Consensus versus Concreteness: Tensions in Designing for Scale

Authors:

Marisa Cannata (Corresponding author)
Vanderbilt University
marisa.cannata@vanderbilt.edu
PMB 414
230 Appleton Place
Nashville, TN 37203
ORCID: 0000-0003-2336-028X
Twitter: macannata

Tuan Nguyen
Kansas State University
nguyetd1@ksu.edu
261 Bluemont Hall
Manhattan, KS 66506
ORCID: 0000-0002-2920-3333
Twitter: edu_tuan

Funding note: This research was conducted with funding from the Institute of Education Sciences (R305C10023). The opinions expressed in this report are those of the authors and do not necessarily represent the views of the sponsor.

Word Count: 12,840

Author bios:

Marisa Cannata is a Research Associate Professor in the Department of Leadership, Policy, and Organizations at Vanderbilt University and Director of the National Center on Scaling Up Effective Schools. She is the co-editor of *School Choice and School Improvement*, published by Harvard Education Press, and *Mapping the High School Reform Landscape*, a yearbook of the National Society for the Study of Education published by Teachers College Record. Dr. Cannata has a Ph.D. in Educational Policy from Michigan State University.

Tuan Nguyen is an assistant professor in the Department of Curriculum and Instruction at Kansas State University. His main research interests include teacher leadership and school improvement, teacher labor markets, and financial aid and postsecondary success. He holds a dual-degree in Math and Physics from the University of Oklahoma and a Master in the Arts of Teaching from Washington University in St. Louis. He taught middle and high school math for several years before obtaining his Ph.D. at Vanderbilt University.

A prior version of this paper was presented at the annual meeting of the American Educational Research Association in April 2015.

Consensus versus Concreteness: Tensions in Designing for Scale

Abstract

Substantial research on reform implementation highlights numerous challenges to implementing innovations at scale with depth and sustainability, yet new reforms continue to encounter many of the same challenges. This has led to calls for researchers to work in partnership with practitioners to design, implement, and scale educational innovations. While these approaches hold promise, little is known about the internal operations of these improvement approaches and the experiences of their participants. Through a case study of a research-practice partnership that uses a continuous improvement approach to design and development, this paper explores how the collaborative design process shaped the resulting innovation design. The data come from observations of design team meetings, member feedback forms, and interviews. The findings highlight tensions between achieving the necessary concreteness in the design through a process that valued collaboration and consensus. The paper has implications for researchers and practitioners who want to use continuous improvement processes to scale and sustain educational improvements.

Keywords: School reform; Scale-up; Continuous Improvement; Case studies

Consensus versus concreteness: Tensions in designing for scale

Extensive research on school reform implementation has consistently identified the challenges of scaling educational innovations (Berends, Bodilly, and Kirby 2002; Datnow, Hubbard, and Mehan 2002; Glennan et al. 2004; Stringfield and Datnow 1998), yet, new reform efforts continue to encounter many of the same challenges (Payne 2008). Recent scholarship on the relationship between research and practice suggests traditional approaches do not adequately address the reality of practitioner needs or lead to interventions that can be implemented and scaled with depth and sustainability (Bryk, Gomez, and Grunow 2011; Coburn et al. 2010). The emergence of design-based implementation research, improvement science, and networked improvement communities reflects this need for researchers to work in partnership with practitioners and other stakeholders to design, implement, and scale educational innovations (Bryk et al. 2015; Penuel et al. 2011; Tett, Crowther, and O'Hara 2003). These improvement approaches are likely to increase as funders increasingly incorporate requirements for networked-based approaches to improvement and research-practice partnerships (Cohen-Vogel et al. 2015; Coburn, Penuel, and Geil 2013; Sawchuk 2017). While the particulars of network and partnership-based approaches to school improvement may differ, they are united in a recognition that achieving success at scale is more than the identification of effective practices but also the ways in which the practices are implemented at the local level (Penuel et al. 2011).

The rise in networked approaches to improvement reflects a shift in the scaling up scholarship, moving away from focusing on the faithful implementation of a proven program in diverse contexts and towards a focus on sustained improvement in educational outcomes (Cannata and Rutledge 2017; Sabelli and Harris 2015). Leading scholars of school reform have

suggested we abandon the traditional approach of identifying a program that has rigorous evidence of impact on student outcomes and replicate it in other schools (Elmore 2016; Fullan 2016). Indeed, achieving improvement at scale has been compared to social movements where reform advances "through non-hierarchical networks that include educators both at the grassroots and in leadership positions" (Rincón-Gallardo and Elmore 2012; Niesz and Ryan 2018). Networked-based approaches have seen success in locales as diverse as Ontario, Long Beach, Mexico, and India (Rincón-Gallardo and Elmore 2012; Niesz and Ryan 2018; Gallagher, Malloy, and Ryerson 2016; Zavadsky 2016). Case studies of school systems that have experienced sustained improvements found that local ownership, teacher involvement, and continuous improvement were critical to their success (Fullan 2016).

While partnership or networked-based approaches to improvement hold promise, we know little about the internal operations of these approaches, leading to calls for more research on how engaging in co-design process in a partnership shapes the subsequent implementation and scaling of a reform initiative (Coburn and Penuel 2016). In particular, we do not know how do these types of approaches balance the need for both local ownership and specificity of the design. Research has long recognized the need for local ownership of reform and the ability of educators to adapt innovations to their context (Berends, Bodilly, and Kirby 2002; Datnow and Park 2009; McLaughlin 1987). At the same time, challenges exist when major decisions about the focus and content of the reform is left to local decision-makers (Cohen et al. 2013; Nunnery 1998; Watson and Michael 2016). In short, we know little about how to balance the appropriate amount of local adaptation.

This paper is a case study of a networked improvement community that sought to balance the development and implementation of a scalable, district-wide innovation with school-level

adaptation. By examining the development of an innovation focused on building student ownership and responsibility, this paper sheds light on the tension between developing a well-specified innovation and attending to the context of individual schools that are implementing it. Our work contributes to the literature on research-practice partnerships and network-based school improvement by exploring the complexities of local adaptation and the delicate balance of consensus and specificity in a networked improvement community. Moreover, our work highlights the enduring dilemma of creating an environment to foster innovation and local ownership while also maintaining integrity to a shared theory of change. The paper begins by reviewing the literature on networked improvement communities and reform design and implementation, highlighting how four design factors shape subsequent implementation. Second, we provide a brief overview of the research-practice partnership in which the current work is situated. Third, we describe the data that served as the evidence for constructing this case study. In reporting the findings, we first provide explicit examples of the design process and then use these examples to illustrate themes that cut across these factors. We end with a discussion of implications for practice and future research on reform development.

Theoretical framework

Networked Improvement Communities

Networked Improvement Communities (NICs), which have origins in improvement science, work to mobilize collective knowledge-building around complex problems and potential solutions, with various partners each contributing different forms of expertise (Bryk, Gomez, and Grunow 2011; LeMahieu et al. 2017; Russell et al. 2017). NICs form around a specified problem of practice, use common tools or routines in their work, and emphasize shared learning and collective improvement (Cannata, Cohen-Vogel, and Sorum 2017). By bringing together a

"diverse colleagueship of expertise," NICs are intended to "enhance the efficacy of individual efforts, align those efforts and increase the likelihood that a collection of such actions might accumulate towards efficacious solutions" (Bryk, Gomez, and Grunow 2011, 5). While NICs in education are relatively new, they build off continuous improvement efforts in other industries (Engelbart 1992; LeMahieu et al. 2017).

