Navigating the Tensions: How Could Equity-relevant Research Also Be Agile, Open, and Scalable?

Jenna Zacamy, Empirical Education Inc.; Jeremy Roschelle, Digital Promise

August 2022

Digital Promise® Accelerating Innovation in Education
Empirical Education
IES Institute of Education Sciences

Table of Contents

Suggested Citation: Zacamy, J., & Roschelle, J. (2022, August). *Navigating the tensions: How could equity-relevant research also be agile, open, and scalable?* Digital Promise. https://doi.org/10.51388/20.500.12265/159

This project is supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305N210034 to Digital Promise. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

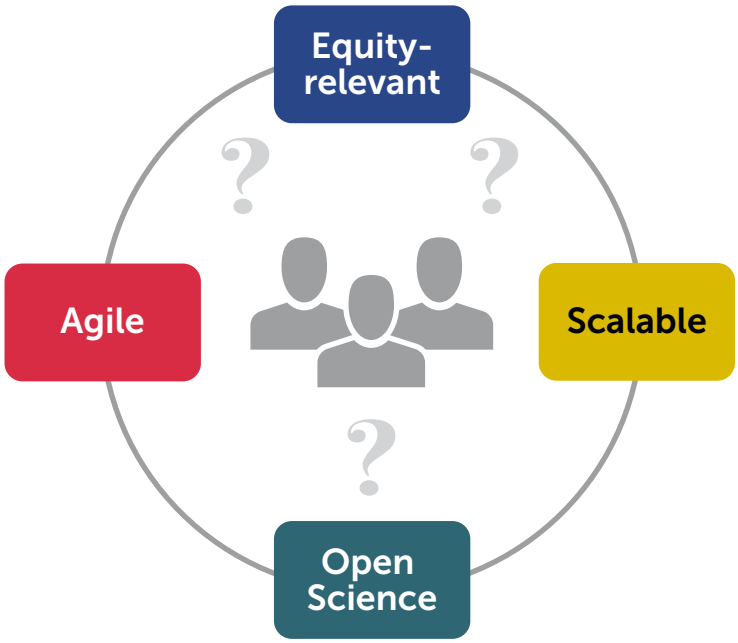
Abstract

Digital learning platforms (DLPs) are beginning to become open to research. Specifically, in our work in SEERNet, developers are extending five platforms, each used in either K-12 or higher education by more than 100,000 users, to enable third-party researchers to explore, develop, and test improvements. SEERNet seeks to enable equity-relevant research aligned with the Institute of Education Sciences (IES) Standards for Excellence in Education Research (SEER) principles. It also seeks to support research that is more agile (or rapid), is more open, and scales from research to impacts on practice. We review the emerging tensions among the goal of equity-relevant research and desires for agile, open, and scalable research. We argue that designing and developing technical capabilities for agile, open, and scalable research will not be enough. Based on a series of interviews we conducted with experts in social sciences and equity-focused research, we argue that researchers will have to rethink how they plan and undertake their research.

Five shifts could help. First, researchers could deliberately reframe their designs away from a comprehensive, monolithic study to smaller, agile cycles that test a smaller conjecture each time. Second, researchers could shift from designing new educational resources to determining how well-used resources could be elaborated and refined to address equity issues. Third, researchers could utilize variables that capture student experiences to investigate equity when they cannot obtain student demographic variables. Fourth, researchers could work in partnership with educators on equity problems that educators prioritize and want help in solving. Fifth, researchers could acknowledge that achieving equity is not only a technological or resource-design problem, but requires working at the classroom and systems levels too. In SEERNet, we look forward to working with the research community to find ways to address equity through research using well-used digital learning platforms, and to simultaneously conduct research that is more agile, more open, and more directly applicable at scale.

Introduction

Digital learning platforms—for example, widely used systems like Canvas—are beginning to become open to research that seeks improvements to teaching and learning. This may bring advantages. By starting with platforms that are already in use, research may be more connected to the specific challenges of teaching and learning in realistic settings. Also, working with platforms that are already widely used may reduce the time required to gain educational users, and thus, research may become more agile with faster cycles of evaluation and improvement and more opportunity to explore whether a promising approach works well among varied students and in varied settings. Regarding equity, it would be highly desirable to identify the barriers or obstacles that emerge for students in realistic settings. It would be desirable to more quickly determine how to support learner variability. Likewise, we need more research that seeks to understand how to activate different assets or skills that individual students bring to learning. With availability of a larger variety of students and educational institutions already using a platform, it should be easier to investigate for whom and under what conditions a new approach works.



Within the SEERNet network of nonprofit organizations, universities, and platform developers, we've begun exploring these possibilities. SEERNet currently consists of five digital learning platforms and a network hub (led by Digital Promise and Empirical Education) that interconnects the platforms with researchers. Each platform will be briefly profiled below. With support from the Institute for Education Sciences, each of these platforms is expanding its technical capabilities to enable researchers to conduct investigations using platform resources and data. Within about a year, IES expects to invite proposals from external researchers who will propose how they could use these platforms to conduct research. SEERNet will also welcome funded research projects into its network. (Additional information about SEERNet is in this blog.)

