Principles of Collaborative Education Research with Stakeholders: Toward Requirements for a New Research and Development Infrastructure

William R. Penuel[1]

Robbin Riedy[1]

Michael S. Barber[2]

Donald Peurach[3]

Whitney A. LeBouef[4]

Tiffany Clark[1]

[1]*University of Colorado Boulder*

[2]*Spencer Foundation*

[3]*University of Michigan*

[4]*University of Denver*

Abstract

A group of collaborative forms of education research sits uneasily within the existing infrastructure for research and development in the United States. Members of this group hold themselves to account to ways of working with schools, families, and communities different from the research models promoted in U.S. policies and endorsed by U.S. federal agencies. Those models privilege individual investigators' priorities for research and regularly yield products and findings with little relevance to practice. Four such models are reviewed in this paper: the Strategic Education Research Partnership, Design-Based Implementation Research, Improvement Science within Networked Improvement Communities, and Community-Based Design Research. Through a participatory process involving developers and advocates for these group members' approaches, we identified a set of interconnected principles related to collaboration, problem solving, and research. Further, we reviewed evidence for the embodiment of these principles in from four U.S. projects belonging to these approaches by examining a total of 13 journal articles, reports, and book chapters published between 2008 and 2018. Understanding, building, and supporting enactments of these principles is a worthwhile endeavor because there is evidence that these approaches to research can promote agency and equity in education. However, supporting these principles requires criteria for judging quality, which peers can use to evaluate individual studies or sets of research; new outcomes by which to measure progress; new venues for developing and giving accounts of research; and an appreciation for the value of developing and cultivating relationships with educators, families, and communities as an integral part of research.

Key words: design, collaboration, systems change, equity, research-practice partnerships

Principles of Collaborative Education Research with Stakeholders:

Toward Requirements for a New Research and Development Infrastructure

Policy makers and researchers have long lamented the limited impact of education research on policy and practice (Lagemann, 2002). With some notable exceptions (see Schneider, 2014), ideas and findings from research have had narrow and short-lived impacts on schools, districts, and communities. Since 2002, a central policy strategy in many countries for improving the value of research to educators has been to develop an infrastructure focused on the systematic development and testing of interventions designed to improve student outcomes (Wiseman, 2010). That infrastructure supports a "research-to-practice" pipeline, that is, a sequential model of innovation that begins with basic research on problem identification, proceeds to development and then testing of an innovation, and finally dissemination and promotion of broad use of effective innovations (Peurach & Glazer, 2012).

A core premise of this strategy is that to have impact, the *quality* of education research must first improve. By and large, definitions of what constitutes quality education research offered by consensus panels in the United States (National Research Council, 2002), federal agencies here in the United States (Institute of Education Sciences & National Science Foundation, 2013), and professional associations (American Education Research Association, 2009) have focused on the need for research to adhere to methods that are appropriate to answering the question at hand. Further, U.S. policies (e.g., Education Sciences Research Act, Every Student Succeeds Act) have strongly encouraged the use of experimental methods (e.g., randomized controlled trials), in part on the grounds that such methods are likely to yield evidence that is most useful to educators, because such methods should yield unbiased estimates of the impacts of programs, thereby reducing the risk of adopting programs that are likely to fail to improve student

outcomes (see Dynarski, 2008). Underlying this approach is skepticism regarding claims about "what works" from program developers and advocates—which necessitates designs that can provide unbiased estimates of impact—as well as maintaining distance from both developers of programs and practitioners, so as not to compromise the integrity of the research.

A distinct group of approaches to research and development that seeks to improve outcomes through more collaborative engagement with schools, families, and communities offers us a context in which to construct what we believe to be a more expansive view of quality in education research. This group includes approaches that go by many different names, from improvement science (Bryk, Gomez, Grunow, & LeMahieu, 2015) to design-based implementation research (Fishman, Penuel, Allen, Cheng, & Sabelli, 2013), to practice-embedded educational research (Snow, 2015), to community-based design research (Bang, Faber, Gurneau, Marin, & Soto, 2016). As we argue in this review, these four approaches accord a central role to "non-researchers" in evaluating quality; they employ a wide range of methods; and their accounts of research reflect a common set of principles that are distinct from and ask more of research than the U.S. research and development infrastructure promotes and actively supports. Importantly, these approaches embrace a more expansive set of outcomes for education research, including directly supporting the accomplishment of particular aims for education, either as pursued by schools today (improvement goals) or as might be imagined for educational systems that do not yet exist (transformation goals), and they imply the need for additional criteria for judging the quality of research.

Taking these various approaches as ones that share commonalities or "family resemblances" (Wittgenstein, 1953), we ask: *What accounts do practitioners of these approaches give of their work, and how do these accounts help us reconsider what counts as "quality" in research?* To

answer these questions, we describe this group in terms of common themes and distinctive emphases. We analyze the accounts that their practitioners give of their work, that is, accounts not just as reports but as arguments for the trustworthiness of their claims. These accounts specify the people and places to whom they hold themselves to account for quality, the objects that are evaluated, and the substance of their arguments. As elaborated below, "practitioners" of these approaches include, but are not limited to, researchers in academic institutions; they can include those who are often participants in research, that is, educators and community members. In the concluding part of the review, we lay out a set of conditions for developing an alternative research and development infrastructure that reflects the commitments of these collaborative approaches to research to produce equitable change in educational systems that are accountable not only to other researchers but also to policymakers, schools, families, and communities. By "equitable," we do include both improvements to existing systems that make a concrete difference to the learning opportunities, experiences, and outcomes of members of nondominant communities, including but not limited to communities linked by race, class, gender and sexual identity, home language, and ability, as well as transformations of education systems that imagine new possible goals for systems and arrangements for learning.

Defining Research Quality in an Era of Evidence-Based Policy and Practice

During the Bush and Obama administrations, U.S. federal policies strongly promoted the development and use of evidence to inform educational decision making as part of a broader movement to promote evidence-based policymaking (see Haskins & Margolis, 2015). For example, the No Child Left Behind (NCLB) Act included over 100 references to "scientifically based research," with expectations that local decision makers adopt interventions and strategies with strong evidence from randomized controlled trials (Honig & Coburn, 2008). The Education

Sciences Research Act ("Educational Sciences Reform Act of 2002," 2002) established the Institute of Education Sciences (IES) at the U.S. Department of Education to support evidence-based practice, guided by a mission "to provide rigorous and relevant evidence on which to ground education practice and policy and share this information broadly." Since its founding, IES has funded hundreds of efficacy and effectiveness studies of interventions, programs, and practices that have generated evidence from studies using experimental and quasi-experimental designs (Institute of Education Sciences, 2012).

To support proposal and study review panels, U.S. federal agencies have published extensive guidance regarding the conduct of studies. The What Works Clearinghouse, for example, has developed and revised multiple times guidelines for evaluating whether an intervention has evidence of sufficient quality from experimental studies to be included in that database (U.S. Department of Education, 2013). In addition, the Institute of Education Sciences collaborated with the National Science Foundation (2013) in the U.S. to produce common guidelines for studies at different stages of the development and testing of interventions and programs. These guidelines named six different types of supported research: foundational, exploratory, design and development, efficacy, effectiveness, and scale-up. For each, the guidelines specified purposes and expected outcomes, theoretical, practical, and empirical support required to justify funding, and ideal research designs. Funding support has varied according to the phase of research, too: both the Bush and Obama Administrations in the U.S. invested in tiered evidence grant-making initiatives like the Investing in Innovation (i3) Program. In a tiered-evidence design, programs with more rigorous evidence of impact are eligible for the most funding, while programs with less rigorous or emerging evidence are eligible for smaller

grants (Haskins & Margolis, 2015). Reliance on tiered evidence is still part of current policy as reflected in the Every Student Succeeds Act ("Every Student Succeeds Act," 2015).

An Infrastructure Focused on Expert-Driven Innovation

Policies and programs to support the generation of evidence to inform policy and practice under the Bush and Obama administrations helped to build a strong infrastructure for innovation centered on a particular model of research and development. The model imagines a research, development, dissemination, and utilization sequence, which Peurach and Glazer (2012) have called the "RDDU" sequence. A core assumption of this sequence, argue Peurach and Glazer, is that this sequence can "enable rapid, large-scale improvement by providing schools with 'research-based' and 'research-proven' programs that can be implemented quickly, effectively, and efficiently" (p. 160). In contrast with this core assumption, they note, research on large-scale reform initiatives shows that the four steps unfold concurrently and recursively over a much longer period of time than is presumed. The process is largely "expert-driven," meaning it is led by researchers who eventually are expected to transfer ownership of designs to educators who will implement those designs with integrity to the designers' intent. In the past two decades, the educational evaluation sector has moved toward this approach despite ongoing and unresolved difficulties in determining implementation fidelity, and new professional societies and advocacy groups have formed to advocate for RDDU principles (Peurach, 2016).

Undergirding this infrastructure and approach to research are two values social and educational scientists have identified as characterizing much of social science research, namely skepticism and the importance of researchers maintaining distance toward the objects and participants of research (Nzinga et al., 2018). Merton (1973) refers to science as adhering to the norm of "organized skepticism" (p. 264), in which accepted truths or premises are subjected to

detached scrutiny. Skepticism inheres specifically within the RDDU infrastructure as skepticism regarding claims about "what works" in education, that is, about programs and practices that are intended to improve student learning outcomes. It is primarily these claims that are to be subjected to scrutiny. The need for distance from one's subject matter focus and from one's study populations is often embraced as a means to protect against bias, particularly as research moves beyond an exploratory phase (Nzinga et al., 2018). Consistent with this value orientation, in the RDDU infrastructure developed for educational research, program developers and researchers closely associated with research teams are expected to play a diminishing role as programs go to scale and are tested for their effectiveness. This is so, in large part, to protect against the charge of bias in developing research designs and reporting study results.

A Family of Approaches Develops Within this Context

What follows is a comparative literature review of four models of research and development that draw on alternatives to the RDDU sequence as a way to organize research and development but that have developed alongside it. In some ways, its growth can be traced to a pair of reports by the National Academy of Education (1999) and National Research Council (2003), which called for closer collaborations between research and practice. One of those reports inspired the creation of an initial infrastructure, the Strategic Education Research Partnership (SERP) Institute, for bringing these ideas to fruition. Other large collaboratives focused on design and testing of innovations also inspired new kinds of collaborations between educators and practitioners. In the last few years, funders like the Institute of Education Sciences and Spencer Foundation (among others), began to take notice and to promote more such collaborations.

The different models reviewed here share a common aim, namely of specifying new ways to organize the enterprise of education research to be more of a "two-way street" between

researchers and key stakeholders in the improvement of practice (Tseng, Easton, & Supplee, 2017). The researchers involved see their goal as generating and systematically testing ideas collaboratively with educators. From their viewpoint, improving educational opportunity is not just a matter of researchers getting better at dissemination or supporting educators in making more use of evidence, but of engaging in principled efforts to change educational systems, how research is conducted, and researchers' relationships with educators, families, and community members. These approaches aim not just to test and develop programs and practices that are effective, but also to develop and test new norms for how research might contribute to better relationships among researchers, educators, families, and communities (Bang et al., 2016).

There have been a number of arguments made in recent years for the advantages of more collaborative approaches to research, which have partly fueled their growth. Collaborative approaches to research are argued to produce more timely and relevant research, for example (Tseng et al., 2017). In addition, research in education suggests such approaches have the potential to produce usable and effective innovations to improve teaching and learning outcomes (Coburn & Penuel, 2016). Others highlight the potential of collaborative partnerships for promoting use of research in decision making, noting that a number of studies underscore the importance of sustained interaction with research and researchers for evidence use (National Research Council, 2012). Still others have argued that the trust built through collaborative engagement may yield more accurate data (Nzinga et al., 2018).

Within this larger RDDU system, these approaches have a shared vulnerability that motivates the current review. As compared to the RDDU infrastructure, the infrastructure within the research and development ecosystem is relatively weak (Peurach, Penuel, & Russell, 2019), a point we revisit in the concluding section of this review. This vulnerability led us to explore the

possibility that the approaches might point the way to the development of a new, alternative infrastructure for research and development in the United States that values engagement rather than distance, designs for utility and relevance, and seeks the input of more stakeholders throughout the research process. This review explores that possibility, by adopting the heuristic that these alternative forms of research constitute a family, with certain "family resemblances" (Wittgenstein, 1953) that allow us to bring into focus requirements for that infrastructure.

In this review, the heuristic of treating each approach as member of a common family was a strategy for helping to identify a set of overlapping, interconnected principles of the approaches, by bringing into focus what makes the approaches similar to one another. Wittgenstein (1953) used this term to explain how it is we can recognize different kinds of "games" as belonging to the same category. He thought of family resemblances not as isolated characteristics held in common, but rather as "a complicated network of similarities overlapping and crisscrossing: sometimes overall similarities, sometimes similarities of detail" (pp. 66-67). We see and present the principles behind these family resemblances as the same way, as a set of overlapping and interconnected principles. Broadly speaking, they are unified by common themes of collaboration, to which we return in the discussion. Our approach to naming the members as constituting a family and *then* identifying characteristics that constituted them as a family was a technique adapted from developmental psychology (Werner & Kaplan, 1963). This technique can be seen as a projective, rather than inductive technique, in that it projects the family across a field of difference, and invites analysts—in this case, advocates and leaders of the approaches studied—to consider the network of similarities among them. We note this interpretation of identifying a "family of approaches" is different from an attempt to trace the lineage and relationships of different approaches, as might be anticipated by our use of this phrase. We see

this type of review as complementary to our approach, which has the potential to identify and catalyze *new* relationships that could be grounded in solidarities perceived among advocates of the approaches that are imagined or seen through this approach.

Methods

This review describes an effort involving developers of four different models of collaborative research in order to identify family resemblances (i.e., overlapping, interconnected, defining features) among them and to identify forms of exchange and collaboration that might build an intellectual community organized around developing and conducting these forms of research. It presents exemplar projects from the four approaches and describes their principles. In the concluding part of the paper, we consider what would be necessary to engage in field building to support further development of these approaches specifically as a means to expand equity of educational opportunity in a fragmented and turbulent ecosystem.

