase 1. Instead, we recommend beginning the testing for early evidence of promise in Phase 2.

Phase 2

In Phase 2, we recommend focusing benchmarks on usefulness, usability, and use. (Our guide on measuring implementation provides recommendations on how to measure these benchmarks.) This phase is the venue for human-centered design and user research, in which teams can determine what types of activities are feasible and make sense for which user population. Before establishing initial evidence of promise, teams need a period for user research and iteration. We recommend thinking of this phase in two stages. Stage 1 focuses on iterative development and Stage 2 focuses on a pilot to assess early evidence of effectiveness. Because these activities can take time, especially for teams just beginning to design their solution, we recommend some flexibility in the length of this phase. Two years might be sufficient if recruitment of participating schools, educators, and students in each stage occurs quickly, but this phase might otherwise take longer.

We have published a template and guide on developing measurement and evaluation plans comprising the learning goals for Phases 2 and 3. This tool provides recommendations on several aspects of conducting right-sized, meaningful learning during both of these phases. Topics include defining the solution being studied; specifying a theory of change; summarizing relevant prior research; documenting community engagement that has informed the solution design; and identifying the research questions, study design(s), measures, and timeline for the learning activities.

We recommend the following additional activities in Phase 2:

- **Measure cost per pupil to establish market fit.** If the cost will put the solution out of reach of many local education agencies (LEAs), the research team should identify strategies to make it more affordable.

- **Design communities of practice that support continuous sharing of findings across developers.** This can accelerate each developer's learning. This requires setting developers' expectations that knowledge created during the process will be a shared, rather than proprietary, asset.

- **Collaborate with community members when interpreting data and identifying findings.** This will also be important in Phase 3. Our guide to co-interpreting data with community collaborators provides recommendations on how to do this effectively.

Phase 3

We recommend that Phase 3 define benchmarks around the parameters of studies to conduct and the domains of measurement to include (as well as standards for identifying valid measures). Because the focus of Phase 3 activities includes a continued effort to learn about solutions as they scale, we highlight aspects of rigorous learning that Phase 3 should encompass.

- **Study parameters** include sample size and number of sites; minimum detectible effects; thresholds for meaningfully sized effects on outcomes; and representativeness of the study sample with respect to a defined population and intended users (which could be defined in terms of LEA, educator, and/or learner). We have published guidance documents on study sample sizes for generating different types of evidence and on setting targets—thresholds for meaningful study results—that could be useful references on identifying these study parameters.

- **Domains of measurement** should include measures of implementation fidelity, such as dosage offered, dosage received by learners, and implementation duration; primary outcome measures and proximal or mediating outcome measures; measures of cost that distinguish fixed costs at the district, school, or educator level from marginal costs per student or per unit of dosage intensity; and measures of LEA, educator, and learner characteristics that help assess for whom the solution is effective as it serves larger numbers of users. This last category can support assessment of product–market fit by identifying the users for whom the solution is more effective. Domains can also continue to include users' perceptions of the usability and usefulness of the solution, to illuminate conditions that enable or inhibit effectiveness.

This page has been left blank for double-sided copying.

This page has been left blank for double-sided copying.

Mathematica Inc.

Our employee-owners work nationwide and around the world.

Find us at **mathematica.org** and **edi-global.com**.