NICs are both design communities and learning communities that engaged in research and development while also arranging human resources and knowledge-based tools to organize improvement work (Bryk, Gomez, and Grunow 2011). A common protocol of inquiry, such as Plan-Do-Study-Act cycles (PDSA), helps to provide a common language and system of measurement that facilitates a disciplined improvement approach (Bryk, Gomez, and Grunow 2011; LeMahieu et al. 2017). Core to NICs' appeal is that it is perceived as a way to address the challenges inherent in designing, implementing, and scaling up interventions. These include lack of teacher buy-in (Glennan et al. 2004; Nunnery 1998), inadequate attention to the organizational context in which practices are to be implemented (Bodilly et al. 2004), and conflicts with existing district programs (Berends, Bodilly, and Kirby 2002; Sanders 2014; Stringfield and Datnow 1998). By involving educators in the design of change ideas specific to their context, while working in a network to accelerate learning, NICs are expected to support the successful scaling of improvement initiatives that allow for adaptation to local context.

Involving local actors in developing reforms for their context has been intentionally included in some reforms (Rowan et al. 2009), probably because lack of attention to context has been a key stumbling block (Datnow and Park 2009; Supovitz 2008). Further, local adaptation becomes inevitable as reform designs are not enough in and of themselves to successfully reform schools and improve student performance (Berends, Bodilly, and Kirby 2002). This attention to

local context is particularly important for achieving scale as innovations must be able to fit with contexts that vary greatly in organizational structure, buy-in, capacity, and funding while coping with change, promoting ownership, building capacity, and enable effective decision-making (Cohen et al. 2013; Dede and Honan 2005; Peurach and Glazer 2012; Tett, Crowther, and O'Hara 2003).

At the same time, allowing too much local development has drawbacks. First, reforms are most effectively implemented and have larger impacts on student learning when they have a well-specified design (Cohen et al. 2013; Rowan et al. 2009). Second, to successfully implement a reform, educators need sufficient training on what is expected of them; the greater the access to technical expertise, the easier it is for educators to understand what they should be doing (Berends, Bodilly, and Kirby 2002; Desimone 2002). Furthermore, improvement efforts that intentionally build in local adaptation require capacities such as time, expertise, and collaborative ability that teachers may not have, particularly in low-performing schools (Berends, Bodilly, and Kirby 2002; Cohen et al. 2013; Datnow et al. 1998).

In sum, successful improvement efforts must balance fostering local adaptation and ownership of change with maintaining integrity to core elements of the improvement effort. This is especially meaningful in networked approaches to improvement where researchers, practitioners, and other stakeholders all play critical roles in the adaptation and implementation of a co-created reform (Cannata, Redding, and Nguyen in press; Coburn and Penuel 2016). Parsing the research on the opportunities and challenges with local development of reforms suggest that network leaders need to negotiate how to provide the necessary specificity and support while providing room for alignment with local context and a sense of ownership (Cohen

and Ball 1999). To address this, we turn to design factors that can facilitate or hinder the complex task of balancing specificity and local adaptation.

Design factors that shape adaptation and implementation

Shiffman and colleagues (2008) identify four design factors that shape how reforms are subsequently adapted and implemented: design emphasis, innovation complexity, implementation supports, and innovation engagement. These design factors are important because, regardless of how the reform ideas are developed, empirical research demonstrates that these features influence whether the ideas lead to deep change in practice (Shiffman et al. 2008). Below, we discuss how each factor contributes to reform adaptation and implementation.

The design emphasis factor focuses on what constitutes the major elements of the design, such as the features that are considered most central to the reform, whether there is a reliance on a particular organizational structure, and how the core features are sequenced (Shiffman et al. 2008). Design emphasis provides the overall picture of where and how the reform practices are intended to take place. The specificity of these practices is important as more specific practices provide more guidance for implementation (Desimone 2002). Reforms are most effectively implemented and have larger impacts on student learning when they have a well-specified design with clear routines for educator behavior (Cannata, Redding, and Nguyen in press; Cohen et al. 2013; Rowan et al. 2009). For reforms to succeed, there needs to be clarity in the major elements and what is expected from teachers and administrators (Berends, Bodilly, and Kirby 2002; Desimone 2002). The ability of network-based approaches to provide clarity and specificity around the major features of the reform design, then, are important for improving student outcomes.

The second design factor focuses on the complexity of the innovation, which highlights the difficulty local actors encounter while enacting the design (Shiffman et al. 2008). The more complex the design, the more local actors will experience challenges in making sense of the reform practices and implementing them in practice (Supovitz 2008). This complexity can be disaggregated into two components: level of abstraction and technical difficulty. Similar to design emphasis, if the design is too abstract, it will not provide the necessary specificity for local actors to translate them from ideas into actions (Desimone 2002; Nunnery 1998). The technical difficulty focuses on the level of skills, and thus amount of teacher learning, required to effectively enact a design component (Desimone 2002). The greater the degree of complexity in either component, the greater the difficulty practitioners will experience with implementation. Network-based approaches to improvement at scale, with their focus on how systems influence individual behavior (Bryk et al. 2015) can appear more complex. However, they use a combination of a shared theory of change (i.e., a driver diagram in the context of NICs) and specific change ideas to reduce the level of abstraction while helping educators see how specific change ideas are nested within the larger system (LeMahieu et al. 2017; Russell et al. 2017)

The third design factor focuses on the implementation supports available to local actors, which may include professional development, classroom-based assistance, modeling of practices, example materials, and common planning time (Bodilly 1996; Desimone 2002; Nunnery 1998; Shiffman et al. 2008). Establishing networks of support where teachers have frequent and deep connections to reform expertise and learning opportunities embedded in daily practice fosters sustainability and teacher commitment to reform practices (Camburn 2010; Coburn et al. 2012). The specificity of implementation supports is an important part of network-based improvement

approaches, as the process of implementation is itself tested and iterated upon (LeMahieu et al. 2017; Fishman et al. 2013).

The final design factor is innovation engagement, which focuses on how the innovation builds commitment and engagement in local actors (Shiffman et al. 2008). Teacher commitment and buy-in to the reform is critical to successful implementation and scale (Glennan et al. 2004). One way of engaging local actors in a reform is to have them self-select into the reform or help to co-construct the innovation, perhaps through participative decision-making about the reform (Camburn 2010; Devos, Tuytens, and Hulpia 2014). Other ways to engage local actors is to alter their work in meaningful ways, relate the reform practices to shared problems, and demonstrate the effectiveness of the reform (Shiffman et al. 2008). This process of engagement, with a focus on developing reform ideas that are user-centered and relevant to improving the everyday work of educators and informed by the knowledge of a diverse set of stakeholders, is a core principle of network-based improvement approaches (Bryk et al. 2015).

This paper provides a case study of a single NIC to bring effective practices to scale within a large district, guided by the following questions around the design factors: *How did the innovation's core design emphasis develop? How was the innovation's complexity addressed? In what ways were the implementation supports discussed as part of the design process? How did the process engage both members of the design teams and teachers in the school?* The approach was designed specifically to address past challenges in scaling effective practices and uses elements of design-based implementation research and implementation science (Bryk et al., 2015; Fishman et al., 2013). As policy implementation research shifts to have development and ongoing implementation become joint work of researchers and practitioners (Cohen-Vogel et al. 2015), this paper offers insight into the daily work of network-based improvement approaches

and how these new forms of educational improvement work enact and balance innovation design and development (Rosing, Frese, and Bausch 2011).