Formulating the Project

This review is the product of a project that grew out of a meeting sponsored by the Spencer Foundation in November 2016, which focused on exploring overlaps among various forms of "continuous improvement research" in education and related disciplines. The meeting involved fewer than a dozen scholars closely tied to three of the four approaches reviewed here, that is, researchers involved in the Strategic Education Research Partnership (SERP), Design-Based Implementation Research (DBIR), and Improvement Science. Others in attendance represented outsiders to the approach, more closely associated with and advocates of the RDDU approach. They had been invited as critical friends to the meeting.

That meeting was organized around a central question: If these different forms of research shared defining features in their otherwise distinctive approaches to research, how might looking

at them as sharing family resemblances help to guide the building of an alternative kind of research and development infrastructure to support their collective development? We wanted to explore what constitutes the presumed to exist resemblances among the group and not treat them as a set of isolated approaches reflective of the academy's tendency to create new individual labels for approaches that are well-established. Further, if this were a set of approaches in development, we sought to explore what its guiding *logos,* or approach to study, might be, and how as a community we might develop standards of quality through such existing venues as peer review and new venues that bring researchers and practitioners closer together.

A number of the participants in that original meeting became part of the project that emerged from this meeting, led by two of the authors and funded by Spencer. The scope and aims were jointly negotiated by the foundation and authors. Its aim was to convene developers and leading practitioners of a subset of these approaches to identify commonalities and distinctive characteristics of them and to articulate a pathway by which standards of quality might be developed for this group of approaches. Of course, we recognized that the very premise of the project—that these approaches might share enough resemblances to warrant a common name and collective investment—biased us toward the search for similarities. In addition, the idea that one can identify overlapping, interconnected characteristics of different approaches to research assumes that there are identifiable regularities of the particular approaches, even if these regularities are evolving and instantiated in different ways within different projects.

The focused scope of the project meant that participation would need to be limited to participants within the United States. This decision necessarily limits the applicability of findings to other countries, and it likely led to the exclusion of some approaches that should have been represented. Some countries, however, have research and development ecologies similar to that

of the United States, where RDDU infrastructures and collaborative approaches are both present. The Netherlands is one such example, where a strong tradition of curriculum co-design and design research has developed (Pieters, Voogt, & Roblin, 2019) within a system that also prizes impact research (Koninklijke Nederlandse Akademie van Wetenschappen, 2018). Similarly, researchers advocating for more pragmatist approaches to education (e.g., Biesta, 2007) have argued against models of evidence-based education in the United Kingdom that exclude considerations of relevance to practice, much in the way that advocates of the approaches we selected have. Thus, the study findings hold some potential relevance for countries with similar research and development ecosystems, despite the exclusion of non-U.S. participants.

Inclusion and Exclusion Criteria for Selecting Approaches

There are many different approaches to research that engage researchers and educators collaboratively, and many that also focus on solving practical problems. Coming to understand any one approach requires deep analysis of the approach, including its antecedents and related fields. We were constrained, by resources and our appreciation of the complexity of different approaches, to select a small number of approaches to analyze, recognizing the limits it places on our analysis. We sampled purposefully for similarities and differences, to help generate a set of principles that might be refined in later study. We applied four criteria to selecting approaches:

> (1) the approach needed to engage participants—and not just researchers—from different organizations in the process of collaboratively designing and testing solutions to identified problems;
>
> (2) the approach needed to address problems that require systemic change, where "system" referred to components and links within a bounded network, community, or local education agency;

> (3) the approach needed to be relatively mature: that is, to have a range of potential projects and teams to choose from, if we were to select representatives to bring together to identify principles and commitments; and
>
> (4) as a *set of approaches,* we sought some level of breadth in order to test the limits of any initial consensus regarding commonalities among the different approaches.

This fourth criterion was especially important to us, since the framing of our goal biased us toward looking for commonalities rather than distinctive features of approaches.

There were exclusion criteria as well. We excluded approaches where educators are called on to cooperate with researchers solely for the purposes of testing the impact of an innovation. We recognized such approaches require the consent and deep engagement of educators, but because educators do not have a say in the design of the innovation being tested, we consider such approaches to be outside the group of approaches we sought to investigate. Application of this criterion led us to exclude individual exemplars of the RDDU approach, as well as at least one successful model of curriculum research based on that approach (Clements, 2007), because it does not involve collaborative design. Application of this criterion also led to the exclusion of partnership approaches to research that might have been included. For example, we did not include Research Alliances (e.g., the University of Chicago Consortium on School Research; Roderick, Easton, & Sebring, 2007), which do not engage participants in the design and testing of solutions. Though such approaches offer great promise, they differ in the degree to which researchers have a stake in and share risk in the success or failure of improvement endeavors from the included approaches. We also excluded approaches that did not aspire in some way toward scale. In applying this criterion, we excluded two approaches initially discussed that were

not fundamentally oriented toward spread or broad reach at all. This included certain forms of design-based research that are focused on testing innovations in a small number of classrooms to develop theory (Cobb, Stephan, McClain, & Gravemeijer, 2001; Shaffer & Squire, 2006), as well as practitioner research (Cochran-Smith & Lytle, 1993, 2009). However, in future iterations of this activity—and with additional resources—including each of these excluded approaches in activities to identify family resemblances might be especially productive for specifying more clearly the overlapping similarities that define and distinguish this family of approaches from others in education. Remaining were the four approaches described in detail below. Each fit our criteria but differed in ways we anticipated would lead to productive dialogue, in terms of their different roots, typical partners, and ways of conceptualizing scale.

Deciding on Participants

We invited a total of 14 researchers and representatives from intermediary organizations whose work involves connecting research and practice to join us for a day and a half meeting in spring 2017. Roughly half of the participants were developers or leading practitioners of their respective approaches to research. For each model, we decided it would be important to have at least two representatives of approaches present. We first generated a list of 20 candidate attendees from projects that represented the model, based on knowledge of the authors of the breadth and maturity of work, as evidenced by both publications and funding. After eliminating one candidate project (because it was too new), invitations were extended to eight model representatives, seven of whom were able to attend the meeting. Only SERP was able to send just one representative to the meeting, while the others were represented by two distinct projects.

The other half of the participant list was comprised of scholars familiar with one or more of the approaches and representatives of intermediary organizations who were familiar with

different models. One scholar, for example, was was engaged at the time in a comparative study of different approaches to human-centered design and was investigating a DBIR partnership. One representative of an intermediary organization served as a broker between district leaders and scholars interested in both improvement science methods and DBIR. That district was also one where SERP had a strong presence, and this representative was familiar with their approach. The inclusion of these participants—who were by no means "skeptics" of the approach—was purposeful, to facilitate looking across the different approaches toward common features.

Generating Commonalties and Differences at the Meeting

Our meeting aimed to take advantage of the expertise in the group to collaboratively generate a list of similarities and differences among the four approaches. This approach was similar to that used by Munter, Stein, and Smith (2015) to identify overlapping and distinctive commitments of advocates of dialogic and direct approaches to mathematics instruction. Their approach relied on bringing people with diverse viewpoints to clarify what was at stake with competing visions for instruction, allowing for an iterative cycle of presentations and dialogue.

Over the course of the meeting, we generated ideas about both similarities and differences among the approaches, with respect to their core principles and practices. People involved in the development of each approach gave overviews of their approach, and we heard two presentations from researchers employing each approach that illustrated what it looked like in practice. At the meeting, we also had reflectors — two of whom came from intermediary organizations in education and were not researchers — provide their thoughts about each approach.

We designed the meeting to maximize time for both individual participant input and collective sensemaking. Prior to the meeting, we asked approach developers to nominate and share articles or reports that best embodied the principles and an overview of their approach.

Also prior to the meeting, participants wrote 2-page reflections on what they saw as commonalities, based on reading the papers provided. During the meeting, we set aside time for individual journaling after presentations, shared electronically with the meeting organizers, as well as group discussion. The journaling activities encouraged participants to note what institutional supports were available to them, and what changes to research infrastructure would support their work, in addition to identifying commonalities and differences between approaches. We approached the task of identifying similarities and differences across approaches from multiple angles. After each presentation, we asked participants to compare what they heard to earlier presentations in order to look for similarities and differences. We engaged in a "Venn Diagram" activity, in which each pair of approaches was contrasted to one another and people wrote about and placed sticky notes on where they thought approaches overlapped or were distinctive. These became the basis for our own team's sensemaking about participant input.

This last activity led us to conclude what should have been obvious to us as a team from the start: that the lack of opportunities for exchange among this group prior to the meeting would make it especially hard to identify differences among approaches. Advocates of approaches disagreed with many of the sticky notes characterizing ways their approach differed from the approach paired with theirs. For example, for only two models did outsiders say "flexibility is central" and "much energy is spent on cultivating relationships," about NICs and CBDR. However, both SERP and DBIR writings give great emphasis to both of these qualities. For example, SERP leaders (Donovan, Snow, & Daro, 2013) argue that successful research requires a "major investment in establishing relationships" (p. 404) and "flexibility" (p. 413). Discussion did not resolve these differences, either.

By contrast, we had much greater success with identifying points of commonality, in particular with respect to what we describe here as *shared principles* of this group of approaches to research and development. We assembled input from the meeting and proposed a set of shared principles to the group, and after reviewing all written products related to an exemplar project chosen in consultation with approach developers, we presented an analysis of each project using the shared principles. This paper focuses on those shared principles and how they are exemplified in specific research projects, selected through a process described below.

Generating and Member Checking Descriptions of Approaches

Having elaborated these shared principles, our team constructed accounts of a single exemplar project for each approach. An exemplar project was defined as one that was educationally focused, had a long history in terms of at least one iteration in the educational improvement effort, and was confirmed to be an appropriate selection by the developer or advocate of the approach. We also looked across the exemplar projects to identify similarities and differences in the depth to which written accounts depicted the full picture of the approaches in action. This allows us to make initial recommendations for peer review processes and collaborative exchanges that might be necessary to develop shared standards of quality.

Once we identified exemplar projects, we began to collect accounts of each, using both online databases (e.g., Google Scholar) and through direct communication with researchers involved in a project. Our methodology for selecting and reviewing the written material proceeded as follows. First, we gathered all documents that described the empirical results of research focused on the educational improvement effort. This included peer-reviewed journal articles, book chapters, and white papers available on project websites. We next collected supplemental documents (e.g., book chapters and white papers) that provided a narrative of the

life course of the exemplar project, including the motivation for the partnership formation, rationale for selected research questions, and explanations for how the focus of the partnership evolved over time to address the needs of the educational partner. After that, we developed a set of questions for helping us select information from these accounts related to the shared principles and commitments. Two researchers per approach then gathered this information from each document into a matrix. They focused on one question at a time, for example: how does this account describe the aims of the project, how are those aims justified, what is the setting of the research, and what is its sampling strategy. The entire research team met to discuss the preliminary codes for each question and to adjudicate discrepancies after each of the three rounds of analysis. This ensured that the assigned readers were interpreting the shared principles and associated guiding questions in a consistent manner. For instance, two researchers disagreed about how expertise was valued in the DBIR project. During a research team meeting the differences of interpretation were discussed, the team referred to the other literature on the project, and one researcher per approach made notes of the questions to be asked during later member checking. Four researchers looked across the codes to synthesize an account of each approach that integrated information from the different sources. Finally, one person per approach used the synthesis to generate a narrative account of the shared principles. The draft analysis for each approach's exemplar project was sent both to meeting participants and to the project lead for review and feedback to increase the credibility of our analysis (Lincoln & Guba, 1985). Two meeting participants provided feedback on the drafts and our team also spoke with one researcher who had been directly involved in each of the projects. The meeting participants provided their thoughts regarding the utility of our analysis and suggestions for further research, while the project leads surfaced instances where shared principles were present during the life

course of the project but were not visible in written work. Our accounts were then revised to incorporate the meeting participants' and the researchers' feedback.

Descriptions of the Four Approaches

In this section, we provide a brief overview of each of the four approaches reviewed here. This is by no means a complete analysis of each approach, and it is intended primarily to provide a context for a discussion of the principles of the family of approaches. We have sought to avoid language that implies distinctiveness of each approach relative to the others, though we know there are distinctions. Indeed, we were unable to come to agreement on those distinctions. For more detailed accounts of these approaches and exemplars, see Penuel and colleagues (2017).

The Strategic Education Research Partnership Approach

Initiated in 2003, the Strategic Education Research Partnership (SERP) approach was the first of the four reviewed approaches to be developed, and it was the first to define a set of principles for a collaborative problem solving approach. The origins of the SERP approach and institute trace back to a National Research Council committee report by that name (National Research Council, 2003). Calling for a more "vigorous connection between research and the practice of education," the report defined SERP as (1) a program of research focused on building and testing solutions to persistent problems of practice, (2) two-way partnerships between leading researchers and educators in school districts that would function as field sites for the research, and (3) an organization that could serve as an infrastructure for a network of field sites that could learn from one another's activity (National Research Council, 2003). Though the $500 million investment called on in the report to build the network of field sites did not materialize, in 2005, the Spencer Foundation and The William and Flora Hewlett Foundation provided funding to launch an initial partnership with Boston Public Schools (Donovan et al., 2013).

Informed by research in other sectors (e.g., medicine, agriculture, and transportation), SERP is premised on the idea that sustained, collaborative efforts of researchers, policymakers, and practitioners can produce high-quality, effective, and relevant work (Donovan, 2013). The descriptor "problem solving" approach used to characterize the approaches we reviewed follows from this premise, that there is benefit in designing partnerships "to follow the contours of problems" (Donovan, p. 318) as they identify systemic, sustainable solutions to problems.

In SERP, field sites initiate problem identification (Donovan et al., 2013). Once the local partner identifies the problem, SERP locates researchers with expertise in the problem area of interest and experience for further refining the problem and who can help design improvement strategies. These researchers are charged with bringing research knowledge to help frame problems and with drawing on the best evidence available on successful interventions to solve the focal problem. This serves as the starting point for regular partnership meetings during which the evidence is reviewed and adapted to meet the unique needs and organizational characteristics of the local context (Donovan & Snow, 2018).