While exploring the future of SEERNet, we've encountered tensions between the desire to conduct **equity-relevant research** and other desired qualities of research. Regarding **open science**, the privacy, security and other safety issues around sharing data related to educational settings are already challenging. Data that could identify individual students are often involved in equity-relevant research and require strict precautions, which heightens the challenge to make datasets open. One challenge is that states and local education agencies differ in their policies around data use and data sharing. The differences in available data make it difficult to conduct cross-site and/or cross-state research that includes the student-level information that is often desired to conduct equity-relevant research. For example, some states require local educational agencies (LEAs) to opt in or opt out sharing their data with private vendors that deliver services under state contracts, and this requirement is often extended to researchers.

As another example, some LEAs have extensive research applications, while others require only an administrator to sign a data sharing agreement. While the FERPA "studies exception" allows schools or LEAs to enter a written agreement and disclose student personally identifiable information (PII) from education records to an organization, the study must be conducted "for, or on behalf of" the educational agency and researchers need to take precautions not to redisclose the PII to "unauthorized" individuals (Information about FERPA is available at https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html).

Open science is not just about making datasets open, but also about making other aspects of the research design, like interventions, methods, and measures, more open and transparent. Regarding agile research, researchers currently spend lots of time designing measures, gaining IRB and school approval to use them, and then collecting new data.

Digital learning platforms could save time if they are already instrumented to collect the data that researchers need; however, digital platforms often do not collect the data necessary to determine whether a research-based approach is solving an equity problem. Regarding scale, digital learning platforms do have large numbers of student and teacher users, but engaging with larger numbers increases risks, and thus, recruiting schools (or obtaining access through the platforms themselves) may not be fast or easy merely because an existing platform is used. Beyond the usual precautions like IRB approval, educational constituents do not necessarily want to open up everyday teaching and learning systems to allow research to be conducted about teachers and students—not unless the value back to teachers and students is very clear (as required by FERPA, see above).

With SEERNet partners, we can imagine a world in which equity-relevant research is more open, agile, and scalable. But we recognize that technical solutions (e.g. adding new research capabilities to existing digital learning platforms) will not be enough. We argue that researchers will have to rethink how they plan and undertake research.

Herein, we briefly review the equity-relevant features of improvement-oriented research that are possible within the five platforms in SEERNet, noting that the platforms vary significantly in regard to the equity-relevant features of research they support. We note that the tensions among equity-relevant research, open science, agile research, and the desire for scale can be heightened or reduced by choices made while designing the research. To illustrate this, we present a case study based on completed research to illustrate why merely shifting typical existing study designs into the SEERNet context may accentuate tensions. Based on a set of expert interviews we conducted with social scientists who address equity in their research, we argue that tensions may be lessened when investigators consider alternative ways to frame and organize their research, for example. We note that other fields are navigating similar tensions; for example, the American Psychological Association is also working to promote equity in open science. We suggest five shifts in research design, which are further described below:

1. From monolithic, comprehensive studies to smaller, agile cycles.
2. From develop-and-confirm to elaborate-and-refine.
3. From identity variables to experience variables.
4. From researcher-centered to educational-partner centered.
5. From technology- and student-focused to classroom- and system-focused.

Features of Equity-Relevant Research

The field does not presently have a shared definition of equity-relevant research. The field has advanced considerably from an early orientation to equality of opportunity and outcomes, and many candidate frameworks are available to guide equity-relevant research (Texas Center for Equity Promotion, 2021; Hawn Nelson, et al., 2020; W.K. Kellogg Foundation, 2022). SEERNet takes its name from the IES SEER Principles (U.S. Department of Education, Institute of Education Sciences, 2022), which recently expanded from eight principles to incorporate a ninth equity principle that states: Researchers who are designing and testing interventions must clearly demonstrate how those interventions address education inequities, such as by improving learners' outcomes and/or their access to resources and opportunities (U.S. Department of Education, Institute of Education Sciences, 2022). The ongoing discussion on how researchers conceptualize equity within their studies will be important to this community as SEERNet research will be funded by IES.

As this discussion continues to unfold, we focus on these two concepts: (1) Equity respects variability among learners and the corresponding need to acknowledge differences in individual learners' strengths, needs, experiences, and opportunities (2) Equity also respects that education occurs in systems, and the systems we have today work differentially well for groups of students that share common features of their identity. In one striking example, researchers have found that males perform better than females on multiple choice assessment items relative to constructed response assessment items; some state assessments appear to be biased toward males because they feature a greater proportion of multiple choice questions (Reardon et al, 2018). This also has implications for digital learning platforms, many of which can employ either multiple choice or constructed response questions: would it be important to monitor and adjust the usage of these different question types? Of course, differences between males and females are not the only differences that matter. Specific researchers may focus their efforts on racial inequities in light of unfair variability in access to resources and opportunity, cultural or linguistic inequities, and more. One conventional way of summarizing how equity-relevant research is different from other research is that equity-relevant research answers not only the question of "what works" or "what constitutes an improvement" but also "for whom and under what conditions."

In the context of SEERNet, researchers should be enabled to explore variability in students' learning opportunities and experiences in a digital learning platform, with sensitivity to variation also in settings and contexts. The exploration should enable an asset-oriented approach (e.g., González, Moll, & Amanti, 2006; Lee, 20201) to learners' differential strengths. Teams of researchers and innovators should be able to propose new approaches within a digital learning platform (or new variations on the content and interactive features already there) to improve teaching and learning. Researchers should be enabled to investigate or evaluate the promise of those innovations to address equity issues that have been identified and prioritized. SEERNet is beginning to curate specific examples of equity-focused research questions. Those examples are being co-designed with researchers, practitioners, and the DLPs. While that work is developing, we can begin by considering several broad research goals that reflect essential considerations for equity-centered research. Research goals could include:

- **Understanding usage:** Who uses the digital learning platform and how does use vary? What factors block or enable desired patterns of usage?