Context

The NIC described here established three key features to build buy-in among local implementers and ensure alignment with district and school contexts: 1) The design effort was based on research conducted in the district to identify effective practices, 2) The process used a continuous improvement approach with iterative cycles to build knowledge for design and implementation, and 3) Practices were co-developed by researchers and practitioners serving on district and school design teams (Cohen-Vogel et al. 2016). The work began in 2011-12 with an intensive study of district high schools to identify programs, practices, and processes that differentiated the higher and lower performing high schools. The findings from this initial research established the "design challenge" of developing Student Ownership And Responsibility (SOAR) that became the focus of subsequent work (Cannata, Smith, and Taylor Haynes 2017). SOAR included (a) changing students' beliefs and mindsets to increase self-efficacy and (b) engaging students to do challenging academic work. The design process described below ultimately settled on two core strategies: developing growth mindsets and problem-solving skills in students.

A District Innovation Design Team (DIDT) was established and charged with developing an innovation that would be implemented in three high schools (known as innovation schools). The DIDT had 23 members, including two to three representatives from each innovation school, representatives from six other high schools, five representatives from the district central office, three university-based researchers, and a coordinator who served as a liaison between the external personnel and the district. The DIDT was facilitated by an external organization. Over

seven months, the DIDT met monthly for two days to learn about the design challenge, conduct needs analysis, and develop an innovation prototype. We refer to this phase, which took place from February to August 2013, as Phase 1. In Phase 2, School Innovation Design Teams (SIDTs) were established in each of the innovation schools and consisted of six to eight individuals, nearly all of whom were teachers. SIDTs were charged with taking the innovation prototype that the DIDT developed and engaging in further development, testing, and adaptation to their school, as well as planning for full implementation. During the 2013-14 school year, the DIDT/SIDT had twelve full-day meetings, four webinars, and two after school meetings. At the conclusion of the two phases described in this paper, schools were expected to begin full implementation (in the 2014-15 school year).

During both phases, the activities were divided into those that were intended to build capacity in members and those intended to make decisions related to the innovation design. We focus here on seven activities that were specifically focused on the innovation design:

- The introduction of five practices that the research team identified as having evidence of improving SOAR-related outcomes (Phase 1, session 3).
- After four sessions of needs analysis and brainstorming, the initial attempt to formalize the core strategies that comprise the SOAR innovation (Phase 1, session 5).
- The first attempt at revising the SOAR core strategies (Phase 1, session 6).
- The review of a growth mindset lesson for students (Phase 2, session 2).
- An initial discussion of research about problem-solving processes (Phase 2, session 2).
- After piloting growth mindset and problem-solving lessons, several school-based implementation planning activities across multiple sessions. Between these sessions, SIDTs

were expected to use continuous improvement processes to test individual practices on a small scale before including them in their implementation plan (Phase 2, sessions 8 and 9).

A senior district leader attended nearly all sessions for a short period of time but did not participate as a full DIDT member. He guided the work by participating in periodic meetings with the researchers and facilitators about how the work was progressing; voicing his support for the collaborative, bottom-up design process; and encouraging principals to have their schools participate. The innovation schools were selected through negotiation between the researchers, senior district leader, and principals, on the basis of two criteria: value-added scores placed them in the bottom half of high schools in the district and a sense that school conditions would allow for a collaborative design process to be successful. Table 1 provides descriptive information on the innovation schools and the composition of their SIDTs. All schools had experienced some success with recent improvement efforts, although the nature of those efforts varied. DIDT representation from the innovation schools was selected by the principal. DIDT members from the central office and other high schools were selected through negotiation between the deputy superintendent, facilitators, and principals in non-innovation schools.

The authors (along with other researchers) were both participants and observers of this design process. The lead author served on the DIDT along with two other researchers. The research team also contributed expertise in several ways throughout the partnership, some of which placed researchers as prominent participants in the process. For example, the lead author led several activities and gave presentations that served to define SOAR and its core elements for the design teams. Researchers also provided training on how to engage in the continuous improvement cycle, modeled how to interpret data, and led activities on growth mindset and problem-solving. Other ways in which the researchers shaped the design process were less

visible to design team members. For example, researchers conducted a scan of existing research to identify programs related to SOAR that had evidence of effectiveness and brought in curricular materials related to growth mindset interventions. While these were important contributions, the activities that introduced these materials were led by the facilitators. At the end of every session, all researchers who attended that session prepared a reflection form that was shared with the facilitators. Researchers also engaged in discussion with the facilitators about subsequent design team activities in biweekly meetings. Thus, the researchers were clearly in a position of power within the design team. At the same time, the use of an external organization to serve as facilitator moderated that power.

Data and methods

In addition to the three researchers who served on the DIDT, other project researchers attended all design sessions to take fieldnotes, audio record session discussions, and collect artifacts. The reflection form completed by researchers after each session served as additional fieldnotes. Interviews were conducted with participants at several points in the process. Facilitators were interviewed twice in each phase and DIDT members were interviewed at the end of both phases. A random sample of SIDT members was interviewed at the end of Phase 2. In total, there were 174 hours of audio files, 42 fieldnote logs, 556 artifacts distributed or produced during meetings, 57 meeting minutes, 13 sets of participant feedback forms, and 44 interview transcripts that were collected and analyzed.

Following data collection, the research team conducted an in-depth reconstruction of the DIDT/SIDT process. First, we engaged in a process of data reduction with the audio recordings. Audio data were not transcribed in their entirety due to their length and complexity. Instead, researchers listened to each recording and utilized reflection forms to partially transcribe and

synthesize data falling within our analytic framework. This framework includes: attitudes and engagement; delivery of learning about design challenge and implementation; participant understanding of design challenge and implementation; the extent to which the design process adhered to principles of good design; participant perceptions of the final design; and key points to understand the process that emerged through initial data analysis (see Table 2 for additional detail on the coding framework). Second, all data were systematically analyzed through directed content analysis (Patton 2002), according to this framework. The research team built reliability by simultaneously coding an initial set of documents, consisting of examples of each type of data collected. The team then met to discuss areas of misconception, and to gauge inter-rater reliability. These meetings continued for the duration of the coding and memo writing process.

Third, after this initial reliability-building period, coders engaged in cycles of coding, memo writing and discussion for each session. Researchers were assigned to code all data associated with a particular day-long session and write a memo that synthesizes the evidence for that session around the analytic framework. Fourth, working with the comprehensive session memos, feedback form data, interviews, and meeting minutes, researchers synthesized the data within a single component of the framework (i.e., participant understanding of the design challenge). This process resulted in the production of a summary memo for each phase.

Through this first analytic process, several key themes about the relationship between collaboration and consensus-building, the role of local adaptation, and developing required specificity emerged. We recognized that the framework of design features by Shiffman and colleagues (2008) could help to explain these themes. A second coding and analytic process was used to focus the analysis around the four design factors (design emphasis, complexity, innovation engagement, and implementation support and the three emergent themes of

collaboration, local adaptation, and specificity). In this process, two researchers coded the comprehensive session memos and summary memos that were produced in the prior analytic process, as well as two documents that described the core features of the design. The two researchers engaged in a similar reliability process as the one described above to ensure they were applying the coding framework similarly. Then, we summarized the evidence under each code to draw out the main themes as described below.

Findings

The evidence converges on three main challenges in the design process. These challenges point to inherent dilemmas in maximizing all the design factors as they sometimes conflicted with each other: (1) Members were most engaged when the work was very concrete and deemed feasible in a particular context; (2) Efforts to develop more specificity in the design emphasis were limited by efforts to engage educators in a collaborative process where they had ownership over key design decisions; and (3) The abstractness of the emerging innovation led to difficulties in establishing a shared deep understanding of each core component of the innovation. Finally, the ability of school teams to productively resolve these dilemmas were related to the existing capacity of the school. We first describe the process longitudinally to aid understanding of the process itself, and then draw out these themes using examples from the process.