In the early stages of endeavors, the SERP approach leans most heavily on administrators and teachers who demonstrate a clear vested interest in supporting the work to provide critical feedback in the design and piloting phases. A primary goal of SERP is to design interventions that are practical for both the educational partners as well as educational organizations outside of the partnership. This is done by intentionally developing programs that are feasible to implement within the demands and restrictions of the educational context, and to make the resulting tools, intervention materials and resources for professional learning freely available to others interested in benefiting from the partnership improvement efforts.

The SERP Institute plays a key role in facilitating collaborations between researchers and practitioners; it operates as an intermediary organization that connects researchers and partners at the educational organization of interest in three main field sites and one geographically distributed collective: Boston, Baltimore, the region of San Francisco and Oakland, and a network of schools and researchers called the Minority Student Achievement Network.

Design-Based Implementation Research

Design Based Implementation Research (DBIR) is an approach to research and development that brings together two different research traditions: design-based research and implementation research. Its broad aim is to support equitable change in educational systems such as states, districts, and learning "ecosystems"[1] in communities through collaborative design and testing of solutions to persistent educational problems.

Design-based research is a signature approach of the interdisciplinary field of the learning sciences in which teams organize or "engineer" new approaches to learning in order to study the conditions under which they can be supported (Design-Based Research Collective, 2003). Design-based research typically takes place in small numbers of classrooms, and it can yield both new learning theories and concrete designs and materials for learning (Edelson, 2002).

Implementation research in education often focuses on what happens when policies and programs are brought to scale. It seeks to describe and explain patterns in implementation by examining individual, interpersonal, organizational, and institutional processes (Spillane, Reiser, & Reimer, 2002). Like implementation research, DBIR, too, can yield new theoretical insights related to how and when policies and programs can spread and be sustained (e.g., Frank, Zhao, & Borman, 2004). DBIR seeks to study and engineer these conditions, using theories, design

processes, and approaches of both design-based research and implementation research (Fishman & Penuel, 2018; Fishman, Penuel, Allen, & Cheng, 2013).

DBIR traces its origins to multiple sources and is best understood as a name for forms of research and development that share common features (rather than common methodology). In addition to design-based and implementation research, inspirations for DBIR include participatory evaluation research, community-based research, and social design experiments. The SERP approach is another inspiration, and Word Generation appears in the article as an example outlining the key features of DBIR. Describing DBIR collaborations, Fishman and colleagues (Fishman et al., 2013) named four core principles of this work:

- A focus on persistent problems of practice from multiple stakeholders' perspectives
- A commitment to iterative, collaborative design;
- A concern with developing theory, knowledge, and practical tools related to both classroom learning and implementation through systematic inquiry
- A concern with developing capacity for sustaining change in systems.

DBIR is an approach that is suitable for long-term research-practice partnerships, because it entails designing supports for classroom learning and broad implementation across levels of a system (Penuel & Gallagher, 2017). Projects that use DBIR include actors at multiple levels of the system. They bring together classroom teachers, administrators, researchers, and other stakeholders. DBIR follows in the tradition of design research by working outside lab settings to design and study what makes innovations successful in classrooms and other real learning environments (Fishman et al., 2013). In many projects, researchers assist school staff with the implementation of an intervention, often co-designing tools, resources, and solutions to the identified problems of implementation.

Improvement Science in Networked Improvement Communities

The Improvement Science in Networked Improvement Communities (IS/NICs) approach began as an exploration by Bryk and colleagues into two coordinated approaches that might inform the building of a research and development infrastructure for American education that could support dramatic improvements in outcomes reliably and at scale (Bryk, Gomez, & Grunow, 2011). The leading organization for adapting these approaches has been the Carnegie Foundation for the Advancement of Teaching.

The first approach on which Carnegie draws is *improvement science* as developed within health care (Berwick, 2008) and as inspired by applications of Deming's (1993) writings on performance improvement. Improvement science is a form of systematic inquiry that begins with the definition of a persistent problem and, from there, maps backward to the system that (re)produces the problem, an improvement aim, a system for measuring progress toward that aim, and a design for devising and testing possible solutions using rapid Plan-Do-Study-Act cycles. In improvement science, inquiry is guided by three basic questions: What are we trying to accomplish? What changes can we make that will result in improvement? How will we know that a change is an improvement?

The second approach on which Carnegie draws is *networked improvement communities* as an organizational form through which to operationalize improvement science. As explained by Bryk et al. (Bryk et al., 2015), inventor and engineer Douglas Engelbart (1992) first coined this term to refer to a way that a high performance organization or network of individuals who are engaged in efforts to improve practice might organize itself to "get better at getting better." He asserted that people working on common problems and solutions to those problems could be

more effective not only if they formed a network but also if they engaged in collective work to improve *how* they went about their work of developing and testing solutions.

In Carnegie's approach to networked improvement communities (NICs), members use principles of improvement science in combination with resources and routines to address practical problems and to improve work processes within educational systems. A NIC uses specific improvement science techniques to guide the engagement of local practitioners in working iteratively, over time, to study the problem in local contexts, develop initial theories of problems and solutions, devise and test interventions, and review theories of problems and solutions in light of outcomes. Such work can have focus exclusively on locally designed solutions. It can also focus on incorporating, using, and refining externally developed solutions.

Community-Based Design Research

Community-Based Design Research (CBDR) is an approach to research that is centered in interactions between communities and educational systems. In our view, CBDR belongs in this family of approaches because of its focus on collaborative design and inquiry and because of its attention to public education as a democratic enterprise. It pushes the other approaches to consider ways of broadening and deepening approaches to inclusion in improvement efforts in order to interrupt and address persistent inequity in education. The approach does not draw a distinction between professional researchers and others in inquiry. In CBDR, the positioning of participants as co-researchers is inspired by a concern with democratizing the search for solutions to persistent problems and with repairing harm caused by past—and often continuing—exclusion of community stakeholders such as youth and their families from deliberations about what problems are most important to solve (Bang & Vossoughi, 2016).

CBDR draws from design-based research in its emphasis on the design of new forms of learning in real educational settings. But where design-based research has been largely silent on questions related to power (e.g., "Who designs?") and persistent inequality (e.g., "Why do inequalities of opportunity persist for racially minoritized students?"), community-based design research puts these questions in the foreground (Bang & Vossoughi, 2016; Le Dantec & Fox, 2015). Further, CBDR focuses on expanding participation in design, naming and disrupting historical inequities, and changing institutional relationships (Bang, Medin, Washinawatok, & Chapman, 2010). As with other forms of participatory and community-based research, CBDR places strong emphasis on the values of social justice and of promoting the agency of participants in research, as well as accountability to participants in research (Bang et al., 2016; Cammarota & Fine, 2008; Campano, Ghiso, & Welch, 2015).

CBDR draws primarily from anthropology and cultural psychology, both of which emphasize the ways that people live and learn culturally (Nasir, Rosebery, Warren, & Lee, 2014). Within anthropology, it draws on traditions emphasizing the need to decolonize educational research and adopt a more relational stance toward participants in communities that values them as equal participants in research (Campano et al., 2015; Patel, 2015). From cultural psychology, this tradition of research adopts a view of learning as a cultural process, while at the same time cautioning researchers about taking a monolithic, static view of culture and how people inhabit cultural identities (Gutiérrez & Rogoff, 2003; Nasir et al., 2014). Further, it often centers analysis of issues of power, race, and social status in research in order to diagnose inequity and promote more just and equitable change in systems through what have been called "social design experiments" (Gutiérrez & Jurow, 2016).

Interconnected Principles of the Approaches

We identified eight interconnected principles among the four approaches. Next, we describe these principles and illustrate them through four exemplar projects, each from one of the four approaches to research described above (SERP, NIC, DBIR, CBDR). To provide further organization, we describe two different principles across each of the four elements we have identified as common to this family: beginning with its *collaborative* nature, continuing with its focus on *problems*, following that with a description of commitment to *solving* problems through design, testing, and evaluation, and concluding with its *research* commitments.

Table 1 summarizes how each of the projects representing their approach embodies the principles identified. For space considerations, we have highlighted only some projects under each principle. The reader interested in a more elaborated description of each project's embodiment of the principles is directed to a technical report (Penuel et al., 2017).

Insert Table 1 about here

The *Collaborative* Nature of Research in this Family of Approaches

All the approaches in this family shared a commitment to conducting collaborative inquiry. This collaboration is both a means to supporting a key goal of research and something about which researchers must give an account. We continue, then, by detailing each of these principles. We also animate the shared principles by drawing selectively from each of our four approaches to highlight distinguishing characteristics.

Principle 1. The work should support the agency of participants by collaborating with them as partners in research. One of the impacts sought by researchers in this family is an expanded sense of agency, that is, a greater sense of the possibilities for participants to change their local environments (Campano et al., 2015). With respect to classroom-level instructional

reforms, that might entail supporting teacher agency, whose voices are not always integrated into designs for improving instruction (Severance, Penuel, Sumner, & Leary, 2016). In other instances, the research might seek to amplify the voices of parents and community members as advocates for their own and other children (Booker & Goldman, 2016). In still others, the intent may be to expand the power of marginalized groups to have a say in the aims of school reform and the organization of educational opportunity (Renée, Welner, & Oakes, 2009).

Each of these approaches supports agency by engaging participants as collaborators in core activities of research. Participants may be involved in defining questions to be pursued, specifying measures, data collection and analysis, and communicating findings. Importantly, they also engage participants as co-designers of solutions to problems (Ishimaru & Takahashi, 2017). As such, many practitioners in this family of approaches draw upon traditions of participatory action research and participatory design. Community-based design research (CBDR), for instance, explicitly draws links between its approach and both participatory action research (Whyte, 1991) and participatory design traditions from Scandinavia (Ehn, 1992).

The PRIMES project is an example of a Community-Based Design Research project that engages parents as key collaborators in research. It began with a premise that parents bring important resources and expertise that should be valued in and outside of schools. PRIMES, which stands for Parents Rediscovering and Interacting with Math and Engaging Schools aimed to bridge out of school learning within school math and restore a sense of "epistemic authority" as a basis for supporting parents' sense of mathematical agency (Goldman & Booker, 2009). The emphasis on using collaborative practices as a means of supporting the agency of participants and the accountability of the researchers is typical of CBDR projects (Bang et al., 2016; Cammarota & Fine, 2008; Campano et al., 2015). Using ethnographic methods and a

participatory design process, PRIMES researchers and parents developed tools and resources for families to support math engagement at home and through school-based workshops.

Within the context of PRIMES, deep collaboration between community members and researchers took the form of co-designed family workshops. Recognizing that families have mathematical experiences that are rich in content, but rarely recognized in schools, PRIMES was created to disrupt historical inequities and power imbalances between lower income parents belonging to minoritized groups and the school system. By collaborating with parents to create a program *with* them, rather than *for* them, research with PRIMES directly positioned parents as co-equal with researchers. In this way, parents were able to help define the problems worth investigating and had a say in the design and analysis of the project, for example, by providing crucial feedback regarding what was and wasn't successful in iterations of the parent workshops.

As with many CBDR projects, PRIMES was centered in the community rather than schools, but the local schools were sites of collaboration. By taking research into the domain of children and their parents, families were able to co-design solutions to problems they found relevant.

Principle 2: Accounts of the work should clearly define and describe the role and contributions of partners, particularly their expertise and how it was integrated into the research. The first principle's commitment to supporting agency through participation in research has a corollary principle in communicating about the research. Those communicating about the work have a responsibility to give an account of the role and contribution of key partners in the research. One way that researchers can do this is by analyzing different ways that participants engage in core activities (e.g., Potvin, Kaplan, Boardman, & Polman, 2018). In many instances of CBDR, educators and community members may describe their roles in the first person, as co-authors or speakers who share the stage with researchers (e.g., Campano,

Ghiso, & Welch, 2016). Whatever the approach, it is not sufficient to assert the research is participatory; communication about the research must provide explicit warrants for how the expertise of participants is reflected in the focus, process, and outcomes of research.

The SunBay Digital Mathematics Project represents an example of an effort to clearly define and describe the contributions of participants from both research and practice into a DBIR project. The project involved the adaptation of a curricular intervention focused on middle school mathematics, SimCalc Mathworlds, which focused on helping middle school students develop mathematical ideas about change and entailed solving problems related to rate and proportionality that involved coordinating multiple ways of representing functions (Roschelle & Hegedus, 2013; Roschelle, Kaput, & Stroup, 2000).

For SimCalc researchers, a key goal was to explore how to build capacity of the school system to adapt, integrate, and sustain SimCalc over a long duration and in a way that maintained effectiveness demonstrated in a pair of randomized controlled trials (Roschelle, Shechtman, et al., 2010). The team partnered with a district with support from a local foundation. Participants included researchers, educational leaders, technology designers, and teachers. The primary focus of the work was to integrate the materials into the local instructional guidance system of standards and pacing guides and build teacher leadership to support ongoing implementation.

One way this project is explicit about the contributions of partners and how their expertise was integral is that their descriptions of the work give both a summary and narrative account of contributions. Roy and colleagues (Roy, Vanover, Fueyo, & Vahey, 2012) include a chart of major contributions and expertise of the different stakeholders: university researchers brought their expertise in teacher learning to the task of providing professional development; researchers from SRI provided the curriculum and led SimCalc-specific professional development sessions;

and district leaders provided "cover" for teachers to implement units as well as extensive local knowledge of participating schools, which facilitated implementation.

Other accounts describe how teacher knowledge was integral to the project. Teachers were not simply the recipients of professional development; they were called upon to "provide the core of expertise in the district" (Vahey, Roy, & Fueyo, 2013, p. 187). Teachers were positioned as sense-makers and actively participated to shape their own learning. This was a deliberate move. By ensuring teachers would explore their own learning experiences with the curricular content and activities, partners were able to link teacher thought processes to those of students, which helped teachers and students become co-learners in the classroom.