- **Developing conjectures:** Examining the qualities of learning experiences related to or occurring in the digital learning platforms and how those might be improved to leverage what we know about how people learn (National Research Council, 2000; National Academies of Sciences, Engineering and Medicine, 2018), best practices guides (e.g., What Works Clearinghouse Practice Guides), or insights from involving practitioners in co-design (Roschelle et al., 2006).
- **Adapting or personalizing:** Equity-relevant approaches might be adaptive (see Martin et al, 2020) to specific learner needs, strengths, experiences, or opportunities and may offer students or teachers more "voice and choice" or other features that personalize their experience (Zhang, Basham & Yang, 2020).
- **Investigating "for whom and under what conditions:"** Specific equity-relevant populations of schools, teachers, and students may be prioritized for initial investigations (e.g. targeted universalism, powell, Menendian, & Ake, 2019), and researchers may wish to be attentive to the kinds of variations we discussed above, potentially incorporating these in analytic models as moderator or mediators of student learning.
- **Evaluating generalizability, replicability, and scaling:** SEERNet includes platforms that address topics like mathematics, for which substantial research literatures already exist. Rather than pushing for additional new discovery, worthwhile research goals may extend to understanding how well existing designs or principles replicate, generalize, or scale. As with other social sciences, researchers have described a replication "crisis" in education (Williams, 2022). Generalization (Tipton & Olsen, 2018) and scaling up approaches that have been successful at a small scale are also important problems (McDonald et al, 2006).

In the spirit of working definitions, we briefly describe some of the options in what this might look like. We've based these working definitions on an examination of the kinds of approaches taken in the US Department of Education's (ED) National Center of Educational Statistics' data collections, such as the student questionnaire administered by the National Assessment of Educational Progress, which includes items on classroom experiences and educational supports, or ED school climate surveys (EDSCLS), which include items related to engagement, safety, and environment. Experienced researchers would easily be able to expand and improve upon this starting point, and the basics we describe here would be elaborated in different ways depending on which types of research objectives are chosen.

Research efforts might consider ways to describe different **contexts or settings**. This could include characteristics of school district, school or classroom, or characteristics of a location (e.g. rural or urban settings). Obviously, population characteristics like socioeconomic indicators are readily available in NCES datasets. In addition, we note that equity-relevant indicators are increasingly available in open datasets like the Stanford Education Data Archive (SEDA), which include information on differential growth in student test scores based on student identity, including patterns of closing or expanding achievement gaps over time or the Urban Institute's Education Data Explorer, which draws on data from national data sources on schools/districts/colleges and allows users to build datasets and explore trends related to characteristics, finance, student outcomes, and teachers and staff. Thus researchers could go beyond assumptions about how demographic characteristics correlate with learning to explore specific inequitable patterns in their chosen research locations.

Research efforts might also consider variability at the **classroom, teacher, or student** level. Just from examining what is collected in national data sets, we observe wide ranging possibilities. Obviously, these national datasets include demographic indicators, such as race and ethnicity. However, they also include data from surveys that ask about students' experiences. They include affective indicators, instructional differences (e.g. how much time is spent in small group learning), and questions that ask students about their contexts. Indeed, we've started by looking at extant open datasets, such as those available from NCES or SEDA, to emphasize a starting point where equity-relevant research and open science is less problematic: equity-relevant explorations can start with datasets that are already readily available. It is worthwhile to remember that "for whom and under what conditions" is not entirely about individual characteristics: conditions of education are often inequitable and how an approach plays out in different conditions can be equity-relevant research even when individual demographic characteristics cannot be collected easily.

In SEERNet, we expect this to expand to include **data about variable student behavior in the digital learning environment.** As we will describe shortly, the platforms vary in what data is available, but we expect that all will have data on students' prior performance in the digital learning platform—and this can allow exploration of how students with differences in recent performance use the platforms in similar or different ways. Further, researchers have found that patterns of usage sometimes correlate strongly with experiential variables, like whether a student is experiencing frustration (e.g., Henderson et al, 2019). Thus, it could become possible to ask exploratory questions about what digital learning features appear to further obstruct or enable learning on the platform for students who have previously shown either low performance or experiences of frustration. These are just some of the possibilities to use traces of student experiences in the platform as either direct variables or proxies for other variables of interest. With regard to proxies, the variable of interest may be supporting a specific student group (e.g. students with disabilities), however, it may be possible to initially study new approaches to helping that group of students by focusing on behaviors in the digital learning system that are exhibited for students with disabilities. Initially, one might be able to avoid asking students about their disability status, and thus, avoid the need to collect identifiable information about students.