Phase 1: Deciding on the core components

The first two sessions of Phase 1 focused on community building, understanding the initial research that identified SOAR as the design challenge, and introduction to the work. The first activity that focused explicitly on designing the innovation occurred in Session 3; this activity highlights the role of contextually specific implementation supports in the team's work. During this session, five research-based practices were introduced to the DIDT. The practices

included: an evidence-based model of student engagement; growth mindsets; blended learning; project-based learning; and student feedback processes for self-regulating learning. These practices were selected by the research team based on a review of prior research and conceptual linkage to SOAR. The DIDT was divided into five groups, with each group engaging in a discussion around a short reading about the practice. Then, each group shared what they discussed with the full DIDT.

In two of the groups, members displayed negative attitudes about the practice, citing lack of alignment to district practices or other potential implementation supports. In a discussion of blended learning, the group cited several problems with the district's experience of online credit recovery and a member from Wheatley said, "my immediate reaction is that this is a tool or structure that could be put in place, but I'm not seeing how this is going to happen….We have Plato and they are still struggling." An at-large member added, "we had to restructure [Plato] because it didn't work without daily monitoring from the teacher. Some students can handle it, but many need a lot more from the teacher." Similarly, in the group on project-based learning, members focused on challenges of implementation given the district curriculum frameworks, with a member from Cervantes saying, "The research shows that it works, but you are leaving it up to innovative, individual pockets, but that's because he's going off the curriculum. If a [district] learning specialist comes in, they're coming to see the curriculum." With this perception that the programs would not have adequate implementation supports in the context of their district, both groups ultimately disengaged from further discussion of their assigned practices.

In contrast, the growth mindset group remained consistently engaged and focused on concrete ways to build growth mindsets in students. For example, an at-large member

summarized their conversation by saying, “we really like the idea of scripted conversations so that everyone is hearing the same thing, and having that conversation surrounded around effort. We see it as linked to SOAR through the belief that, ‘through effort, I will grow.’” A Wheatley member focused on adapting techniques that are already common practice to reinforce ideas about growth mindset, “we all do exit tickets; what if there was an exit ticket after every class, what was the best mistake of the day?” In this way, growth mindset became a concrete idea that was feasible to implement in small ways within their current context, in ways that blended learning or project-based learning were not. In the feedback forms at the end of this session, members were asked an open-ended question about what practices they were most interested in including in the innovation. Of the 21 responses, 12 wrote they wanted to include something related to growth mindset. The next highest practice, mentioned by four people, was AVID, which was an existing district program that was not introduced by the facilitators or researchers but repeatedly mentioned by members as relevant to SOAR. In regards to the design factors, this activity suggested that members began to emphasize particular elements in the design due to perceived availability or alignment with particular implementation supports in their context.

In the remainder of this session and in the next one, this discussion of how specific practices would work in their district and meet the needs identified around SOAR continued. By the fifth session of Phase 1, the DIDT had identified a number of ideas and practices for potential inclusion in the innovation. With the school year coming to a close, the task in Session 5 was to formalize the content of the design, making decisions about which practices should be emphasized in the innovation design and how they build into a coherent improvement approach. This activity thus provided key evidence about how the innovation came to have a particular design emphasis. Working with a consensus-building process, only ideas that had consensus

would be included in the innovation; ideas for which there was no consensus were discussed or revised until all concerns were addressed. Two recurring concerns were raised throughout this activity: what to do about ideas they saw as important but were considered not feasible for next year, and how much specificity the design will have for schools.

First, several members expressed a desire for a dedicated time, such as an advisory period, as a key implementation support for the innovation. A member from Wheatley countered that, since school bell schedules had already been finalized, "this is a constraint that can't be altered next year." This led to an extended discussion of how to handle aspects of the design that members did not want to disappear altogether but for which there would not be the necessary supports for implementation in the next year. One member from Walker said, with others agreeing, "if we vote against them now, they will never be revisited." An at-large member summarized these sentiments, saying "My concern is that if it is not listed somewhere, it will *never* happen." Despite these concerns, the pressure to make a decision about specific practices to test in the coming year led them to focus on what was feasible to implement now. A dedicated time would not appear in design discussions again until the end of the next year when one school included it in their implementation plan.

The second recurring concern was the degree of specificity the design needed to have. Many members were concerned that the emerging design emphasis was too abstract and not specific enough. Using a consensus process, facilitators narrowed ideas to the "what" around which they could obtain consensus, leaving points of disagreement as the "how" to be decided later. Members disagreed about the appropriateness of the distinction between "what" and "how." Some members felt as if there should be more detail and that "all details are being hollowed out" of what they have designed. They argued that keeping the design more abstract

made it more difficult to build teacher engagement around it. One member from Cervantes said: "Our effectiveness as an SIDT is predicated on having *something* to work with" and a member from Wheatley added, "we need to design with enough specificity for teachers." Other members were concerned that they would face more challenges engaging teachers if the design was too specific. An at-large member had a counter view and said "we're not going to get buy-in if we constrain this too much." Later in the conversation when talking about student engagement, a Wheatley member said, "This is more 'how'…. This is how you engage students. There is no way we can get around how." Ultimately, the agreed upon design included broad statements such as "Teachers' share examples of global and community real world connections to content-based activities related to SOAR skills" although the structures by which that sharing would take place or the definition of SOAR skills was not defined.

In the feedback forms at the end of the session, multiple members expressed frustration over this lack of specificity and its implication for their ability to engage teachers. One member wrote,

> I think that it's necessary that we first finish up what the plan is...I can't be confused about what the plan is or unclear of what the plan is if it's my responsibility to teach others and bring others in on said plan.

In interviews, DIDT members reflected on this activity and several noted that while they felt positive about the innovation, they thought there needed to be more details for implementation. When asked about whether the innovation met the needs of the innovation schools, one DIDT member described how the facilitators appeared to intentionally keep the discussion at a more abstract level:

> We have problem-solving which is just sort of a brutally wide area.... There's a line between prototype and implementation. There's this, okay, we're, like we'll be having a conversation, and we'll be directed to stop the conversation because we're getting too far into implementation.

This member, like others, noted the lack of concreteness of the innovation suggesting that while members agreed with the emerging design, they were agreeing to a rather abstract statement about what the work in schools would actually entail.

This lack of concreteness and disagreement over the appropriate amount of specificity continued as the work shifted into Phase 2. The final session of Phase 1 began by DIDT members reviewing the description of the innovation that came out of their consensus process as they planned how to introduce it to the new SIDT members the following day. At this point, the innovation consisted of broad statements grouped into four components around growth mindsets, problem-solving, goal-setting, and self-monitoring. The discussion focused around two issues. First, within these broad components, how consistent do the three schools need to be? Members wanted to engage SIDTs in the design process as a way to foster their engagement with SOAR, yet recognized that could mean school practices would diverge from each other. For example, one at-large member asked, "How might we deal with SIDT members that tell us we ought to do this or ought to do that, how consistent do we have to be?" Several members echoed this concern. A researcher responded, "we want something that is similar enough that everyone can benefit from the conversation" about specific school-level adaptations. A member from Walker says, "there needs to be a core 'it' to go district-wide."