Valuing the contributions of all partners involved giving up the idea of "fidelity of implementation" as a goal, that is, asking teachers to implement materials as is, without adaptation. Instead, in the SunBay project, teachers were deliberately encouraged to make adaptations to the curriculum materials and teach them in whatever way made the most sense. The researchers grounded their conviction in the principle of supporting teacher adaptation on the basis of past SimCalc research that had found a wide variety of teaching strategies resulted in effective implementation (Roschelle, Pierson, et al., 2010).

Focused on Addressing *Problems*

This family of approaches is centered on practice, specifically on addressing the problems that arise from attempts to accomplish particular goals for educational equity, helping educators grapple with dilemmas and puzzles encountered in practice, and developing and expanding local "sites of possibility" within educational systems and communities.

Principle 3: The problem takes center stage in research and should be important to a broad range of stakeholders. In these approaches, the "problem" that is the focus of research

and development is always one that is defined not by a single actor (whether the researcher or practice partner) but by a group of stakeholders (Fishman, Penuel, Allen, Cheng, et al., 2013). These approaches assume multiple types of actors should have a role in defining the problem and research questions to be answered. As principles 1 and 2 indicate, each of the members of this family of approaches accords strong value to having the perspectives and expertise of a broad range of stakeholders involved in educational improvement efforts. Advocates of these approaches believe that drawing on these perspectives is especially helpful in leading to a deeper understanding of the problem, as well as to insights valuable in searching for solutions to the problem (Bryk et al., 2011; Donovan & Snow, 2018).

The term "problem" does not always adequately capture the ways that research in this family centers on educational opportunity and practice, however. Problems are always framed in relation to both aims and values (Bryk et al., 2015; Penuel & Gallagher, 2017). That is, a problem exists because it stands in the way of an aim that may be explicit or implicit, or because current practice does not adequately reflect the values that participants hold. An alternate framing is to center their efforts on identifying "sites of possibility" (Bang, personal communication) where common aims and values are already partly reflected in practice, that is, where there is already some evidence of participant agency and educational opportunity. This term was first used by Giroux (1986) as a call to critical researchers in education to balance their analyses of educational inequality with efforts to engage with educators and communities to imagine educational alternatives, including ones within schools. It has since been embraced especially so by community-based researchers seeking to identify and augment spaces for youth and parents to organize for more equitable learning opportunities.

The Carnegie Math Pathways project embodies both this principle and a key principle of improvement science, namely, that it is critical to "make the work problem-specific and user centered" (LeMahieu, Grunow, Baker, Nordstrum, & Gomez, 2017, p. 13). The focal problem around which the Carnegie Math Pathways network first organized was the high failure rate of students in community college mathematics. At that time, approximately 60% of America's community college students began postsecondary studies with a developmental mathematics course, but 80% of them did not move on to earn college-level mathematics credit even after three years of enrollment (Bailey, Jeong, & Cho, 2010). Instead of being an avenue for college success as they had been intended, conventional approaches to developmental mathematics functioned essentially as a roadblock to students' certifications and degree attainment. Led by the hub organization, the network engaged in activities to better understand the problem and the systems in which local instances of the problem operated (Yamada & Bryk, 2016).

In this and other projects facilitated by the Carnegie Foundation for the Advancement of Teaching, networked improvement communities conduct extended analyses of the focal problem both before and throughout the design process. Ahead of design, teams use tools for helping them to "see the system" that is producing the problem and making it persist. One such tool teams create is a "Fishbone" diagram, which comes from Japanese manufacturing, and is so named because it looks like the skeleton of a fish (Ishikawa, 1968). A Fishbone diagram represents team members' ideas about different sources and causes of a problem within the system. Working collaboratively to identify the (often multiple) causes helps illuminate the complexity of a given problem space for a networked improvement community.

Principle 4: The research should attend to context. In each of these approaches, there is a strong emphasis on the context of educational change efforts. Context is in the foreground, rather

than the background, because context and intervention are considered integrally related to one another (Cole & Packer, 2016; Datnow, Hubbard, & Mehan, 2002). Where traditional models seek to identify what things work, *ceteris parabis*, these models never see "all things as equal" and, instead, focus on what works where, when, and for whom (Bryk et al., 2015). To be successful, innovations need to be "adaptively integrated" (Bryk et al., 2015; Hannan, Russell, Takahashi, & Park, 2015) or integrated into existing infrastructures of systems (Penuel, 2019).

These approaches assume persistent problems have systemic causes, not single causes. Indeed, this is where "following the contours of the problem" leads them: deep into the systems that give rise to problems and cause them to persist. This principle relates to the one above, in that within a system, actors have only a partial view from where they sit and perhaps have access to a few other perspectives through their social interactions. From advocates' perspectives, some effort to elicit and integrate perspectives on systems is necessary to diagnose the educational problems that are the focus of the research and development effort (Bryk et al., 2015; Donovan et al., 2013). Teams can and do employ a variety of tools to analyze systems collaboratively, from root cause diagrams (Males, Sears, & Lawler, 2020) to activity system models (Kali, Eylon, McKenney, & Kidron, 2018) to maps showing networks of policies, people, and processes that explain why systems look the way they look (Riedy et al., 2018).

SERP's development of Word Generation, a literacy intervention for upper elementary and middle school students struggling with reading comprehension in core subject areas, illustrates how this family of approaches attends to context throughout the research and development process. Prior to creating initial designs for an intervention, researchers at SERP conducted classroom observations of instruction as well as interviews with teachers and district leaders in order to gain an understanding of the barriers to student reading comprehension. They relied on

teachers to help set constraints on potential solutions (e.g., the amount of class time that could be devoted in core subjects outside English/language arts for teaching comprehension). And, after initial pilots in multiple classrooms, the design team made modifications to accommodate middle school classrooms in different organizational contexts within the same partner district. For example, for those middle schools configured for grades 6-8 (as opposed to the schools that housed grades K-8), instructional planning tended to occur within siloed departments such that teachers only communicated about teaching activities within subject areas. Because the challenges of struggling early adolescent readers required multi-subject area interventions, Word Generation developed opportunities for cross-subject communication without overburdening the teaching staff (Snow, Lawrence, & White, 2009).

Designing and Testing Practical *Solutions* to Problems

A design orientation is a distinctive characteristic of the family and is evident in two principles that focus on practical value and on giving an account of and making use of implementation evidence to improve solutions.

Principle 5: What makes research valuable is that it provides something of practical value to participants and their organizations or communities. A key determinant of the value of a project in this family is that it can inform what individual participants in the research do in their day-to-day work. The "what" can encompass ideas from research, curriculum, tools, and practices—that is, more than just knowledge that might be developed for the benefit of other scholars (Edelson, 2002; Ikemoto & Honig, 2010). Researchers proposing new studies should be able to make an argument for the significance of what would be produced for practice, to demonstrate that proposed innovations or initiatives would in fact yield improvements that the participants valued (Gutiérrez & Penuel, 2014). A key part of practical value is demonstrating

the usability of an innovation. That is, it should not only accomplish the goals intended for improving educational opportunity and outcomes but also be implementable at scale, given supports that can be readily put into place in the context (Fishman et al., 2013).

For these approaches, there is an additional "stakeholder" in the research, namely the organization or community in which the research is taking place. These approaches emphasize the need for research activities and results to help build or transform cultures of organizations in ways that support use of research or use of evidence-based innovations (Henrick, Klafehn, & Cobb, 2018). Collaborative education research is sometimes directly involved in helping to design or redesign educational infrastructures necessary for improvement. These infrastructures include policies and guidance that are likely to increase the likelihood that interventions will be implemented with integrity, that is, in a way consistent with designers' intentions (Penuel, 2019).

The uptake of the Word Generation intervention developed by SERP for struggling readers provides evidence of the practical value of the intervention to teachers and the district. Other schools within Boston chose to implement the program after the piloting work of the program was completed in a handful of schools and the participating educators expressed their appreciation of the program and its results (Donovan & Snow, 2018). Ongoing research by SERP researchers identified conditions that acted as supports and barriers to implementation, helping the district target resources to address inequities in implementation (Elmore & Forman, 2011). SERP subsequently made the Word Generation materials and implementation guides freely available to educators beyond BPS and have reported over 20,000 downloads by registered educators (Donovan & Snow, 2018). While downloads are not necessarily measures of quality of use, they are one indicator of the wide reach of the Word Generation program. Further, uptake in other districts demonstrates the ability of the SERP partnership approach to produce educational

improvement strategies that are feasible for practitioners to implement in real world settings both in and beyond the places where they are developed.

Principle 6: The research should account for the gap between what was intended and what was accomplished. In some research, a study ends with a descriptive or explanatory account of variation in implementation of an intervention or other solution to a problem. In this family of approaches, the study of variation is a fundamental resource for further collaborative learning, innovation, and improvement. From the standpoint of these approaches, there is a strong need for giving an account in research of the ways that projects failed to accomplish their intended goals. In iterative approaches like this tradition's, failure is expected and is thought to be a vehicle for learning (O'Neill, 2016). However, there is a need to describe failures directly, to inform the work of others. Researchers emphasize that implementation involves adaptation, as individuals work to make innovations fit into their local context (DeBarger, Choppin, Beauvineau, & Moorthy, 2013). Further, they seek to learn from these adaptations, by documenting the work of supporting implementation and by identifying "productive deviance," that is, outliers which others can learn from (Bryk et al., 2015), as well as outliers where extra support for implementation may be needed. By doing so, they hope to inform others as to what it takes to make things work in a real educational setting.

The accounts of the SunBay DBIR project openly present the challenges the team faced in the project in ways that embody this principle in action. They include descriptions of how teacher turnover partly undermined their efforts to build a cadre of teacher leaders who could help others learn how to implement the units and advocate for them. In addition, they noted that technology was not always available or in good working order. Given the centrality of technology tools for the curriculum, this was a significant obstacle. Moreover, team members

discovered that teachers' comfort with using technology in the classroom varied widely. They found that some of the participating teachers did not use the technology with the level of intensity they forecasted during the initial professional development (Roy et al., 2012, p. 156).

Research **in the Family**

Research in this family of approaches, like all other research in education (cf., National Research Council, 2002), aspires to answer questions systematically by gathering and interpreting evidence that contributes to new ideas and informs future research. There are specific demands within this family, however, for warranting iterations in design and for spreading ideas, tools, and findings to other educators and community members within and outside the partnership.

Principle 7: The research plan should include specific, logical, and coherent plans for studying and following problems; for designing, testing, and iterating upon solutions; and for constructing and using practical knowledge. Though this family of approaches embraces uncertainty and complexity, and such embraces always require fluidity and adaptability, that does not absolve teams engaged in these forms of research from being explicit and rigorous in detailing their plans for navigating uncertainty and complexity. As should all empirical researchers, these researchers select methods appropriate to the question at hand, use systematic forms of data collection and analysis, and develop claims that are supported by and do not go beyond the evidence available. They expect results to be explained, not just described, and there is a strong value accorded to clearly specified conjectures or a theory of action that can be tested, as part of evaluation of the worth of the project or proposal. The findings, whether or not they are encouraging, should also resonate with key stakeholders and participants in the project.

Importantly, methods and presentations of findings are held to community norms for what counts as evidence. Some of the approaches have articulated some criteria for quality which, when articulated, make apparent a way to evaluate studies within an approach. For example, Improvement Science in education continues to adapt well-established methods from medicine (Berwick, 2008). Similarly, Design-Based Implementation Research is developing principles for argumentation that draw from design-based research (Fishman & Penuel, 2018).

The Carnegie Math Pathways project illustrates the diversity of methods and approaches used in projects within this family and how these are linked to different phases in the research and development processes. Early on in the project, scholars sought to identify key indicators that educators could use to find out which students might need extra support to succeed in developmental mathematics courses. They constructed a Driver Diagram—a model or theory of systems change—informed by a systematic review of theory and evidence regarding student motivation and learning to guide their research and development efforts. They discovered through the simultaneous iterative testing and refinement of a brief intervention and practical measures (i.e., Plan-Do-Study-Act cycles) designed to reduce students' feelings of uncertainty as to whether they belonged in class could improve success rates in classes using curriculum materials introduced by the project (Bryk et al., 2015).

Once they had refined their initial intervention, the team developed studies to help them investigate variation in effects within and across local contexts. In an evaluation of Quantway 1, the first semester of one of the curricular pathways, researchers from the hub organization did not stop at estimating its effect size. They also explored the variation in effects across different student subgroups, teaching faculty, and colleges within the network (Yamada, Bohannon, &

Grunow, 2016). By doing so, they were able to determine that Quantway 1 effects were positive across all student subgroups of sex, race, and ethnicity.

While these studies provided evidence that Quantway 1 had positive effects in classrooms and colleges in the network and *could* work for different faculty in various institutional contexts, they also revealed that Quantway 1 worked especially well in some places but not well in others (Yamada et al., 2016). Exploring that variation within and across local contexts is an opportunity to leverage the power of the NIC approach. Members seek to understand the micro-processes that either produced strong effects or where work routines broke down and effects were not positive. Discerning which students benefited from a new design and which students did not was viewed as a key activity and space for further inquiry by partners (Yamada & Bryk, 2016).

Principle 8: The research should be of value to others outside the partnership. As in other forms of research, researchers within this family of approaches seek to produce knowledge and tools of value to people beyond the immediate setting for research. Researchers sometimes refer to this as "generalizability," but the meaning is somewhat different for scholars in this tradition, who emphasize that any idea, practice, or program will need to be adapted in a new context (Gutiérrez & Penuel, 2014; Henrick, Cobb, Penuel, Jackson, & Clark, 2017). Moreover, these scholars emphasize that the tools and practices may be taken up by practitioners elsewhere, without the mediation of researchers. The notion of "transferability" of research from qualitative inquiry (Lincoln & Guba, 1985) is relevant here to describing how ideas, tools, and conclusions might be transferred or re-contextualized by others for their own purposes.

One of the biggest impacts of the PRIMES community-based design research project on the research community has been to help open up the space of designing for mathematical learning in families within the learning sciences and to set a precedent for using the approach developed

within PRIMES of observing family practices and identifying mathematical practices has been a common feature in this new line of research. For example, some members of the PRIMES team, joined by other teams from the NSF-funded Learning in Informal and Formal Environments (LIFE) Center, employed the ethnographic methods used to generate examples of mathematics problem solving in families to inform participatory design of mobile applications to foster fun mathematics learning in families (Esmonde et al., 2012). Other researchers have used the approach of identifying stories of mathematics use from families to explore the ways family goals, commitments, hopes, and demands of practical activity shape and motivate mathematics problem solving at home (Pea & Martin, 2010).