There are also choices that SEERNet researchers will have about how and when to collect equity-relevant data needed to address their research objectives. Data might be collected from students, from teachers, from parents, or from a school district or state data system. It might be collected early in the research process (e.g. on a consent form), as students use a platform, or as part of a post-test or other measure of learning impacts. It could be merged in from other datasets available from the school or the state or in aforementioned open datasets. The ways in which researchers can collect context and setting data are still unfolding, as are the ways in which platform developers can incorporate the collection of this information into their systems. This will, at least initially, require collaboration between the researcher and the DLP, an example of which is described by Carnegie Learning in their blog post, From Design to Implementation, found at https://seernet.org/?p=408.

Overall, a key point of this section is that research teams that seek to conduct equity-relevant research will have a lot of choices and options. As this paper continues, this will become important in a couple of ways. First, we'll describe limitations in what the SEERNet platforms provide already; the consequence of this is that equity-relevant research with these platforms will likely involve making choices about how to collect additional data. Second, we'll observe that some choices that teams make could run counter to the spirit

of SEERNet as an easier, more agile alternative to traditional research processes. Third, we'll suggest some directions for moving forward with SEERNet that might reduce tensions among equity-relevant research, open science, and scale.

Variation in SEERNet platforms and resulting tensions

The five platforms in SEERNet provide different extant data collection capabilities for equity-relevant research. One platform, **Arizona State University's Learning At Scale** (ASU's L@S), will eventually provide investigators with the ability to explore very detailed information about individual students. As a university, ASU has significant background information about each student, including highly sensitive information about their family income, their demographic characteristics, their prior educational experiences, course profile and outcomes data, and more. Correspondingly, an open science approach will be slow and difficult to realize; we expect several years of preliminary policy development work before ASU can be open to third party research.

A different SEERNet platform, **ASSISTments / eTrials**, includes information about schools and districts, but almost no personal information about students. Indeed, the main data point available about an individual student is their prior performance on math tasks in ASSISTments itself. By having almost no PII available, ASSISTments has been able to take a strong stance on open science and is ready for researchers to use now. However, if researchers want data about student characteristics or experiences at the individual level, they need to determine provisions to collect that data and link it in themselves.

Researchers will also need to protect the resulting dataset before sharing it on an open science platform. We note one important additional point: ASSISTments has been studied with the BROMP protocol (Pardos et al., 2014), which established correlations between observable student experiences (e.g. a student expressing frustration) and clickstream data in the system. Thus, clever research teams could address equity while staying firmly within the open science commitments of the platform team if they focus on student experience variables, and don't initially focus on the lack of demographic variables. We'll return to and elaborate on this point later.

Figure 2: The design of eTrials prioritizes open science, but little individual student data is available. The design of ASU L@S prioritizes extensive individual student data, but open science will be harder to achieve. The other three platforms fall in the middle, leaving each researcher team to make choices about their priorities.

Kinetic at OpenStax aims to support researchers to test novel supports for learning that are positioned within their open textbook series. Kinetic will support Qualtrics surveys, and many students who use OpenStax are 18 years of age or older, which means they can consent to participating in research directly through a survey. Consequently, although Kinetic does not now collect information about how students

vary, the survey feature would allow individual investigators to do so. On the open science front, the corresponding burden would be on the researchers to protect PII when they publish any datasets they collect via such surveys. Kinetic's longer-term plan is to develop a secure data enclave which will allow researchers to run analyses with full datasets without access to PII. This approach will significantly reduce privacy risks while also making the breadth of identified data available to support research.

Terracotta / Canvas does not intrinsically provide data to researchers; rather the research team must form a partnership with an educational institution. In the partnership, the institution would agree to turn on the Terracotta plugin to their Canvas installation. Further, because a partnership is required and educational institutions have much data about individuals, they could agree to link such data with the data that research teams collect from Canvas using Terracotta. As with OpenStax, the burden would fall on the partnership between researchers and educators to protect data when sharing it, including complying with the applicable laws for protecting data of minors who are students.

UpGrade is an open-source A/B testing platform currently being used with Carnegie Learning's grades 6-12 math learning software, MATHia. UpGrade does not have a built-in way to connect student data to outside sources; however, if MATHia customers (that is, districts and schools) use a rostering service with unique student IDs, those IDs could be used to link to the district's demographic or assessment data (or other external data sources). One unique aspect of UpGrade's current application within MATHia is that it brings in an important scale consideration—they have a product that is being used by customers and want to be sure that what is being tested would be desired by customers if it worked (this is outlined in their study vetting process). Thus, the research plan must appeal not only to science, but also to what the market wants.

How research choices could accentuate tensions

SEERNet aims to support research that is equity-relevant, that moves toward open science, and that is agile but also scalable. Satisfaction of all these properties does not come for free, but is dependent on the choices a research team makes. Here we provide one case example, based on a historical study with one of the SEERNet platforms, ASSISTments.

A project called "An Efficacy Study of Online Mathematics Homework Support: An Evaluation of the ASSISTments Formative Assessment and Tutoring Platform" was funded by IES in 2012. The overall purpose of the study was to determine whether using ASSISTments to support homework for a full school year, in a state that provides each student with a laptop to take home, increases student math achievement compared to business-as-usual supports for homework. Results, which demonstrated positive impacts for ASSISTments, are described in two publications (Murphy et al., 2020; Roschelle et al., 2016). Although it could be considered a successful and useful efficacy trial, it was hardly agile—it took more than three years to complete—and was not particularly generalizable (indeed, subsequent projects were funded to generalize, replicate, and scale up the work). It also did not produce an open dataset, for reasons we describe shortly. This study can serve as an existence proof that one could design research with a SEERNet platform while achieving none of the goals of SEERNet. (The subsequent section will recommend alternatives).