The second concern raised was about the amount of time to teach lessons around the four SOAR components. The sense of time quickly became a constraint on what they could do. One Wheatley member said, "I believe we will only get 30 to 40 minutes of time every six weeks" to implement in classrooms. A Cervantes member responded, "30 minutes is not enough time to explain goal-setting, problem-solving, self-monitoring." Another Cervantes member suggested that "this is too broad, we need more specificity." A Walker member pointed out that time

constraints differ between the schools, so they shouldn't focus so much on what is happening in a 30-minute block of time. Another member continued to express concern that "if it is vague or not immediately classroom ready, the SIDT at different schools will struggle to get these specific pieces in place."

Ultimately, a facilitator suggested they make the innovation seem less overwhelming given time constraints by prioritizing two of the four components. With all members quickly agreeing that growth mindset was foundational and self-monitoring did not make sense without the others, the discussion focused on whether to prioritize goal-setting or problem-solving. There was a heated debate about these components, revealing that members' concerns that the innovation was overwhelming for SIDT members may be less about the number of components and more about the unresolved abstractness and level of complexity in their ideas. One Walker member argued they should begin with goal-setting by saying, "I don't problem-solve just to problem-solve, I problem-solve to attain a goal." A Cervantes member countered by suggesting a lack of problem-solving skills was a bigger obstacle for students, "problem-solving is the method they use to achieve it…they don't know the steps to get them there, or if they fail a step they don't know how to rebound." The discussion uncovered a good deal of ambiguity about what a focus on goal-setting or problem-setting meant. For example, an at-large member asked "Is goal-setting content-specific or universal?" Another at-large member wondered why they could not "just collapse goal-setting into problem-solving?"

At the facilitator's suggestion, the team took a vote on whether to prioritize problem-solving or goal-setting. Yet, there was pushback about this suggestion, with a member saying, "I can't make a logical vote since these things aren't clarified enough yet." With a vote of 8 to 7, problem-solving and growth mindset became the two main components of the innovation's

design emphasis. The uncertainty felt by DIDT members after this decision was evident in their response to attempts by the facilitators to achieve a positive tone in the meeting. One facilitator suggested they clap after deciding on problem-solving to mark this occasion, but no one did. The discussion following the vote indicated members did not share a deep understanding of what it meant to implement growth mindset and problem-solving. For example, an at-large member asked, "How does mindset tie into problem-solving?" As they transition into a break after the discussion, a facilitator asked them to reflect on the good work they accomplished this morning. Again, no one responded.

Phase 2: Getting concrete with school-based testing

The next day, in Phase 2, Session 1, the DIDT members introduced the new SIDT members to the design process in which they engaged and the innovation and definition of SOAR. The tone was largely positive as they focused on a broad introduction to their work. In Session 2, which occurred a month later, SIDT members began the in-depth work of designing the practices they would test in their schools. In contrast to the decision around problem-solving or goal-setting, which left design team members unsure and frustrated, one of the first activities of this session had high engagement, and also emphasized the point that members were the most engaged when they were working within particular contexts and concrete practices.

At this point in the process, there was a clear emphasis on teaching about growth and fixed mindsets. Using a lesson plan and materials built from prior growth mindset interventions, members were assigned to cross-school groups to review the materials and provide feedback on the lesson materials. That members enjoyed this focused review of a lesson plan was clear. Engagement was extremely high; all members participated throughout the 90 minutes during which they reviewed the lesson, with minimal off-topic conversations. In all four groups, all

members contributed ideas or questions. Nearly all comments were constructive, as members critically reflected on the materials. At the end of the activity, one member said, “there is potential for awesomeness.” A member from Cervantes noted the collaboration and said, “I wished we could always plan lessons this way.” The suggestions for improvement included: discussions around whether the assigned reading was at the appropriate reading level for high school students, whether a video may be more useful for engaging students, how to make use of the anticipation guide, whether the time estimates were realistic to complete the lesson over two periods, and how to incorporate instructional strategies that were common in the district. These suggestions demonstrate how members were engaged around very concrete aspects of an innovation that would be introduced to students.

At the same time, the activity also surfaced misunderstandings about mindsets. For example, while one group was discussing the anticipation guide that has statements oriented around a fixed mindset, a Cervantes member said the statement about “affirmations/praise” should be clarified. A Walker member replied, “I think this means that you shouldn’t praise what they are, i.e., ‘you’re pretty,’ and should say ‘good job’ instead.” Another Cervantes member added, “I think ‘fixed number of connections’ is confusing, are they talking about physical connections or connections you can make on your own?” The confusion about whether they were teaching students about mindsets or about how the brain worked also came out through the most negative comments of the activity from any group. One member from Walker pushed back on the idea that they were trying to have more students adopt a growth mindset, calling it “indoctrination.” This member further clarified, “I think students ought to be presented with facts regarding both [growth and fixed mindset] and then make an informed decision about which to adopt.”

Despite some ambiguity around their understanding of growth mindset and how the brain works, the productivity and engagement of this design moment stands in contrast to an activity the following day focused on the other core component, problem-solving. In this activity, a facilitator and a researcher shared different research-based frameworks for problem-solving that included making a distinction between well-structured and ill-structured problems. This activity was designed to build a common understanding of problem-solving and the types of problem-solving they want to encourage through the SOAR initiative, but it was not focused on concrete activities they would do with students. Engagement was low for most of this activity, with only a few members asking or answering questions. At one point, three members headed to a back table for a side conversation. The lack of clarity about the discussion's relationship to the emerging innovation, was apparent when, after about 20 minutes of discussing the difference between well-structured and ill-structured problems, a member from Walker asked, "why are we talking about this?" The facilitator responded, "because we need to expand our research base for problem-solving." By the end of the discussion, most members were disengaged and not paying attention. This conceptual discussion of problem-solving was followed in a later session by a more concrete discussion of a problem-solving lesson plan. However, it took place in a webinar format, which limited members' ability to engage collectively.

By the second half of Phase 2, teams had tested the growth mindset and problem-solving lessons in their classrooms and they were revised based on evidence collected during those tests. With lessons for each of the two core components, the design work shifted to how the SIDTs would implement the lessons as a coherent innovation in their school. In Session 8, SIDTs were given time to develop an implementation plan for their school and then asked to identify a specific part of that plan to test now.

As the SIDTs focused more on their school context, differences began to emerge in what the design emphasized between schools. For example, the Wheatley team began by identifying the supports they could draw upon for implementation. One member said, "we need to consider the system's constraints regarding implementation" and another asked "to what degree are school assemblies or other school meetings set in stone for the year?" With the logistical challenges in mind, Wheatley outlined a plan to spend the second day of school with each teacher teaching the same lesson on growth mindset, that was divided up into seven segments. Each component would be delivered in a different period. By the end of the day, students would have participated in the complete lesson across all their classes. SIDT members were concerned about how teachers would respond to this plan, with one member saying that teachers could think "I'm only teaching one thing, but I'm just teaching parts of it at a time." This plan also integrated an existing school-wide literacy practice into the growth mindset lessons to align SOAR with other school priorities. The SIDT chose the second day of school because, "day 1 is critical for teachers to set the tone for their classroom…day 2 is more acceptable to tell teachers what they need to do." This day 2 activity was accompanied by a plan to devote nearly a full day of professional development for their faculty, introduce problem-solving at the six-week mark, and provide ongoing support for growth mindset and problem-solving practices throughout the year. Throughout this discussion, all SIDT members were engaged and focused on turning the abstract ideas of growth mindset into implementable practices and considered how teachers would respond to these practices. They decided to test out an idea for a behavioral reflection sheet intended to give students who are misbehaving a chance to think about and reframe their behavior instead of referring them to the office.