The research has also informed an emerging literature that focuses on changing conceptions of the relationship between families and schools. For example, research on the topic of "brokering" and parents' roles in supporting students in finding opportunities to pursue educational opportunities across different settings related to STEM has drawn on lessons from PRIMES (Ching, Santo, Hoadley, & Peppler, 2016). Inspired by the idea of "systemic repair," scholars seeking to shift away from school-centered parent engagement models have also pointed to PRIMES as an example for how to do so (Jay, Rose, & Simmons, 2017).

Discussion

The principles that animate four forms of research we investigated in this project elaborate important facets of collaborative approaches to designing and investigating solutions to significant educational problems. The different facets speak to the ways these approaches seek to support the agency of participants in research through collaborative engagement, and they also hold themselves to account for evidence of participant agency. The research is centered on the concerns and aims of participants, rather than solely on researchers' goals for knowledge

building in their academic fields. As with the focus on agency, researchers are expected to give an account of the broader context of their work and articulate *for whom* particular concerns are central and how they know. The research itself—as all research is or aspires to be—is systematic and seeks to be useful to others beyond the immediate context, but it prizes the production of knowledge and tools that have practical value to participants.

Contemporary collaborative approaches to research and development have different roots within education, but the principles uncovered in this study suggest an underlying commitment to some core tenets of *pragmatism*. These include a commitment to conducting inquiry that is grounded in particular questions or problems and concerned with studying and changing situations in which human beings find themselves (Legg & Hookway, 2019). These approaches prize being "problem centered" (Bryk et al., 2015) or "problem oriented" (National Research Council, 2003) in their inquiry, rather than purely theoretical. Similarly, the approaches share a desire to produce knowledge that is useful for practice. Design-based research in particular seeks to avoid epistemological and theoretical approaches that lack utility for informing design decisions toward the improvement of practice (Design-Based Research Collective, 2003). Other approaches included similarly prize what can improve practice, even if it requires drawing on constructs from different sub-disciplines to do so (see, e.g., Bryk et al., 2015). These approaches also share pragmatism's commitment to democratic ideals of shared inquiry. In some writings, there is an explicit focus on promoting "democratic dialogue" through research (Bang & Vossoughi, 2016, p. 185), while others speak to how research can support schools in their mission to educate young people for democracy (National Research Council, 2003, p. 89). Ultimately, it is this commitment to *shared* inquiry—that is to collaborative engagement with

education stakeholders in research and development—to transform systems that make these approaches stand apart from the RDDU-based approach to educational improvement.

Describing the approaches as sharing a pragmatism and also of valuing collaboration begs the question of "from whose perspective." Indeed, each of these approaches must grapple with the questions of not just the "how" and "for what" of designs for change, but also the "for whom" and "with whom" in concrete terms and in concert (Philip, Bang, & Jackson, 2018). The "for what" of these projects and approaches vary with respect to whether they seek to improve existing educational systems or transform them altogether. The people "for whom" and "with whom" designs are undertaken are always positioned in ways that are already raced, classed, and gendered within unequal societies. Democratizing research as an ideal necessarily requires attending to how the expertise and knowledge of actors in collaborative research is taken up, ignored, resisted, and changed (Biesta, 2007). Otherwise, the aspirations of collaborative research for democratization, inclusion, and equity cannot be met.

Toward Criteria for Quality in Research Design, Conduct, and Reporting

Up to this point, we have not addressed one of the key goals of the initial project, to identify a set of criteria for quality for this family of research. Without naming specific practices shared by approaches, it may be premature to specify when and how, for example, particular methods for fostering collaboration in research, might best be used. However, the principles that guide the approaches do point to some criteria for judging the quality of research, which might be applied to projects of the kind we have examined here. We offer these criteria as provisional ones, and in so doing, seek to address questions that arise from the effort to distance this family of approaches from an RDDU approach. Here, we highlight both differences and potential areas of overlap.

Research is accountable to research peers and to stakeholders. One distinction from the RDDU approach in terms of quality criteria pertains to who can judge its quality. In an RDDU approach, peer review is held up as one basis for judging quality. The success of the exemplars of the approaches reviewed here both in publication and securing funding speaks to the value these scholars do place on peer review. For advocates of this family of approaches, however, peer review is insufficient. Quality is at least partly determined by judgments of the value to stakeholders, because of the commitment to collaboration and to create tools and products that are useful in addressing significant educational problems. Whether a project was successful depends in part on the evaluations of and uptake of ideas and findings by the stakeholders.

We need not presume that education stakeholders de-value qualities that researchers assert makes research trustworthy. For example, studies of education leaders who have regular interactions with researchers show they value high-quality research syntheses that have undergone peer review (e.g., Hopkins, Weddle, Gluckman, & Gautsch, 2019). At the same time, research suggests that educator partners might value some traditional criteria over others. For example, research suggests that educators prize research on populations "like theirs," and thus take threats to external validity as seriously as threats to internal validity (Coburn, Honig, & Stein, 2009; Coburn & Talbert, 2006).

Saying that quality is determined both by peers and stakeholders in research leads to many open questions for this family of approaches to address. What happens if a study meets the approval of peer reviewers, but is judged to be of limited use to collaborators in schools and communities? What happens when different stakeholders do not agree as to the relative strengths and weaknesses of a study's design or conclusions? What should be the role of researchers in educating partners about the relative merits of alternative designs? And what should be the role

of partners in educating researchers about the context? In this form of research, stakeholders' goals for research are likely to arise as partners negotiate the design of a study, but the success of the endeavor requires that the research be carried out in a timely manner, such that findings are still relevant. In other words, the conditions of systems cannot be so volatile that no matter what research is conducted, it is likely to be irrelevant by the time it is completed.

The project's problem focus or foci must be warranted. While all research is focused on and designed to answer a particular question, this family of approaches' stance about grounding inquiry in specific problems demands that researchers give an account of the problem from different stakeholder perspectives. It is not sufficient to define the problem in relation to existing bodies of research; some evidence must be presented that the problem is of importance to the key stakeholders involved. Though it is possible that through careful problem definition, collaborators arrive at a common understanding of the problem, it is just as likely that they have different conceptions of it (Penuel, Coburn, & Gallagher, 2013). It is also typical for multiple problems to be taken up simultaneously (Quartz et al., 2017). High quality research within this family of approaches might include a description of how problems were determined, by whom, and what differences in perspectives remained or emerged throughout the life of a project.

A challenge in this approach is to articulate a problem that is worthwhile to address through research. Some problems may already have answers in the existing literature and not merit additional study. Other problems as presented do not get at underlying issues or take as "given" institutional structures and systemic inequities that should be challenged. The disciplined approach taken by improvement science in this regard provides one set of methodologies for ensuring that problem definition takes into account what is known already, the perspectives of different actors, and perceptions of root causes. But needed also is a *critical* pragmatic stance

toward problems, that is, where the operation of power in situations of inequality can generate commonsense notions of "problems" that themselves need to be challenged (Feinberg, 2015).

Research designs must submit change strategies to systematic testing. The four approaches share with all other approaches to education research a commitment to systematic inquiry: developing plans for studying problems, using systematic methods to design, test, and iterate solutions, and both constructing and warranting new knowledge. A commitment to systematic inquiry indicates that advocates of these approaches embrace uncertainty in the scientific enterprise, with an openness to what they might find, with knowledge of the possibility that their ideas or proposals for change might prove disappointing. Attending carefully to relationships with participants adds something extra to the requirements for high quality research. It often requires slowing down the research process to develop relationships and come to understand the context before embarking on a search for solutions (Gutiérrez & Jurow, 2016).

The kind of engagement promoted in these approaches makes it especially challenging to avoid charges of bias and to present findings that may put participants' goals or actions in an unfavorable light (Kirshner, 2010; Nzinga et al., 2018). A number of partnerships have devised formal strategies to support researchers in delivering "bad news" to partners, and such strategies might need to become more formalized and widespread (Connolly, Plank, & Rone, 2012). A central challenge, then, is to balance the "organized skepticism" of science with a position of active engagement and accountability to stakeholders.

As noted above, the pragmatic stance of this family of approaches makes it a methodologically pluralistic one for which a single standard of inquiry cannot therefore be applied. As the review indicates, this family of approaches encompasses mixed methodologies, ethnography as well as experiments, and different kinds of studies that fit the question at hand.

The pluralism reflects the fact that the kinds of research questions posed are not solely of the "what works" variety; instead, questions that relate to matters of importance to partners may be descriptive or contextual in nature, or pertain to implementation (see also, Conaway, Keesler, & Schwartz, 2015). This presents a challenge to advocates of the approach, who must often develop skill in multiple methodologies to address the questions arising from joint work.

Multiple qualities and multiple kinds of research products should be evaluated. Unlike, RDDU research, *what* is to be evaluated for quality extends beyond the quality of a research study; it also necessarily encompasses the quality of other products of research. Particularly important are judgments about the strategies, tools, materials, or routines that are designed to improve learning. The potential impact of such strategies for others adopting them is of course important, but so is an assessment of their utility and perceived value to stakeholder groups. Also, potentially important are ideas or concepts produced by partnerships, which can move from projects into systems in ways that shift how people think about an issue, with consequences for polices (Farrell, Coburn, & Chong, 2019). Or, experiences within a research collaboration may create opportunities for participants to take on new methods of inquiry in their day-to-day lives, so that they could see themselves as knowledge producers (Cammarota & Fine, 2008).

As with the first criterion of quality, this criterion raises questions about who evaluates such qualities and how. While there exist independent review processes for curriculum resources in many states, there are not similar processes in place for reviewing professional development strategies or tools for participatory design. Just who might be equipped to review these products of research, and how standards might be set up to review them, is an open question. This leads us to consider next a bigger question, related to infrastructures that might be necessary to advance this set of research approaches as a group.

New Infrastructures Needed to Cultivate Collaborative Approaches to Research

If collaborative approaches were simply "new wine in old bottles" or "nothing but" traditional research "done well," then we might imagine its development needing to unfold in a fashion similar to how methods or subfields have developed in the past, namely by re-purposing existing institutional forms. We might push for clearer definitions and distinctions among approaches. We might re-design the content of graduate classes, create a new journal, hold conferences and create networks dedicated to collaborative research, and perhaps build new program areas and departments within Schools of Education dedicated to producing scholars who are prepared to engage in collaborative research. No doubt, such steps are necessary because few educational researchers have the skills necessary both for conducting rigorous research and engaging different publics productively. And, some steps, like changing existing professional societies (e.g., the Society for Research on Educational Effectiveness), have already begun and now reflect multiple perspectives on research quality.

Ultimately, we suspect such efforts will fall short, however, in helping to support and sustain this family of approaches to research. Instead, we propose that what could bring the advocates of these approaches together successfully is their implicit call for more re-organization of the research and development enterprise. The *collaborative* dimension challenges dominant approaches that emphasize the need for research to be in service either to the research community *or* to the immediate demands of practice, as well as calling for new forms of deliberation over the purposes and trajectories of research. The *problem*-focused dimension demands new forms of accountability of research to educators, families, and the community, not just to peer reviewers from the research community. The search for *solutions* demands new warrants for the value of the research that is conducted. And to judge the quality of *research*

requires the participation and engagement of new kinds of stakeholders and accounts of the process of research that traditional journal articles do not demand. On these matters, our participants agreed on the need for new arrangements between research and practice, ones that will require significant effort on the part of researchers, educators, and community members to overcome cultural and institutional barriers to joint work.

One way that the current infrastructure will need to change is in the media and methods with which we communicate our research. With respect to journals, if we create new ones, they will need to be Open Access, so that anyone (but especially participants and stakeholders) can access the research accounts we produce. At present, an additional limitation of journals is space: no single article can support giving an account of how each of the shared values we have identified was realized within a project. And we will need to engage in different forms of writing and speaking, from blogging to developing brief presentations of key findings that can engage policy makers, local decision makers, parents and guardians, as well as additional community members. These forms of communication need not be viewed as completely separate from traditional academic writing, either, but as complementary and as informing one another (Rose, 2018). Examples of these forms already exist, both within popular press blogs that are focused on linking research and practice (Hill & Loeb, 2020) and also in the form of collaboratively produced resources that support changes to practice (Morrison & Bell, 2018).

In addition, institutional incentives of researchers and other co-investigators will need to become better aligned to the goals of this kind of research. Developing public scholarship among university scholars requires a reframing of academic work as a unity of teaching, research, and service (Kellogg Commission on the Future of State Land-Grant Universities & National Association of State Universities & Land-Grant Colleges, 1999). This demands a realignment of

incentives linked to tenure and promotion of faculty (Fischman, Anderson, Tefera, & Zuiker, 2018). It also demands pathways for researchers outside universities or in research staff tracks, for educators who might be engaged in co-design with researchers, for administrators collecting practical measures data as part of a Plan-Do-Study-Act cycle, and for family and community members who are directly supporting implementation of a new initiative. There are already examples of such pathways, such as within Regional Education Laboratories in the United States, though these pathways are not always valued by those in the academy. In short, it demands a redistribution of labor across the worlds of research, practice, policy, and community.

A key reason to invest in new forms of infrastructure for research and development is that the current infrastructure limits the potential value of education research for practice. It incentivizes and produces primarily studies of the impact of programs and policies, which is only a small proportion of research that education leaders find valuable. Educational leaders are far more likely to turn to research that can help them design professional development or to monitor implementation of initiatives than to research that helps them select programs (Penuel, Briggs, et al., 2017; Penuel, Farrell, Allen, Toyama, & Coburn, 2018). In general, leaders are more likely to ask questions about how to make programs work for all students, classrooms, and schools, than questions about whether a program works or not (Means & Penuel, 2005). As a result, a research and development infrastructure that prizes impact research over continuous improvement limits the likelihood that research will be relevant to education leaders (Peurach, 2016).