As with many IES-funded efficacy trials, this study was conceived of as a large randomized controlled trial, powered at the classroom level. As is typically the case with large-scale RCTs, recruitment was challenging, resulting in two years of recruitment and two cohorts of data collection (in different school years). Hence the study was not agile.

For equity-relevant variables, the study team planned to use student-level demographic variables, including race, gender, ethnicity, and special education status. This study was conducted in Maine because the state already provided laptops for every student to take home, and thus, technology access would have little impact on student use of an online homework intervention. At the time, Maine was the only state with a statewide laptop program. The demographics of Maine include many rural and low-income students, but more than 85 percent of students are white.

The research team painstakingly negotiated access to the above demographic variables based on student IDs and data available to the University of Maine through its relationship to the state data system. A condition of access was that the data could not be released as an open dataset, hence it really was not an open science effort. Nonetheless, this was mostly all for naught. The analysis of these indicators revealed that they all correlated with each other and that prior mathematics achievement was as good a variable as any to look at differential impacts. These impacts were positive: students with lower prior mathematics achievement (or any of its correlates) gained more from ASSISTments implementation than did students with higher prior mathematics achievement (Murphy et al., 2020).

Further, it wasn't really necessary even to have the state-level dataset as a prior mathematics achievement covariate because within-ASSISTments measures have been established as pretty good predictors of the state measures (Feng, Heffernan & Koedinger, 2006). One could say the research team (including an author of this paper, Roschelle) painted themselves into an open science limitation for little reason. Incidentally, data analysis was also non-agile because only the University of Maine had access to covariates, and thus there were weekly cycles for each post-hoc exploratory analysis because the primary research team could not directly interact with the data and had to ask the team at the University of Maine to try the analysis and report back.

Relative to all the effort that went into collecting demographic variables (which were not that useful), in retrospect the study team wishes it had collected more teacher-level data. We observed a possible mediating effect of teacher homework review practices. Teachers who had ASSISTments appeared likely to amplify the benefits of the system if they used homework data in two ways: 1) to select specific homework problems to review and 2) to examine the most common wrong answer of their students and address possible misconceptions.

If the study team had conducted a preliminary year of agile teacher implementation research (which SEERNet could enable) instead of focusing so much on one big study, it might have noticed this possible mediating effect earlier and worked on ways to best measure it within ASSISTments. This might have resulted in additional nuance in our outcomes that would be quite useful to districts and teachers who seek to implement ASSISTments. (However, peer reviewers in the IES competition for efficacy proposals appear to heavily weigh any "incomplete" measures against a proposal's score, and thus, we were loath to propose that any work on measures would be needed in the course of the four-year project.) As it turned out, we did have some potential indicators of mediation, but our analysis failed to show an effect, which we believe may have been a measurement problem (Murphy et al., 2020).

What worked best about the study was really the scalability of ASSISTments, which played out well in recent years as use of digital learning platforms soared. ASSISTments usage has greatly increased, and this study's findings are often cited in promoting the platform. For example, an IES-sponsored review conducted during the COVID-19 pandemic recommended ASSISTments based on this study (Sahni et al., 2021).

Our overall point, however, is that the IES efficacy trial described above would be no more agile, open science, replicable, or generalizable with SEERNet than it was when proposed 10 years ago. It depended on covariates data from the state of Maine that the team was not permitted to release; it involved development and deployment of many supplementary measures that took time to collect and analyze; and the study was designed as one large monolithic and comprehensive effort that took years to complete. Further, there are some presumptions about fundable research and about the kinds of studies that address IES research goals, that unless addressed, may continue to turn SEERNet based studies away from goals of equity-relevant, agility, and openness. For example, peer reviewers appear to expect a single, comprehensive, bulletproof study and may be less inclined to give high scores to agile sequences of studies.

Resolving the Tensions within SEERNet

To seek progress on the complex tensions among equity-relevant research, open science, agile research, and scale, we discussed the challenges with a convenience sample of a dozen experts in the field—people with significant experience in conducting school-based improvement and social science research, often with technology.

In line with the preceding example, the experts noted that some typical ways of proposing and conducting research may not translate into SEERNet advantages. We discuss five directions for moving forward below. Please also note the National Academies of Sciences, Engineering, and Medicine (2022) report that resulted in recommendations on how IES might change how it funds research in the future. The recommendations in that report align with some of the recommendations from the experts we talked with.

Shift 1: From a monolithic study to agile cycles

IES funds many projects that have one major empirical data collection and analysis cycle; such a project is more monolithic, and it may take three years before major empirical findings are published. Further, the emphasis on a monolithic study tends to lead researchers to plan a very comprehensive data collection, which in turn makes it harder to fund the study, to gain IRB approval, and to recruit school participants. Monolithic studies are rarely agile; they proceed slowly. Further, many of the data collections are planned, not because of an especially strong conjecture or theory that must be confirmed or disconfirmed, but to cover the bases for the intended moderator or other analyses of a more exploratory nature. Conducting the exploratory research as agile cycles may lead to a more focused set of conjectures worth testing at scale and may simplify and sharpen a culminating data collection. One could wait to collect extensive data until there is a clear reason to believe that the more extensive data will reveal findings that could not be easily predicted from a smaller dataset (e.g. when intercorrelations among prior math achievement and demographic variables are strong, there may be less rationale for collecting personal data early in the agile cycles). Further, it's very hard to recruit for large, comprehensive studies. Smaller agile cycles may help in building trust with schools and participant communities and gain their buy-in to collecting particular equity-relevant variables down the road, after there's been some evidence that warrants collecting those variables.