Cervantes began with all members participating in a brainstorming practice, until one member said, “we’re getting lots of good ideas; we need to distinguish which practices we think we will implement this year versus next year or in the future.” The discussion then focused on additional piloting of growth mindset and problem-solving practices this year. Specifically, they focused on identifying other teachers who could be “early adopters” to pilot the lessons they had previously piloted as a way to engage additional faculty, as well as testing whether a peer editing process helps to build growth mindset. Similar to Wheatley, the SIDT was positive, with most members engaged throughout the discussion.

The Walker SIDT encountered more difficulty in coming to agreement on their specific implementation plan. One idea that emerged was similar to Wheatley’s, with the content they wanted students to get being divided among different teachers and delivered on the first two days of school. A key difference was that teachers would be delivering the same lesson to each new group of students throughout the day, with teacher assignments based on their content area. For example, science teachers would provide the mindset lesson, math teachers would provide a lesson on problem-solving, and English teachers would focus on goal-setting. This was first proposed after a science teacher noted, “All the biology teachers teach how the brain works. So the first lesson on growth mindset could be done there.” Yet, other SIDT members were not happy with this decision, with one saying, “It might feel like we are dumping this on everybody and [teachers will] think that it will go away.” Another member said, “I don’t see how this is going to work in my department [English/language arts].” As they discussed potential pushback with implementation, one member asked about the challenge of teacher and administrator support, “Is it the fear of not being able to get the faculty and administrators to buy in?” Ultimately, the group discussion fragmented into three separate conversations and they did not

decide on a specific practice to test before the next session. The challenges in Walker of coming to agreement on their implementation plan was also evident in how their plan changed just a week later after sharing their progress with the principal, which introduced a goal-setting template in a homeroom class and moved away from the SIDT's work around growth mindset and problem-solving.

In Session 9, SIDTs shared with the network the results from the practices they had tested. A member from Wheatley shared the positive response they had from students on the behavioral reflection form. In their school-based discussion following this network sharing, the SIDT began discussing how to incorporate it into the innovation. A member whose role was not exclusively classroom teaching expressed major reservations about this form being used more extensively in the school, which led to a tense discussion between two members. The fieldnote log noted, "Discussion is entirely between [two members]; feels a little tense; other participants are remaining very silent." After an at-large member intervened by saying "I don't want to lose sight of the fact that this form is just a way to help teachers be effective in implementing the bigger innovation of problem-solving or SOAR; the form isn't the end-all," the two members agreed to table their disagreement and plan other components. All members engaged again as they fleshed out details of the innovation around problem-solving, with one member saying, "we want to make sure that we convey that the problem-solving skills are transferable across subject areas." As they continued planning, the fieldnote log stated, "Emphasis on creating a common language around SOAR, growth mindset, and problem-solving; most of the discussion seems to be revisiting things they have already worked out." By the end of the discussion, they had a fairly detailed implementation plan for the following year.

Cervantes had shared the results of engaging more teachers in piloting the growth mindset lessons and a testing a peer editing process, but began their school-based planning session by noting the importance of reflection. One member, who taught in the AVID program, suggested a process similar to what is used in AVID. She said they could "look at progress report grades and make a reflection and then reflect again when grades come out. Compare the two reflections." Other members agreed, and they developed a plan of weekly advisories to engage in this reflection and teach about growth mindset and problem-solving. The written implementation plan they produced noted that problem-solving will involve both the advisory (to talk about "life problems") and academic classes (to talk about "academic real world problems"). The SIDT quickly agreed on the major components of the innovation for students (i.e., the topics of the weekly advisory discussions). Most of their discussion focused on how to resolve the challenge of providing enough professional development for teachers. One member worried, "If we don't teach it explicitly, it's going to end up being [poor quality]...There is already too much put into department meetings, so that is not a good place." At the advice of a central office member, they decide on "a half-day waiver day and a half-day on a day on the Friday before in-service starts," as well as ongoing supports in schoolwide faculty meetings.

In contrast to Wheatley and Cervantes, who ended the session with a concrete implementation plan, Walker again had difficulty coming to agreement. After some discussion of the goal-setting template they had developed with their principal, one member introduced a new approach:

> Twelve, 10-15 minute lessons that teachers do every 3 weeks. Consistent teaching about growth mindsets and problem solving in homeroom. Maybe watch a video and have a discussion. If each of us come up with a lesson, then we aren't coming up with everything.

Another member agreed, saying, “The team will break growth mindsets and problem solving into twelve chunks, six chunks for each maybe, we’ll work on it.” Yet, the disagreement in the group was evident as another member brought up the plan from Session 8 of “introducing the lessons not on the first day of school, but maybe the second or third day. Teachers would be interested in doing it as long as they don’t have to plan the lessons.” Other members went back to fleshing out details of the goal-setting sheet. When asked how they will support teachers in implementing the lessons, the response was that they “would like to make their lessons a series of videos or Powerpoints that would require teachers to only press play.”

Comparing the final implementation plans across the three schools, they all included a design emphasis on growth mindset and problem-solving, but they achieved this in quite different ways. Wheatley had plans to give detailed lesson plans for all teachers in all periods of the second day of school, in addition to introducing problem-solving later in the year, and the use of a behavioral reflection form to build students’ mindsets and problem-solving abilities. Cervantes had a plan of a weekly advisory time to reflect on grades and deliver lessons around growth mindset and problem-solving. Walker’s plans were less developed, and included more of a focus on goal-setting. Notably, while the schools had plans to teach lessons about growth mindset and problem-solving, none of the plans made direct use of the lessons the DIDT/SIDT had previously discussed and piloted, as the SIDTs planned to break them up into smaller units.

Cross-cutting themes

Looking across the design process, three overall themes emerged. We refer to these themes as the tensions around concreteness, complexity, and collaboration. We also saw differences between schools in their ability to productively resolve these tensions.

Concreteness, complexity, and collaboration. The tension between concreteness,

complexity, and collaboration played out in several ways. First, members were most engaged when the design work focused on concrete practices they considered feasible in their particular context. This was evident in the differential engagement between groups reviewing potential innovation ideas that various groups deemed feasible/infeasible in their district context, as well as the differences in engagement between the growth mindset and problem-solving discussions. It was also evident in the high engagement when the SIDTs were focused on developing their school implementation plans. In short, engagement was high when members were focused on concrete practices they could see implementing with students the next day (or a pre-identified date). Engagement was lower when the conversation focused on more abstract ideas or on practices they considered unaligned with their current context.

The need for concreteness was important as reform designs do need to provide sufficient specificity to guide action. But in the context of a network-based approach to improvement, it also created challenges. That each school, even within the same district, faced a different context highlights the tension with collaborative, network-based design as the school-focused nature of those discussions pulled them away from practices developed by the larger group. For example, despite the very positive experience reviewing and then piloting the growth mindset lesson that was meant to be common across schools, the lessons themselves were not included in implementation plans as written. The lessons developed by the facilitators provided much needed concreteness to their work that was welcomed by participants. Yet SIDTs still needed to adapt them to their specific school context, such as how much time (and over how many days) they had to deliver the lessons.

A second way in which this tension appeared was in developing a sense of ownership and commitment in the collaborative design process by all members while also trying to make

concrete decisions. The successes in engaging DIDT/SIDT members in ways that built commitment often also appeared to limit their ability to adequately achieve other design needs—such as the appropriate level of concreteness. This was evident in Phase 1 when members attempted to use a consensus process to move from brainstorming to decisions about the innovation. As a group, they agreed with the decisions, but also were concerned that their decisions were not concrete enough for teachers to implement. In Phase 2 when SIDTs were developing specific implementation plans for their schools, members made concrete decisions in ways that led the schools to diverge from each other. DIDT members saw the need for a common design emphasis, yet struggled with how to give SIDTs ownership over their school's design without mandating certain practices. The response was to keep the design at a more abstract, less concrete level to continue to foster authentic engagement in the collaborative process.