A second limitation of the current infrastructure is that it prizes researchers' ideas for what will improve education over all other education stakeholders' ideas. The common guidelines produced by the Institute of Education Sciences and National Science Foundation require no evidence from students, educators, families, or community leaders that the problem a proposed

innovation addresses is an important one to them (Gutiérrez & Penuel, 2014). Further, there are no requirements for educators to be involved in helping design innovations that researchers are testing in early stage research or that researchers consider how the educational systems in which they are embedded are likely to affect implementation. In fact, from the standpoint of the current infrastructure, the best innovation is one that will be effective anywhere no matter who implements it. The pursuit of this ideal disregards a large body of implementation research, which suggests that educator ownership is critical to develop early and to successfully scale and sustain innovations (Datnow, Hubbard, & Mehan 2002).

One consequence of excluding the voices of educators and community members from deciding the focus of research is that the topics investigated are not necessarily relevant to stakeholders, particularly those historically marginalized in setting directions for policies and programs (Tuck & Yang, 2014). The focus of programs and outcomes to be assessed may be narrow and privilege purposes for education valued by the most powerful groups in American society. The diversification of voices that shape the production of research can center different purposes and lead to important new discoveries about learning and how to support it (Medin & Bang, 2014). As with all research, there must be a coherent, explicit chain of reasoning that can be evaluated, and the coherence of the kinds of collaborative education research we have explored here depends on an expanded set of enacted commitments that emphasize the interconnections among rigor, relevance, and agency among research participants and audiences.

Limitations and Qualifications

The principles we have highlighted here are common to a limited set of collaborative approaches to research and development. They are not static or universal, nor are they unaffected by the positionality of the authors, one of whom is a developer and advocate for one of the

approaches reviewed. Indeed, a limitation of this review is that we have not undertaken a detailed analysis of who the participants and advocates of these approaches are, their positionality not only as scholars, but within larger networks of power. Nor have we explored in detail the way that collaboration is inevitably raced, classed, and gendered within these approaches to research. As noted above, this is an important limitation of our review, because it could easily lead one to conclude that there are no important distinctions among these approaches that are linked to differences in *who* is engaged in this research, and there is ample evidence from different scientific fields that "who is asking" matters (Medin & Bang, 2014).

Our approach to identifying family resemblances contributes directly to this limitation. It produces overlapping similarities, without identifying distinctions that matter to those who are advocates for the approach. A different approach might be needed to help specify these, one grounded more in the identifying of *contrasts,* rather than of overlapping and interconnected features, so that novel meanings might be derived (Marton & Pang, 2013).

At present, there are multiple, parallel efforts underway to name some of the overlapping spaces in which we locate ourselves as authors and actors who engage in collaborative forms of research (e.g., York, Valladares, Valladares, Garcia, & Snyder, 2020). As others continue to develop and articulate what is distinctive about these approaches and their premises, they will choose to highlight different principles than we have here. Our primary claim to the trustworthiness of our typology of principles is based on the participatory nature of the process we used to derive the principles and on the systematicity with which we have analyzed exemplar accounts and subjected them to our colleagues' review.

Now that the overlapping and interconnected principles have been identified, moreover, we are in a better position to evaluate whether some additional approaches should have been

included but that were not. Some forms of teacher inquiry, such as lesson study, might be considered a member of the family, when they map onto methods of one of the approaches (e.g., Lewis, 2015). Similarly, teacher inquiry instituted at the systems level could also count as a member of this family. As with some of the models reviewed, the question of "what scales," however, needs to be considered, because it may not be abstract representations of practice generated through formal communication of research findings (Oliver et al., 2018).

Our attempt to represent a particular family of approaches is not intended to discredit other approaches to research that fall outside its scope. We are particularly concerned to preserve a space for inquiry that seeks to develop new possibilities for learning and organizational change at the edge of what is now considered feasible to implement in educational systems today. Such projects in "social dreaming" (Gutiérrez, 2005) are essential for breaking through the impasses created by inequitable institutions and creating new arrangements among educational organizations in communities. At the same time, we think existing RDDU research could be strengthened with a greater support for researchers building collaborative relationships with the youth, families, communities, and educators in schools and district offices who are participants in their research. In our view, the benefits in terms of enhancing the potential relevance of research, expanding our understanding of problems facing children and the strengths that they themselves bring to addressing them, and creating new possibilities for change outweigh the risks associated with collaborative engagement. There is already room within the current guidelines for federal research, as well as within the Institute of Education Sciences' mission, to consider how relevance to practice might be made into a criterion or set of criteria for defining quality in research (Gutiérrez & Penuel, 2014).

It will not be enough to change education research if educational systems do not also change. Such systems are turbulent and characterized by high turnover, which threaten long-term collaborative research efforts (Peurach & Glazer, 2012). Although turnover is not always fatal to collaborative research, it can be, especially when it occurs in an already unstable environment (Farrell et al., 2018). To improve systems at scale requires qualitatively different support for leader learning and for creating a culture of inquiry (Peurach, Penuel, & Russell, 2019). It may also require deeper engagement with the wider ecosystem of organizations that influence school improvement, from policymakers to publishing companies to advocacy groups of different kinds.

As a final limitation, we consider some limitations within the studies reviewed. Beyond some shared pragmatic commitments, key terms used across the different approaches take on different meanings across the traditions that speak to some potential differences among them. "Practice" is one such term. Rarely defined in the scholarship within this group of approaches, the term often refers to classroom practice or more broadly to what takes place in educational systems as opposed to in research ("the worlds of research and practice"; National Research Council, 2003, p. 1). The term is also used to refer to "disciplinary and cultural practices" (e.g., Booker & Goldman, 2016). Sometimes, it is used to characterize a purpose of research, as in to "develop new tools and practices that produce new learning arrangements" (Fishman, Penuel, Allen, Cheng, et al., 2013, p. 141). Agency is another such term, which may sometimes be interpreted to be related to the "authority to act" to change one's practice, but may also refer to the actions of a collective to create new forms of activity (e.g., Severance et al., 2016).

No doubt, advancing research will require a more careful elaboration of the meanings of these key terms; here, we seek through case study to elaborate on their practical significance for the unfolding of projects as described in studies. Drawing on rich definitions of the term

"practice" offered in the science studies literature (e.g., Pickering, 1995), as well as discussions of agency within sociology (e.g., Emirbayer & Mische, 1998) and cultural-historical activity theory (e.g., Virkkunen, 2006), would be useful places to start. Tighter links between theory and methodology could serve the development of this family of approaches to research, in that it could support knowledge development about the approaches themselves (Engeström, Sannino, & Virkkunen, 2014). This search for coherence sits in tension with the pragmatism of many of these approaches, which employ theories and methods that match more local demands and goals.

Conclusion

As many fellow education researchers before us have (e.g., Lagemann, 2002), we reiterate the call for a practice of educational research that is in closer dialogue with policy and practice. We agree with others that for too long, education research has had a limited impact on policy and practice. However, we see the need for a change in both language and practice that reflects better the value of respecting the agency and expertise of participants in research articulated here. These approaches call for a more collaborative engagement of research with policy and practice (Penuel & Spillane, 2014), in which impact is reconceptualized as a "two-way street" where policy, practice, and research influence and change one another.

The practices of research that embody the values shared among this family of approaches to education research cannot survive without new infrastructural elements to support their enactment. There are far too many incentives to continue "business as usual" within the RDDU model in ways that make conducting collaborative research difficult, such as the incentives for investigator-driven ideas for intervention. In this respect, the exercise of constituting and analyzing them as a family of approaches is more than just a conceptual effort to identify points

of commonalty. It is also an activity to identify points of solidarity around which members of the family might come together to advocate for and build such infrastructures.

Such advocacy and field-building efforts will require allies beyond this family of research approaches. There are strong points of intersection in this respect with recent calls for publicly engaged scholarship (Fischman, Anderson, Tefera, & Zuiker, 2018; Giles, 2016; Kellogg Commission on the Future of State Land-Grant Universities & National Association of State Universities & Land-Grant Colleges, 1999; Oakes, 2018). For example, Oakes (2018) recently called for "embedding research in educational contexts, with the dual goal of partnering with practitioners to improve practice and producing knowledge that is viable beyond that particular context" (p. 101). This call resonates strongly with this family of approaches' call for collaborative engagement between research and practice. There are also strong resonances with recent calls for democratizing evidence in policy making (Tseng, Fleishman, & Quintero, 2018). Here, the effort is centered on how to construct more "two-way streets" (Tseng et al., 2017) between research on the one hand and policy and practice on the other, for the purpose of promoting evidence-based decision making. Democratizing the movement to improve the use of evidence depends on people in communities and schools having a say in what questions are asked and what research is conducted.

Both in finding allies and in building new research infrastructures, our own project has reinforced for us the centrality of *principles* in the endeavor to improve education research. We are by no means the first to point this out (e.g., House & Howe, 1999), but what unites these approaches are a set of commitments to engagement that, from the perspective of insiders to the approaches, set them apart from "traditional" research. First and foremost, these collaborative approaches are efforts to humanize and democratize the field of education research.

References

*Articles referenced as representative of the focal projects of this review.

American Education Research Association. (2009). Standards for reporting on humanities-oriented research in AERA publications. *Educational Researcher, 38*(6), 481-486. https://doi.org/10.3102/0013189X09341833

Bailey, T., Jeong, D. W., & Cho, S.-W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review, 29*(2), 255-270.

Bang, M., Faber, L., Gurneau, J., Marin, A., & Soto, C. (2016). Community-Based design research: Learning across generations and strategic transformations of institutional relations toward axiological innovations. *Mind, Culture, and Activity, 23*(6), 28-41. https://doi.org/10.1080/10749039.2015.1087572

Bang, M., Medin, D., Washinawatok, K., & Chapman, S. (2010). Innovations in culturally based science education through partnerships and community. In M. S. Khine & M. I. Saleh (Eds.), *New science of learning: Cognition, computers, and collaboration in education* (pp. 569-592). New York, NY: Springer.

Bang, M., & Vossoughi, S. (2016). Participatory design research and educational justice: Studying learning and relations within social change making. *Cognition and Instruction, 34*(3), 173-193. https://doi.org/10.1080/07370008.2016.1181879

Berwick, D. M. (2008). The science of improvement. *The Journal of the American Medical Association, 299*(10), 1182-1184. http://doi:10.1001/jama.299.10.1182

Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the deficit in educational research. *Educational Theory, 57*(1), 1-22. https://doi.org/10.1111/j.1741-5446.2006.00241.x.

*Booker, A., & Goldman, S. V. (2016). Participatory design research as a practice for systemic repair: Doing hand-in-hand math research with families. *Cognition and Instruction, 34*(3), 223-235. https://doi.org/10.1080/07370008.2016.1179535

Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. Hallinan (Ed.), *Frontiers in sociology of education* (pp. 127-162). Dordrecht, the Netherlands: Verlag.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Cambridge, MA: Harvard Press.

Cammarota, J., & Fine, M. (2008). *Revolutionizing education: Youth participatory action research in motion*. New York: Routledge.

Campano, G., Ghiso, M. P., & Welch, B. (2015). Ethical and professional norms in community-based research. *Harvard Educational Review, 85*(1), 29-49. https://doi.org/10.17763/haer.85.1.a34748522021115m

Campano, G., Ghiso, M. P., & Welch, B. J. (2016). *Partnering with immigrant communities: Action through literacy*. New York, NY: Teachers College Press.

Ching, D., Santo, R., Hoadley, C., & Peppler, K. A. (2016). Not just a blip in someone's life: Integrating brokering practices into out-of-school programming as a means of supporting and expanding youth futures. *On the Horizon, 24*(3), 296-312.

Clements, D. H. (2007). Curriculum research: Toward a framework for "research-based curricula". *Journal for Research in Mathematics Education, 38*(1), 35-70. DOI: 10.2307/30034927.

Cobb, P. A., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences, 10*(1&2), 113-163. https://doi.org/10.1207/S15327809JLS10-1-2_6.

Coburn, C. E., Honig, M. I., & Stein, M. K. (2009). What's the evidence on districts' use of evidence? In J. D. Bransford, D. J. Stipek, N. J. Vye, L. M. Gomez, & D. Lam (Eds.), *The role of research in educational improvement* (pp. 67-87). Cambridge, MA: Harvard Education Press.

Coburn, C. E., & Penuel, W. R. (2016). Research-practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher, 45*(1), 48-54. https://doi.org/10.3102/0013189X16631750.

Coburn, C. E., & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education, 112*, 469-495. DOI: 10.1086/505056

Cochran-Smith, M., & Lytle, S. (1993). *Inside/outside: Teacher research and knowledge*. New York, NY: Teachers College Press.

Cochran-Smith, M., & Lytle, S. (2009). *Inquiry as stance: Practitioner research for the next generation*. New York, NY: Teachers College Press.

Cole, M., & Packer, M. J. (2016). Design-based intervention research as the science of the doubly artificial. *Journal of the Learning Sciences, 25*(4), 503-530. https://doi.org/10.1080/10508406.2016.1187148.

Conaway, C., Keesler, V. A., & Schwartz, N. L. (2015). What research do State Education Agencies really need? The promise and limitations of state longitudinal data systems. *Educational Evaluation and Policy Analysis, 37*(1S), 16S-28S. https://doi.org/10.3102/0162373715576073.

Connolly, F., Plank, S., & Rone, T. (2012). *Baltimore Education Research Consortium: A consideration of past, present, and future*. Baltimore, MD: Author.

Datnow, A., Hubbard, L., & Mehan, H. (2002). *Extending educational reform: From one school to many*. New York: Routledge/Falmer.

DeBarger, A. H., Choppin, J. M., Beauvineau, Y., & Moorthy, S. (2013). Designing for productive adaptations of curriculum interventions. In B. J. Fishman, W. R. Penuel, A.-R. Allen, & B. H. Cheng (Eds). *Design-based implementation research. National Society for the Study of Education Yearbook, 112*(2), 298-319.

Deming, W. E. (1993). *The new economics*. Cambridge, MA: Harvard University Press.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5-8.