Shift 2: From develop-and-confirm to elaborate-and-refine

Although this is rarely commented upon, the traditional IES research cycle assumes that what researchers do is develop interventions and then confirm those interventions work. Although this makes sense

research-wise, schools these days show somewhat less interest in interventions coming from outside their buildings, in part because they are flooded with commercial educational technological products. Further, confirmations that an intervention works are only one factor in whether it scales—schools consider many other factors besides research evidence in deciding what to implement widely. Are there alternatives to develop-and-confirm as the core research activities?

As we considered the SEERNet platforms and talked with experts, other perspectives came into view. When working toward equity-relevant research, researchers could spend more time elaborating what the equity-relevant challenge is, potentially through exploratory or hypothesis-generating research. Traditional educational research often assumes that the equity-relevant question is known. Proposals often start by talking about well-known achievement gaps (and thereby take on a deficit frame) and then skip to their conception of an intervention worth developing and testing. What is often missing is elaboration of the questions that fill the gap between the broadest impacts of inequities (e.g. long-term trends in test scores) and what happens in classrooms as they use digital learning platforms, e.g. the proximal phenomena of inequity. What is also left out is what schools are already using that might be close to a solution or part of a solution. Through the process of elaboration, researchers might identify phenomena that can be easily measured in the platforms and indicators of overcoming barriers in the platform—and thus discover ways to measure and address inequities at scale without collecting much data about individual students (e.g. without collecting much PII).

A further consideration is that with more attention to exploratory research, investigators might uncover successful equitable practices that occur infrequently in digital learning platforms but might spread more widely if systematically supported. There may be less focus on injecting something newly designed by a researcher into an education setting and more on researchers and practitioners collaborating to refine good things that happen already so they can be supported universally.

Shift 3: From identity variables to experience variables

To address questions of "for whom" an improvement or approach works, researchers typically collect "policy-relevant" demographic variables (e.g., race/ethnicity, socioeconomic status, gender). These variables are routinely collected and stored in school administrative data and are "available" at the student level at scale (with appropriate permissions/data sharing agreements in place). These variables, however, focus on student identity which are only a proxy for students' actual educational experiences.

The work within SEERNet could shift more directly toward assessing interactions between improvements and students' unique experiences, such as their sense of belonging in their school/class, their self-efficacy in a subject area, how they are affected by disruptions, or how they are supported by teachers. These types of data would require a survey or some form of primary data collection, but it is likely that they could be more easily collected within the digital learning platform. Other data collected within the platforms may also provide insights into instructional experiences (e.g., content being assigned and at what pace) or what types of improvements have supported prior successful learning experiences for students. Experience variables may also pose less risk of identifying individual students and may move us away from further stigmatizing students who belong to certain identity groups. We acknowledge the risks are still present and researchers will need to take precautions to protect sensitive data and critically review their work to ensure it is promoting beneficial impacts for all learners.

Shift 4: From research-centered to partnership-centered

Along with the above shift, it is well worth noting that experts recommend that SEERNet could be more valuable if schools or instructors were full partners in the research. Taking a researcher-practitioner partnership approach could more closely link a school's improvement goals to the purpose of the research, making a school a willing partner in sharing their data to better accomplish their goals. It could also help with site recruitment, as the reason to participate would be to shift from being science-centered to being school-centered. Furthermore, this partnership allows schools to have an active role in ensuring that what is being asked in the research is consistent with their students' experiences and access to learning, thus improving the validity of the research. Notably, the primary benefit of this type of partnership is that in addressing local questions, it is more likely that the research being produced will be used to inform improvement or change practice.

Investing in a partnership-centered approach takes time, focused attention, authentic interactions, and purposeful negotiations between researchers, practitioners, and the community. This may be challenging for any one research team to scale to multiple sites. However, establishing longer-term collaboration within a few sites would allow for iteratively building upon what is learned and deepening and broadening what is studied within those sites. As these partnerships disseminate what is learned locally, other partnerships can be formed to test if similar/same improvements within the platforms generalize to their context.

Shift 5: From technology- and student-centered to classroom- or system-centered

Finally, the experts weighed in strongly as seeing the teacher—and not just the technology—as the driver of change where equity is concerned. Within all SEERNet platforms, there could be a stronger emphasis on "does this improvement give teachers information and guidance they need to address equity?" and not only on "adapting to each student to achieve equity." Cohen, Raudenbush & Ball (2003) noted that the core of instructional improvement requires transforming relationships among teachers, students, and resources in an environment (see Figure 3).