Finally, the underlying complexity of the innovation ideas further created challenges in sustaining the collaborative design process. The focused discussions on growth mindset and problem-solving revealed that not everyone had a deep understanding of these ideas. Further, discussions of the ideas without a focus on the concrete ways they would introduce them to students led to disengagement, as was evident in the activity designed to build a common understanding of problem-solving. The need for concreteness also limited their ability to discuss more abstract ideas such as how complex ideas fit together to have a cohesive whole. This tension of concreteness and complexity arose most clearly when the DIDT had the contentious discussion and majority vote at the end of Phase 1 to drop goal-setting and self-monitoring to focus on growth mindset and problem-solving. At this point, the ideas outlined in the innovation were still at an abstract level and members felt pressure to present a feasible innovation to SIDT members without overwhelming them. Members' desire for concrete practices focused their

attention on the specific implementation supports they had available (i.e., 30 minutes every six weeks) and how they would fill this time. Reducing the number of ideas on which to focus, rather than focusing on the inherent complexity of their ideas, thus made sense.

Role of existing school context and capacity. A networked-based approach to improvement needs to contend with collaboration both across and within schools. In addition to the challenges of maintaining network wide collaboration and consensus as described above, we also saw patterns in each school's ability to resolve these dilemmas within the specific context of their school. Without having to manage differences in context across schools, in some ways the task of resolving concreteness, complexity, and consensus was simpler. Yet we also saw how individual member contexts, such as subject area or assignment, shaped reactions to the emerging innovation. This was evident in the tense disagreement in Wheatley around the behavioral reflection form and subject-area disagreements in Walker.

Further, the ability of schools to develop concrete practices for implementation in a collaborative process was greatly related to pre-existing conditions in each school, such as prior experience with teacher-led initiatives and a strong culture of trust and support. These pre-existing conditions facilitated or impeded the school members in moving from the abstract to the concrete while negotiating with each other to reach consensus. At Wheatley, the SIDT members were part of a school-wide teacher-led initiative prior to the innovation. Many of them had leadership roles in the initial process of gathering information for implementation and subsequent support. At Cervantes, there was an established culture of trust and support among the faculty and administration, particularly those on the SIDT. During both phases, the school administration had been very supportive of the SIDT's work. In contrast, there was a lack of prior experience and a limited culture of trust and support at Walker. Additionally, there was also

no clear leader on the Walker team. The individual who often tried to assume that role, while technically capable, had a position in the school that led others to discount their perspective. As a result, there was often a lack of consensus from the Walker team and it was difficult for them to agree on a plan of action and develop concrete practices. At the end of Phase 2, one SIDT member from Walker disagreed so strongly that they removed themself from the discussion.

In contrast, Wheatley's SIDT had a professional development plan and outlined roles for how they would support each other during the lesson development. They had also finalized their plan to incorporate their literacy technique with growth mindset to begin the next school year. At Cervantes, the SIDT members had gathered feedback from their faculty about what they thought their students needed to be have ownership and responsibility of their learning and incorporated the feedback when they created a sequence of lessons for each content area. With the support of the administration, they were able to secure time and resources for the faculty to come to a retreat to develop the lessons as a group before school started.

Conclusion and discussion

This case study of a collaborative design process in a research-practice partnership illustrates the complexities of the co-construction of an innovation between researchers, practitioners, and other stakeholders and highlights the need for a delicate balance between specificity of the design and collaboration. We showed how a collaborative process fostered high engagement as researchers and practitioners co-constructed the innovation, and how the team struggled to define the core strategies in sufficient detail to allow for implementation planning in a way that maintained the co-constructed design. There appeared to be a tension between achieving the necessary concreteness or specificity in the innovation design that would be implemented across contexts and a process that valued local ownership, collaboration and

consensus. This tension is an enduring dilemma for organizations and leaders in their quest to establish an innovative environment where there is room for independent thinking as well as supports for negotiations to challenge established norms and practices that can set guidelines and practices to achieve specific goals (Cohen-Vogel et al. 2018; Rosing, Frese, and Bausch 2011). Without specificity in the language and implementation supports necessary for the innovation, members could reach consensus, but this resulted in agreement on vague statements that lacked the appropriate concreteness needed for implementation. While exploring the subsequent implementation of the innovation is beyond the scope of this paper, prior research would suggest that this abstraction in the innovation's core design emphasis will create challenges for implementation (Shiffman et al. 2008). With less concrete and more abstract guidance for implementers, both teachers and administrators are likely to struggle to understand what is expected of them and thus little change in actual practice (Rowan et al. 2009; Sanders 2014). Despite adopting this approach to build buy-in, the result may lessen buy-in as implementers experience frustration (Nunnery et al. 1997). Our findings suggest teachers want input into the process but do not want ambiguity about what they are doing. Developing an innovation design that allows local input and adaptation to context without burdening educators with even more demands is a delicate balance (Nunnery et al. 1997). Moreover, there appears to be a need for flexibility at different time periods to sometimes emphasize innovation and at other time to emphasize implementation (Rosing, Frese, and Bausch 2011).

Because the innovation remained at an abstract level, there was significant room for design team members to make their own sense of the innovation. When individuals are confronted with unfamiliar ideas, they engage in sensemaking to integrate new ideas with existing understandings, which in turn influences their willingness to implement and adhere to

programs with fidelity (Coburn 2006; Waterman 1992). With little to guide that sensemaking process, it is likely that members attached the ideas to their own various existing practices (Spillane, Reiser, and Reimer 2002). Changing practice in ways intended by a particular reform requires specificity in design to provide clear guidance of implementation at all levels (Desimone 2002; Supovitz 2008). Districts can play a key role in shaping the sensemaking process in schools by establishing clear expectations for enacting the reform and how the reform fits in a coherent organizational context (Sanders 2014; Supovitz 2008). Organizational learning across levels in a district can be embedded and supported when there are specific tools or routines that can serve to reify the knowledge gained, enact it in practice, and allow members to see the continuity and progression of the work (Knapp 2008; Louis 2008; Redding, Cannata, and Miller 2018; Stein and Coburn 2008).

Moreover, as in other reforms taken to scale, decisions made in regards to one design feature had implications for how the others were enacted (Shiffman et al. 2008). The process of deciding on the central features of the innovation (design emphasis) while maximizing collaboration and involvement of local stakeholders (innovation engagement) also contributed to the abstractness (complexity). Indeed, these findings reinforce the idea that design is not just about a discrete practice but is a puzzle that encompasses relationships with the schools and the infrastructure built to support the work (Cohen et al. 2013). This implies that successful design must proceed with implementation in mind. That innovation development cannot be separated from implementation concerns is recognized in the increased attention to partnership-based approaches to reform such as design-based implementation research and improvement science. For example, design-based implementation research is distinguished by its attention to both the reform design itself and to the process of implementation (Fishman et al. 2013). At the same

time, engaging in improvement science requires practitioners to adopt radically new ways of selecting and implementing reforms (Bryk et al. 2015). Practitioners not only need expertise in the technical aspects of the reform necessary for implementation, but also expertise in design and development, space for such work to take place, and latitude to fail in a safe environment.