Donovan, M. S. (2013). Generating improvement through research and development in educational systems. *Science, 340*, 317-319. DOI: 10.1126/science.1236180

*Donovan, M. S., & Snow, C. E. (2018). Sustaining research-practice partnerships: Benefits and challenges of a long-term research and development agenda. In B. Bevan & W. R. Penuel (Eds.), *Connecting research and practice: New models for equity and ethics* (pp. 33-50). New York, NY: Routledge.

*Donovan, M. S., Snow, C. E., & Daro, P. (2013). The SERP approach to problem-solving research, development, and implementation. In B. J. Fishman, W. R. Penuel, A.-R. Allen,

& B. H. Cheng (Eds). *Design-based implementation research. National Society for the Study of Education Yearbook, 112*(1), 400-425.

Dynarski, M. (2008). Bringing answers to educators: Guiding principles for research syntheses. *Educational Researcher, 37*(1), 27-29.

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *The Journal of the Learning Sciences, 11*(1), 105-121. https://doi.org/10.1207/S15327809JLS1101_4

Ehn, P. (1992). Scandinavian design: On participation and skill. In P. S. Adler & T. A. Winograd (Eds.), *Usability: Turning technologies into tools* (pp. 96-132). New York: Oxford University Press.

Elmore, R. F., & Forman, M. L. (2011). *Building coherence within schools*. Boston, MA: SERP.

Emirbayer, M., & Mische, A. (1998). What is agency? *The American Journal of Sociology, 103*(4), 962-1023. https://doi.org/10.1086/231294

Engelbart, D. C. (1992, August). *Toward high-performance organizations: A strategic role for groupware.* Paper presented at the GroupWare '92 Conference, San Jose, CA.

Engeström, Y., Sannino, A., & Virkkunen, J. (2014). On the methodological demands of formative interventions. *Mind, Culture, and Activity, 21*(2), 118-128. https://doi.org/10.1080/10749039.2014.891868

Esmonde, I., Blair, K., Goldman, S. V., Martin, L., Jimenez, O., & Pea, R. D. (2012). Math I Am: What we learn from stories that people tell about math in their lives. In B. Bevan, P. Bell, R. Stevens, & A. Razfar (Eds.), *LOST opportunities: Learning in out-of-school time* (pp. 7-28). New York: Springer.

Every Student Succeeds Act, 20, Pub. L. No. 114-95 (2015).

Farrell, C. C., Coburn, C. E., & Chong, S. (2019). Under what conditions do school districts learn from external partners? The role of absorptive capacity. *American Educational Research Journal, 56*(3), 955-994. https://doi.org/10.3102/0002831218808219

Farrell, C. C., Davidson, K. L., Repko-Erwin, M., Penuel, W. R., Quantz, M., Wong, H., Riedy, R., & Brink, Z. (2018). *A descriptive study of the IES Researcher–Practitioner Partnerships in Education Research program*. Boulder, CO: National Center for Research in Policy and Practice. Retrieved May 23, 2020 from http://ncrpp.org/assets/documents/NCRPP-Technical-Report-No-3_Full-Report.pdf.

Feinberg, W. (2015). Critical pragmatism and the appropriation of ethnography by philosophy of education. *Studies in Philosophy of Education, 34*(2), 149–157. DOI: 10.1007/s11217-014-9415-6

Fischman, G. E., Anderson, K. T., Tefera, A. E., & Zuiker, S. J. (2018). If mobilizing educational research Is the answer, who can afford to ask the question? An analysis of faculty perspectives on knowledge mobilization for scholarship in education. *AERA Open, 4*(1), 1-17. https://doi.org/10.1177/2332858417750133

Fishman, B. J., & Penuel, W. R. (2018). Design-based implementation research. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (pp. 393-400). New York, NY: Routledge.

Fishman, B. J., Penuel, W. R., Allen, A.-R., & Cheng, B. H. (Eds.). (2013). *Design-based implementation research: Theories, methods, and exemplars. National Society for the Study of Education Yearbook.* New York, NY: Teachers College Press.

Fishman, B. J., Penuel, W. R., Allen, A.-R., Cheng, B. H., & Sabelli, N. (2013). Design-Based Implementation Research: An emerging model for transforming the relationship of

research and practice. In B. J. Fishman, W. R. Penuel, A.-R. Allen, & B. H. Cheng (Eds). *Design-based implementation research. National Society for the Study of Education Yearbook, 112*(2), 136-156.

Frank, K. A., Zhao, Y., & Borman, K. (2004). Social capital and the diffusion of innovations within organizations: Application to the implementation of computer technology in schools. *Sociology of Education, 77*(2), 148-171. https://doi.org/10.1177/003804070407700203

Giles, D. E. (2016). Understanding an emerging field of scholarship: Toward a research agenda for engaged, public scholarship. *Journal of Higher Education Outreach and Engagement, 20*(1), 181-191.

Giroux, H. A. (1986). Critical theory and the politics of culture and voice: Rethinking the discourse of educational research. *Journal of Thought, 21*(3), 84-105.

*Goldman, S. V., & Booker, A. (2009). Making math a definition of the situation: Families as sites for mathematical practices. *Anthropology and Education Quarterly, 40*(4), 369-387. https://doi.org/10.1111/j.1548-1492.2009.01057.x

Gutiérrez, K. D., & Jurow, A. S. (2016). Social design experiments: Toward equity by design. *Journal of the Learning Sciences, 25*(4), 565-598. https://doi.org/10.1080/10508406.2016.1204548

Gutiérrez, K. D., & Penuel, W. R. (2014). Relevance to practice as a criterion for rigor. *Educational Researcher, 43*(1), 19-23. https://doi.org/10.3102/0013189X13520289

Gutiérrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher, 32*(5), 19-25. https://doi.org/10.3102/0013189X032005019

Hannan, M., Russell, J. L., Takahashi, S., & Park, S. (2015). Using improvement science to better support beginning teachers: The case of the Building a Teaching Effectiveness Network. *Journal of Teacher Education, 66*(5), 494-508. https://doi.org/10.1177/0022487115602126

Haskins, R., & Margolis, G. (2015). *Show me the evidence: Obama's fight for rigor and results in social policy*. Washington, DC: Brookings Institution.

Henrick, E. C., Cobb, P., Penuel, W. R., Jackson, K., & Clark, T. R. (2017). *Assessing research-practice partnerships: Five dimensions of effectiveness*. New York, NY: William T. Grant Foundation. https://wtgrantfoundation.org/library/uploads/2017/10/Assessing-Research-Practice-Partnerships.pdf

Henrick, E. C., Klafehn, A., & Cobb, P. A. (2018). Assessing the impact of partnership recommendations on district instructional improvement strategies. In P. Cobb, K. Jackson, E. Henrick, T. M. Smith, & t. M. Team (Eds.), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 209-220). Cambridge, MA: Harvard Education Press.

Hill, H. C., & Loeb, S. (2020, January 22). Introducing: A new series on the practical takeaways from research. *Education Week*. Retrieved from https://www.edweek.org/ew/articles/2020/01/22/introducing-a-new-series-on-the-practical.html

Honig, M. I., & Coburn, C. E. (2008). Evidence-based decision making in school district central offices: Toward a policy research agenda. *Educational Policy, 22*(4), 578-608. https://doi.org/10.1177/0895904807307067

Hopkins, M., Weddle, H., Gluckman, M., & Gautsch, L. (2019). Boundary crossing in a professional association: The dynamics of research use among state leaders and researchers in a research-practice partnership. *AERA Open*. https://doi.org/10.1177/2332858419891964

House, E. R., & Howe, K. R. (1999). *Values in evaluation and social research*. Thousand Oaks, CA: Sage.

Ikemoto, G. S., & Honig, M. I. (2010). Tools to deepen practitioners' engagement with research: The case of the Institute for Learning. In C. E. Coburn & M. I. Honig (Eds.), *Research and practice in education: Building alliances, bridging the divide* (pp. 93-108). Lanham, MD: Rowman & Littlefield.

Institute of Education Sciences. (2012). *National Board for Education Sciences Meeting, 2011-2012*. Washington, DC: U.S. Department of Education. Retrieved from: https://ies.ed.gov/director/board/pdf/20126003.pdf.

Institute of Education Sciences, & National Science Foundation. (2013). *Common guidelines for education research and development*. Washington, DC: U.S. Department of Education. Retrieved from: https://www.nsf.gov/pubs/2013/nsf13126/nsf13126.pdf.

Ishikawa, K. (1968). *Guide to quality control*. Tokyo, Japan: JUSE Press, Ltd.

Ishimaru, A. M., & Takahashi, S. (2017). Disrupting racialized institutional scripts: Toward parent-teacher transformative agency for educational justice. *Peabody Journal of Education, 92*(3), 343-362. https://doi.org/10.1080/0161956X.2017.1324660

Jay, T., Rose, J., & Simmons, B. (2017). Finding 'mathematics': Parents questioning school-centered approaches to involvement in children's mathematics learning. *School Community Journal, 27*(1), 201-230.

Kali, Y., Eylon, B.-S., McKenney, S. E., & Kidron, A. (2018). Design-centric research-practice partnerships: Three key lenses for building productive bridges between theory and practice. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, design, and technology: An international compendium of theory, research, practice, and policy* (pp. 1-30). Basel, Switzerland: SpringerNature.

Kellogg Commission on the Future of State Land-Grant Universities, & National Association of State Universities & Land-Grant Colleges. (1999). *Returning to our roots: The engaged institution* (Vol. 3). Washington, DC: National Association of State Universities and Land-Grant Colleges, Office of Public Affairs.

Kirshner, B. (2010). Productive tensions in youth participatory action research. In W. R. Penuel & K. O'Connor (Eds). *Learning research as a human science. National Society for the Study of Education Yearbook, 109*(1), 238-251.

Koninklijke Nederlandse Akademie van Wetenschappen. (2018). *Maatschappelijke impact in kaart*. Amsterdam, the Netherlands: Author.

Lagemann, E. C. (2002). *An elusive science: The troubling history of education research*. Chicago, IL: University of Chicago Press.

Le Dantec, C. A., & Fox, S. (2015). Strangers at the gate: Gaining access, building rapport, and co-constructing community-based research. In D. Cosley, A. Forte, L. Ciolfi, & D. McDonald (Eds.), *CSCW '15: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 1348-1358). Vancouver, BC, Canada: Association for Computing Machinery

Legg, C., & Hookway, C. (2019). Pragmatism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Palo Alto, CA: Stanford University Press.

LeMahieu, P., Grunow, A., Baker, L., Nordstrum, L. E., & Gomez, L. (2017). Networked improvement communities: The discipline of improvement science meets the power of networks. *Quality Assurance, 25*(1), 5-25.

Lewis, C. (2015). What is improvement science? Do we need it? *Educational Researcher, 44*(1), 54-61. https://doi.org/10.3102/0013189X15570388

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage.

Males, L. M., Sears, R., & Lawler, B. R. (2020). Equity and justice in the preparation of secondary mathematics teachers. In W. G. Martin, B. R. Lawler, A. E. Lischka, & W. M. Smith (Eds.), *The Mathematics Teacher Education Partnership: The power of a Networked Improvement Community to transform mathematics teacher preparation* (pp. 57-90). Charlotte, NC: Information Age Publishing.

Marton, F., & Pang, M. F. (2013). Meanings are acquired from experiencing differences against a background of sameness, rather than from experiencing sameness against a background of difference: Putting a conjecture to the test by embedding it in a pedagogical tool. *Frontline Learning Research, 1*(1), 24-41. https://doi.org/10.14786/flr.v1i1.16

Medin, D., & Bang, M. (2014). *Who's asking? Native science, Western science, and science education*. Cambridge, MA: MIT Press.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago, IL: University of Chicago Press.

Morrison, D., & Bell, P. (2018, December). *Equity and STEM pedagogy: Explorations of STEM teaching tools that link research and practice*. Paper presented at the American Geophysical Union, Fall Meeting, San Francisco, CA.

Munter, C., Stein, M. K., & Smith, M. S. (2015). Dialogic and direct instruction: Two distinct models of mathematics instruction and the debate(s) surrounding them. *Teachers College Record, 117*(11), 1-32. http://www.tcrecord.org/Content.asp?ContentId=18115

Nasir, N. S., Rosebery, A., Warren, B., & Lee, C. D. (2014). Learning as a cultural process: Achieving equity through diversity. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 686-706). New York, NY: Cambridge University Press.

National Academy of Education. (1999). *Recommendations regarding research priorities: An advisory report to the National Educational Research Policy and Priorities Board*. Washington, DC: National Academy of Education.

National Research Council. (2002). *Scientific research in education*. Washington, DC: National Academy Press. https://www.nap.edu/read/10236/chapter/1

National Research Council. (2003). *Strategic education research partnership*. Washington, DC: National Academies Press. https://www.nap.edu/read/10670/chapter/1

National Research Council. (2012). *Using science as evidence in public policy*. Washington, DC: The National Academies Press. https://www.nap.edu/read/13460/chapter/1

Nzinga, K., Rapp, D. N., Leatherwood, C., Easterday, M., Rogers, L. O., Gallagher, N., & Medin, D. L. (2018). Should social scientists be distanced from or engaged with the people they study? *Proceedings of the National Academy of Sciences, 115*(45), 11435-11441. https://doi.org/10.1073/pnas.1721167115

Oakes, J. (2018). Public scholarship: Education research for a diverse democracy. *Educational Researcher, 47*(2), 91-104. https://doi.org/10.3102/0013189X17746402

O'Neill, D. K. (2016). Understanding design research–practice partnerships in context and time: Why learning sciences scholars should learn from cultural-historical activity theory

approaches to design-based research. *Journal of the Learning Sciences, 25*(4), 497-502. http://doi:10.1080/10508406.2016.1226835

Oliver, M., Avramides, K., Clark, W., Hunter, J., Luckin, R., Hansen, C., & Wasson, B. (2018). Sharing teacher knowledge at scale: teacher inquiry, learning design and the representation of teachers' practice. *Teacher Development, 22*(4), 587-606. https://doi.org/10.1080/13664530.2017.1381642

Pea, R. D., & Martin, L. M. (2010). Values that occasion and guide mathematics in the family. In W. R. Penuel & K. O'Connor (Eds). *Learning research as a human science. National Society for the Study of Education Yearbook, 109*(1), 34-52.