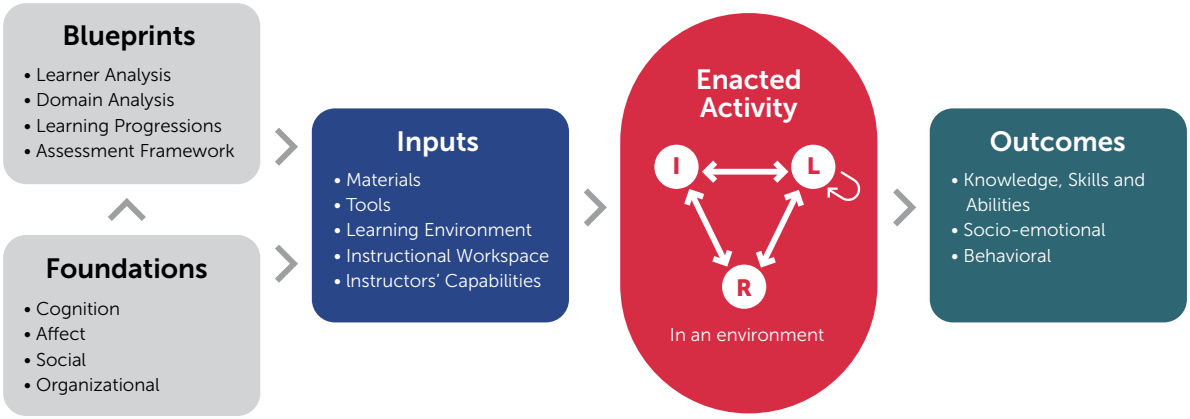


Figure 3: For learning to improve, the activities that instructors and students enact with curricular resources must change. SEERNet researchers should not only focus on the Learner (L) to Resource (R) interaction (i.e. student-to-technology) but also consider how Instructor (I) to Resource (R) and Instructorto Student (I to S) relationships need to change to achieve more equitable learning outcomes. From Foshay, W. R., & Roschelle, J. (2017). Learning Sciences and Instructional Design: Big Challenges, Multi-field, Multidisciplinary Solutions. In L. Lin & J. M. Spector (Eds.), The Sciences of Learning and Instructional Design: Constructive Articulation Between Communities. Taylor-Francis/Routledge.

Thus, the focus might also shift away from a sole focus on the student-technology relationship to incorporate the teacher and a systems view. The K-12 platforms are not intrinsically student-level interventions. In K-12, three SEERNet platforms, ASSISTments/eTrials, Mathia/UpGrade, and Terracotta (Canvas), are typically deployed to whole classrooms and used both during classroom time and (optionally) outside of classroom time. In higher ed, OpenStax provides textbooks, but students use textbooks to participate in a course. Likewise, the ASU focus is on course improvement, not platform improvement per se. The reason for considering this broader focus is that the equity-relevant research issues may arise more strongly in the broader classroom and systems context. Further, the malleable factors may include the role and instructional practices of the teacher. Research teams could explore how to improve the platforms to support a shift in classroom or systems equity practices that go beyond the boundaries of the platform itself. Finally, the nature of the data that need to be collected to answer equity-relevant questions may shift. For example, building on the prior section, it may suffice to look at how the system shifts to support engagement of students who express stronger and weaker "belonging" in the classroom process, and the only use of identity variables might be to establish population characteristics.

The research question: "In classrooms serving English learner populations and students who express high or low sense of belonging, does our novel approach increase student engagement in comprehension activities with assigned readings, do students express more confidence in participating in classroom discussions, and do their course grades increase?" The undescribed novel approach might be a student-level intervention, but also could be a way of providing feedback from low-belonging students to their teachers, so the teacher can better adapt readings, discussions, and interactions to their needs. The risk of open science would be decreased because the key variables in the dataset would be about student belonging, course behavior, and grades— which do not necessarily have to include demographic or identity variables. Further, the focus of intervention can be on helping the course instructor improve, not only to deliver something specifically different to individual students.

Conclusion

We've centered this discussion around four values that could have synergies but also can be in tension. Equity-relevance, open science, agile research, and scalable impacts are all desirable attributes of future efforts in our field. But we have argued that the existence of technical capabilities to use existing digital learning platforms for research will not be enough. Choices in research design and approach will strongly affect how well a particular research project satisfies these mutually desirable criteria.

We've emphasized that researchers have a wide array of choices they can make about how to design equity-relevant research. The SEERNet platforms can help researchers plan work that is equity-relevant, open science oriented, agile and scalable, but the platforms offer no guarantee of success. As a historical case study demonstrated, some choices that could be made with SEERNet platforms could result in research that achieves none of these values.

We recommend that researchers who are interested in planning SEERNet research investigate what the five platforms offer, but that they do not limit their research vision based on the technical possibilities of the platforms. Researchers will need to start by defining what equity means for their research goal. They will need to be creative in making many kinds of choices, including how to structure their work toward a broad research goal into agile cycles and how to align their research question to something that exists in a platform and can be elaborated and refined. They will need to consider whether striving for traditional demographic variables is the best choice in evaluating "for whom" an approach works, or whether they could more easily collect student experience variables to investigate inequities. Considering whether a partnership with educational institutions is appropriate is important too, both to build trust but also to make sure that outcomes would be something that educators want to scale. Finally, research teams may want to be careful about assuming they will address equity by changing a resource in a digital learning platform or by focusing primarily on individual students. They may wish to think about how students use digital learning platforms in classroom and school systems and what drivers for equity must be activated at these levels.

Acknowledgements

Thank you to our SEERNet advisors, to Jessica Tsang for consulting with us, and to those who participated in expert interviews.