Penuel, W. R. (2019). Infrastructuring as a practice of design-based research for supporting and studying equitable implementation and sustainability of innovations. *Journal of the Learning Sciences, 28*(4-5), 659-677. https://doi.org/10.1080/10508406.2018.1552151

Penuel, W. R., Coburn, C. E., & Gallagher, D. (2013). Negotiating problems of practice in research-practice partnerships focused on design. In B. J. Fishman, W. R. Penuel, A.-R. Allen, & B. H. Cheng (Eds.), *Design-based implementation research: Theories, methods, and exemplars. National Society for the Study of Education Yearbook.* (pp. 237-255). New York, NY: Teachers College Record.

Penuel, W. R., & Gallagher, D. (2017). *Creating research-practice partnerships in education*. Cambridge, MA: Harvard Education Press.

Penuel, W. R., Peurach, D. J., LeBoeuf, W. A., Riedy, R., Barber, M., Clark, T. R., & Gabriele, K. (2017). *Defining Collaborative Problem Solving Research: Common values and distinctive approaches.* Boulder, CO: University of Colorado.

Penuel, W. R., & Spillane, J. P. (2014). Learning sciences and policy design and implementation: Key concepts and tools for collaborative engagement. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 649-667). Cambridge, UK: Cambridge University Press.

Peurach, D. J. (2016). Innovating at the nexus of impact and improvement: Leading educational improvement networks. *Educational Researcher, 45*(7), 421-429. https://doi.org/10.3102/0013189X16670898

Peurach, D. J., & Glazer, J. L. (2012). Reconsidering replication: New perspectives on large-scale school improvement. *Journal of Educational Change, 13*(2), 155–190. https://doi.org/10.1007/s10833-011-9177-7

Peurach, D. J., Penuel, W. R., & Russell, J. L. (2019). Beyond ritualized rationality: Organizational dynamics of instructionally-focused continuous improvement. In M. Connolly, D. H. Eddy Spicer, C. James, & S. D. Kruse (Eds.), *SAGE International Handbook of School Organization* (pp. 465-488). Thousand Oaks, CA: SAGE.

Philip, T. M., Bang, M., & Jackson, K. (2018). Articulating the "how," the "for what," the "for whom," and the "with whom" in concert: A call to broaden the benchmarks of our scholarship. *Cognition and Instruction, 36*(2), 83-88. https://doi.org/10.1080/07370008.2018.1413530

Pickering, A. (1995). *The mangle of practice: Time, agency, and science*. Chicago, IL: University of Chicago Press.

Pieters, J. M., Voogt, J. M., & Roblin, N. N. P. (Eds.). (2019). *Collaborative curriculum design for sustainable innovation and teacher learning*. Dordrecht, the Netherlands: Springer.

Potvin, A. S., Kaplan, R. G., Boardman, A., & Polman, J. L. (2018). Configurations in codesign: Participant structures in partnership work. In B. Bevan & W. R. Penuel (Eds.), *Connecting research and practice for educational improvement: Ethical and equitable approaches* (pp. 135-149). New York, NY: Routledge.

Quartz, K. H., Weinstein, R., Kaufman, G., Levine, H., Mehan, H., Pollock, M., Priselac, J. Z., & Worrell, F. C. (2017). University-partnered new school designs: Fertile ground for research–practice partnerships. *Educational Researcher, 46*(3), 143-146. https://doi.org/10.3102/0013189X17703947

Renée, M., Welner, K., & Oakes, J. (2009). Social movement organizing and equity-focused educational change: Shifting the zone of mediation. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second International Handbook of Educational Change* (pp. 158-163). London: Kluwer.

Riedy, R., Van Horne, K., Bell, P., Penuel, W. R., Neill, T., & Shaw, S. (2018). Mapping networks to help education leaders gain insights into complex educational systems. In J. Kay & R. Luckin (Eds.), *13th International Conference of the Learning Sciences* (Vol. 1, pp. 656-662). London, UK: International Society of the Learning Sciences.

Roderick, M., Easton, J. Q., & Sebring, P. B. (2007). *Developing new roles for research in new policy environments: The Consortium on Chicago School Research*. Chicago, IL: The Consortium on Chicago School Research.

*Roschelle, J., & Hegedus, S. (2013). Introduction: Major themes, technologies, and timeline. In S. Hegedus & J. Roschelle (Eds.), *The SimCalc vision and contributions: Democratizing access to important mathematics* (pp. 5-12). Dordrecht, the Netherlands: Springer.

*Roschelle, J., Kaput, J., & Stroup, W. (2000). SimCalc: Accelerating students' engagement with the mathematics of change. In M. Jacobson & R. Kozma (Eds.), *Innovations in science and mathematics education: Advanced designs for technologies of learning* (pp. 47-75). Hillsdale, NJ: Earlbaum.

*Roschelle, J., Pierson, J., Empson, S., Shechtman, N., Dunn, M., & Tatar, D. (2010). Equity in scaling up SimCalc: Investigating differences in student learning and classroom implementation. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the disciplines: Proceedings of the 9th International Conference of the Learning Sciences* (Vol. 1, pp. 333-340). Chicago, IL: International Society of the Learning Sciences.

*Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., & Gallagher, L. P. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. *American Educational Research Journal, 47*(4), 833-878. https://doi.org/10.3102/0002831210367426.

*Roy, G. J., Vanover, C., Fueyo, V., & Vahey, P. (2012). Providing professional support to teachers that are implementing a middle school mathematics digital unit. *Contemporary Issues in Technology and Teacher Education, 12*(2), 145-161. https://www.learntechlib.org/primary/p/35987/

Schneider, J. (2014). *From the ivory tower to the schoolhouse: How scholarship becomes common knowledge in education*. Cambridge, MA: Harvard University Press.

Severance, S., Penuel, W. R., Sumner, T., & Leary, H. (2016). Organizing for teacher agency in curriculum design. *Journal of the Learning Sciences, 25*(4), 531-564. https://doi.org/10.1080/10508406.2016.1207541

Shaffer, D. W., & Squire, K. (2006). The Pasteurization of education. *Advances in Education and Administration, 8*, 43-55. https://doi.org/10.1016/S1479-3660(05)08004-2

Snow, C. E. (2015). Rigor and realism: Doing educational science in the real world. *Educational Researcher, 44*(9), 460-466. https://doi.org/10.3102/0013189X15619166

*Snow, C. E., Lawrence, J., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness, 2*(4), 325-344. https://doi.org/10.1080/19345740903167042

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*, 387-431. https://doi.org/10.3102/00346543072003387

Tseng, V., Easton, J. Q., & Supplee, L. H. (2017). Research-practice partnerships: Building two-way streets of engagement. *Social Policy Report, 30*(4), 3-16. https://doi.org/10.1002/j.2379-3988.2017.tb00089.x

Tseng, V., Fleishman, S., & Quintero, E. (2018). Democratizing evidence in education. In B. Bevan & W. R. Penuel (Eds.), *Connecting research and practice for educational improvement: Ethical and equitable approaches* (pp. 3-16). New York, NY: Routledge.

U.S. Department of Education. (2013). *What Works Clearinghouse procedures and standards handbook (Version 3.0)*. Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf.

*Vahey, P., Roy, G. J., & Fueyo, V. (2013). Sustainable use of dynamic representational environments: Toward a district-wide adoption of SimCalc-based materials. In S.

Hegedus & J. Roschelle (Eds.), *Democratizing access to important mathematics through dynamic representations: Contributions and visions from the SimCalc research program* (pp. 183-202). New York, NY: Springer.

Virkkunen, J. (2006). Dilemmas in building shared transformative agency. *Activités, 3*(1), 43-66. https://doi.org/10.4000/activites.1850

Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to the psychology of language*. New York, NY: Psychology Press.

Whyte, W. F. (1991). *Participatory action research*. Newbury Park, CA: Sage.

Wiseman, A. W. (2010). The uses of evidence for educational policymaking: Global contexts and international trends. *Review of Research in Education, 34*, 1-24. https://doi.org/10.3102/0091732X09350472

Wittgenstein, L. (1953). *Philosophical investigations*. London, UK: Blackwell Publishing.

*Yamada, H., Bohannon, A., & Grunow, A. (2016). *Assessing the effectiveness of Quantway®*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Retrieved May 23, 2020 from *https://www.carnegiefoundation.org/wp-content/uploads/2016/10/Quantway_propensity_score_matching_10-2016.pdf*.

*Yamada, H., & Bryk, A. S. (2016). Assessing the first two years' effectiveness of Statway®: A multilevel model with propensity score matching. *Community College Review, 44*(3), 179-204. https://doi.org/10.1177/0091552116643162

York, A. J., Valladares, S., Valladares, M. R., Garcia, M., & Snyder, J. D. (2020). *Community research collaboratives*. Boulder, CO: National Education Policy Center & Stanford Center for Opportunity Policy in Education. Retrieved May 23, 2020 from https://nepc.colorado.edu/sites/default/files/publications/PM%20York%20CRC.pdf.

Table 1. Principles Reflected in Each Project

	SERP-Word Generation	DBIR -SunBay's SimCalc	NIC -Carnegie Math Pathway	CBDR- PRIMES
Principle 1. The work should support the agency of participants by collaborating with them as partners in research.	Teacher input was valued throughout the partnership and affected changes over the course of developing and revising the intervention.	District leaders worked with researchers to modify the curriculum, including making adjustment according to context.	The Carnegie Foundation engaged the diverse members of the network in activities to develop a better understanding of the problem and to implement math curricula that supported student's agency.	Researchers and parents worked to restore a sense of "epistemic authority" in math and helped parents advocate for their children in relation to math education.
Principle 2. Accounts of the work should clearly define and describe the role and contributions of partners, particularly their expertise and how it was integrated into the research.	Accounts of the SERP collaboration with the Boston Public Schools make clear the unique contributions of (and the coordination among) researchers and partners.	Roy and colleagues (Roy et al., 2012) include a chart of major contributions and expertise of the different stakeholders.	Carnegie Math Pathways brought together groups of partners with diverse expertise from research and practice communities, with the Foundation serving as the hub of the network.	The research team engaged in the collaborative design of solutions; parents were equal partners in the program design; parents and educators were in involved in some data analysis.
Principle 3. The problem takes center stage in research and should be important to a broad range of stakeholders.	While problem identification originated in the local context, the SERP researchers also articulated the relevance of the problem to U.S. national policy leaders and teachers across the country.	Key stakeholders involved in defining the problem included the research teams at both universities, district leaders, and—to a more limited extent—teachers.	The NIC was formed with the purpose of solving the persistent and important problem of very low success rates in developmental math in many community colleges in the U.S.	A problem that drove the PRIMES project was the shift parents experienced as their children moved from elementary school to middle school and, with that, the school's expectations of their involvement as parents.
Principle 4. The research should attend to context.	Prior to building the intervention, research partners conducted teacher interviews and classroom observations in order to gain a deeper understanding of the context and barriers to comprehension success among students in BPS.	The SRI-based team conducted preliminary investigations into the unique opportunities in Pinellas County Schools early on to identify key constraints and stakeholders in the work. Materials and standards were tailored to fit the local context.	Procedures and tools were used further to continually adapt to the findings and iterate on solutions the NIC uncovered.	Before beginning the design of solutions to the identified problems, the team engaged in extensive "ethnographic observations" of families and families were also included in helping to make sense of the data collected.

Principle 5. What makes research valuable is that it provides something of practical value to participants and their organizations or communities.	Interventions were designed to be feasible within the time constraints and demands placed on teachers, requiring only 15 minutes once a week for each teacher in mathematics, science, and social studies.	The research team worked with district leaders to adopt pacing guides to fit easily within the needs of the local teachers.	The NIC developed practical solutions for mathematical learning in community colleges, such as developing new roles and structures to support successful and accelerated math attainment.	The project addressed a concern about math success at the transition between elementary and middle school and worked to validate the math families do at home and connect home and academic math practices.
Principle 6. The research should account for the gap between what was intended and what was accomplished.	Publications describe limitations of research designs that affected interpretation of findings, and they made recommendations for future research efforts to improve.	The accounts of SunBay include descriptions of how teacher turnover partly undermined their efforts to build a cadre of teacher leaders who could help others learn how to implement the units and advocate for them.	The team was explicit about the importance of understanding variation in effects within and across local context and determine that while Quantway 1 worked especially well in some places it did not in others.	Accounts of the PRIMES project pointed to limitations in changing the wider environment or engaging in what the team called "systemic repair," as well as issues engaging with parent advocates who did not feel confident in math.
Principle 7. The research plan should include specific, logical, and coherent plans for studying and following problems; for designing, testing, and iterating upon solutions; and for constructing and using practical knowledge.	The research designs, data collection procedures, and analytic strategies were all coherent and appropriate for the questions raised at each phase of development of the intervention, and the research designs grew progressively more complex after the quasi-experimental study of Word Generation showed promise.	The team was able to establish that the successive samples of students involved in SunBay achieved comparable levels of growth as students in the treatment group had in the earlier randomized controlled trials.	NIC methodology includes tools and procedures for analyzing systems and devising solutions, including fishbone diagrams, system improvement maps, journey maps, and driver diagrams. The Math Pathways project also included formal evaluation of the impact of the programs.	The research followed a coherent sequence, beginning with "critical design ethnography" that examines needs and opportunities for systemic repair, followed by an iterative, collaborative design process in which research findings are analyzed collaboratively to inform iterations on the design.
Principle 8. The research should be of value to others outside the partnership.	The Word Generation program materials have been made freely available to any and all educators interested in implementing the program and had a wide impact in Boston Public Schools in particular.	The team published strategies employed in the curriculum in magazines devoted to practitioner audiences as well as in scholarly journals.	Articles and briefs about design and effectiveness of interventions served as context for theory building regarding NICS. Statway and Quantway were made available as two change packages for community colleges.	Parents helped design math resources that were then shared with families not involved in the project. The workshops were made available to others, too, to lead them independent of the research group. A TV show was produced.