This project is supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305N210034 to Digital Promise. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

Carnegie Learning/Upgrade (11 July 2022). From design to implementation. SEERNet. https://seernet.org/?p=408

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational evaluation and policy analysis, 25*(2), 119-142.

Feng, M., Heffernan, N.T., Koedinger, K.R. (2006). Predicting state test scores better with intelligent tutoring systems: Developing metrics to measure assistance required. In: *Ikeda, M., Ashley, K.D., Chan, TW. (eds) Intelligent Tutoring Systems.* ITS 2006. Lecture Notes in Computer Science, vol 4053. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11774303_4

González, N., Moll, L. C., & Amanti, C. (Eds.). (2006). F*unds of knowledge: Theorizing practices in households, communities, and classrooms.* Routledge.

Hawn Nelson, A., Jenkins, D., Zanti, S., Katz, M., Berkowitz, E., Burnett, T. C., & Culhane, D. (2020). *A toolkit for centering racial equity throughout data integration.* Actionable Intelligence for Social Policy, University of Pennsylvania. https://assets.aecf.org/m/resourcedoc/aisp-atoolkitforcenteringracialequity-2020.pdf

Henderson, N. L., Rowe, J. P., Mott, B. W., Brawner, K., Baker, R., & Lester, J. C. (2019, June). 4D affect detection: improving frustration detection in game-based learning with posture-based temporal data fusion. *In International Conference on Artificial Intelligence in Education* (pp. 144-156). Springer, Cham.

Lee, O. (2021). Asset-oriented framing of science and language learning with multilingual learners. *Journal of Research in Science Teaching,* 58(7), 1073-1079. http://doi.org/10.1002/tea.21694

Martin, F., Chen, Y., Moore, R.L. et al. (2020). Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018. *Education Tech Research Dev, 68,* 1903–1929. https://doi.org/10.1007/s11423-020-09793-2

McDonald, S. K., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher, 35*(3), 15-24.

Murphy, R., Roschelle, J., Feng, M & Mason, C.A (2020) Investigating efficacy, moderators and mediators for an online mathematics homework intervention. *Journal of Research on Educational Effectiveness, 13*(2), 235-270. DOI: 10.1080/19345747.2019.1710885

National Academies of Sciences, Engineering, and Medicine. (2022). *The future of education research at IES: Advancing an equity-oriented science.* The National Academies Press. https://doi.org/10.17226/26428

National Academies of Sciences, Engineering, and Medicine. (2018). *How people learn II: Learners, contexts, and cultures.* The National Academies Press. https://doi.org/10.17226/24783

National Research Council. (2000). *How people learn: Brain, mind, experience, and school: Expanded edition.* The National Academies Press. https://doi.org/10.17226/9853

Pardos, Z. A., Baker, R. S.J.d, Pedro, M. S., Gowda, S. M., and Gowda, S. M. (2014). Affective states and state tests: Investigating how affect and engagement during the school year predict end-of-year learning outcomes. *Journal of Learning Analytics 1.1.*, 107–128. DOI: 10.18608/jla.2014.11.6.

powell, J., Menendian, S., and Ake, W. (2019). *Targeted universalism: Policy & practice.* Haas Institute for a Fair and Inclusive Society. haasinstitute.berkeley.edu/targeteduniversalism

Reardon, S. F., Kalogrides, D., Fahle, E. M., Podolsky, A., & Zárate, R. C. (2018). The relationship between test item format and gender achievement gaps on math and ELA tests in fourth and eighth grades. *Educational Researcher, 47*(5), 284-294.

Roschelle, J., Feng, M., Murphy, R. & Mason, C.A. (2016). Online mathematics homework increases student achievement. *AERA Open, 2*(4), 1-12. DOI: 10.1177/2332858416673968

Roschelle, J., Penuel, W., & Shechtman, N. (2006). Co-design of innovations with teachers: *Definition and dynamics.* https://repository.isls.org//handle/1/3563

Sahni, S. D., Polanin, J. R., Zhang, Q., Michaelson, L. E., Caverly, S., Polese, M. L., & Yang, J. (2021). *A What Works Clearinghouse rapid evidence review of distance learning programs* (WWC 2021-005REV). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/Distance_Learning_RER_508c.pdf

Texas Center for Equity Promotion (2021). *Considerations for equity-centered research.* The University of Texas College of Education. https://texcep.education.utexas.edu/equity-centered-research

Tipton, E., & Olsen, R. B. (2018). A review of statistical methods for generalizing from evaluations of educational interventions. *Educational Researcher, 47*(8), 516–524. https://doi.org/10.3102/0013189X18781522

U.S. Department of Education. Institute of Education Sciences. (2022, June 30). Standards for Excellence in Education Research. https://ies.ed.gov/seer/

W.K. Kellogg Foundation. (2022). Doing Evaluation in Service of Racial Equity. https://everychildthrives.com/evaluation-in-service-of-racial-equity-can-unleash-power-voice-and-presence/

Wiliam, D. (2022). How should educational research respond to the replication "crisis" in the social sciences? Reflections on the papers in the Special Issue. *Educational Research and Evaluation, 27*:1-2, 208-214, DOI: 10.1080/13803611.2021.2022309

Zhang, L., Basham, J. D., & Yang, S. (2020). Understanding the implementation of personalized learning: A research synthesis. *Educational Research Review, 31*, 100339