Lewis (2015) describes this interconnection between the innovation and its implementation in context as recognizing that the knowledge required for success resides in both the people (i.e., local implementers) and the program (i.e., the innovation). Yet our finding that pre-existing school-level conditions appeared to strongly influence their capacity to engage in the co-design process raises concerns for engaging in this type of partnership work in contexts that may lack such expertise in designing and leading reform (Durlak and Dupre 2008). Partnership-based models of improvement, with their reliance on local expertise and engagement, must grapple with a critical paradox of school improvement: it takes capacity to build capacity (Cohen et al. 2013; Hatch 2002; Hatch and White 2002). Improvement efforts that intentionally build in substantial local adaptation require capacities such as time, expertise, and collaborative ability to engage in the development work that teachers may not have, particularly since improvement efforts are often focused on schools with limited existing capacity (Berends, Bodilly, and Kirby 2002; Datnow et al. 1998). Indeed, the realization of policy in practice depends on the fit between the capabilities of those that support implementation and the ambitions of the policy (Cohen, Moffitt, and Goldin 2007). Finding this balance between establishing concreteness and fostering collaboration and consensus while taking into account pre-existing local conditions is a key challenge in school improvement.

References

Berends, Mark, Susan Bodilly, and Sheila Nataraj Kirby. 2002. *Facing the Challenges of Whole-School Reform: New American Schools after a Decade*. MR-1498-EDU. Santa Monica, CA: RAND.

Bodilly, Susan. 1996. *Lessons from New American Schools Development Corporation's Demonstration Phase.* RAND. http://eric.ed.gov/?id=ED396377.

Bodilly, Susan, T. K. Glennan, K. A. Kerr, and J. R. Galegher. 2004. "Framing the Problem." *Expanding the Reach of Education Reforms: Perspectives from Leaders in the Scale-up of Educational Interventions. Santa Monica, CA: Rand Corporation*.

Bryk, Anthony S., Louis Gomez, and Alicia Grunow. 2011. "Getting Ideas into Action: Building Networked Improvement Communities in Education." In *Frontiers in Sociology of Education*, edited by Maureen T. Hallinan, 127–62. United States: Springer Publishing.

Bryk, Anthony S., Louis M. Gomez, Alicia Grunow, and Paul G. LeMahieu. 2015. *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Cambridge, MA: Harvard Education Press.

Camburn, Eric M. 2010. "Embedded Teacher Learning Opportunities as a Site for Reflective Practice: An Exploratory Study." *American Journal of Education* 116 (4): 463–89.

Cannata, Marisa, Lora Cohen-Vogel, and Michael Sorum. 2017. "Partnering for Improvement: Improvement Communities and Their Role in Scale-Up." *Peabody Journal of Education* 92 (5): 569–88.

Cannata, Marisa, Chris Redding, and Tuan Nguyen. in press. "Building Student Ownership and Responsibility: Examing Student Outcomes from a Research-Practice Partnership." *Journal of Research on Educational Effectiveness*.

Cannata, Marisa, and Stacey A. Rutledge. 2017. "Introduction to New Frontiers in Scaling Up

Research." *Peabody Journal of Education* 92 (5): 559–68.

Cannata, Marisa, Thomas M. Smith, and Katherine Taylor Haynes. 2017. "Integrating Academic Press and Support by Increasing Student Ownership and Responsibility." *AERA Open* 3 (3): 1–13.

Coburn, Cynthia E. 2006. "Framing the Problem of Reading Instruction: Using Frame Analysis to Uncover the Microprocesses of Policy Implementation." *American Educational Research Journal* 43 (3): 343–349.

Coburn, Cynthia E., and William R. Penuel. 2016. "Research–Practice Partnerships in Education Outcomes, Dynamics, and Open Questions." *Educational Researcher* 45 (1): 48–54.

Coburn, Cynthia E., Jennifer L. Russell, Julia Heath Kaufman, and Mary Kay Stein. 2012. "Supporting Sustainability: Teachers' Advice Networks and Ambitious Instructional Reform." *American Journal of Education* 119 (1): 137–82.

Coburn, Cynthia E., Mary Kay Stein, Juliet Baxter, Laura D'Amico, Amanda Datnow, Randi Engle, Meredith Honig, et al. 2010. *Research and Practice in Education: Building Alliances, Bridging the Divide*. Lanham, Md: Rowman & Littlefield Publishers.

Coburn, Cynthia E., William R. Penuel, and Kimberly E. Geil. 2013. "Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts." New York, NY: William T. Grant Foundation.

Cohen, David K., and Deborah Loewenberg Ball. 1999. "Instruction, Capacity, and Improvement." Philadelphia: Consortium for Policy Research in Education.

Cohen, David K., Donald J. Peurach, Joshua L. Glazer, Karen E. Gates, and Simona Goldin. 2013. *Improvement by Design: The Promise of Better Schools*. Chicago ; London: University Of Chicago Press.

Cohen, David K., Susan L. Moffitt, and Simona Goldin. 2007. "Policy and Practice: The Dilemma."

American Journal of Education 113 (4): 515–48. https://doi.org/10.1086/518487.

Cohen-Vogel, Lora, Danielle Allen, Stacey Rutledge, Marisa Cannata, Christopher Harrison, and Thomas Smith. 2018. "Organizing for School Improvement: The Dilemmas of Research-Practice Partnerships." *Journal of Research on Organization in Education* 2 (1): 1–24.

Cohen-Vogel, Lora, Marisa Cannata, Stacey Rutledge, and Allison Rose Socol. 2016. "A Model of Continuous Improvement in High Schools: A Process for Research, Innovation Design, Implementation, and Scale." *Teachers College Record* 116 (13): 1–26.

Cohen-Vogel, Lora, Ariel Tichnor-Wagner, Danielle Allen, Christopher Harrison, Kirsten Kainz, Allison Rose Socol, and Qi Wang. 2015. "Implementing Educational Innovations at Scale: Transforming Researchers Into Continuous Improvement Scientists." *Educational Policy* 29 (1): 257–77.

Datnow, Amanda, Lea Hubbard, and Hugh Mehan. 2002. *Extending Educational Reform: From One School to Many*. 1st ed. Routledge.

Datnow, Amanda, Barbara McHugh, Sam Stringfield, and Douglas J. Hacker. 1998. "Scaling Up the Core Knowledge Sequence." *Education and Urban Society* 30 (3): 409–32.

Datnow, Amanda, and Vicki Park. 2009. "Conceptualizing Policy Implementation: Large-Scale Reform in an Era of Complexity." In *Handbook of Education Policy Research*, edited by Gary Sykes, Barbara Schneider, and David N. Plank, 1 edition, 348–61. New York : Washington, D.C.: Routledge.

Dede, Chris, and James P. Honan. 2005. "Scaling up Success: A Synthesis of Themes and Insights." In *Scaling Up Success : Lessons Learned from Technology-Based Educational Improvement*, edited by Chris Dede, James P. Honan, and Laurence C. Peters, 1st ed. Jossey-Bass.

Desimone, Laura. 2002. "How Can Comprehensive School Reform Models Be Successfully

Implemented?" *Review of Educational Research* 72 (3): 433–79.

Devos, Geert, Melissa Tuytens, and Hester Hulpia. 2014. "Teachers Organizational Commitment: Examining the Mediating Effects of Distributed Leadership." *American Journal of Education* 120 (2): 205–31.

Durlak, Joseph A., and Emily P. Dupre. 2008. "Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation." *Am J Community Psychol* 41: 327–50.

Elmore, Richard F. 2016. "'Getting to Scale…' It Seemed like a Good Idea at the Time." *Journal of Educational Change* 17 (4): 529–37. https://doi.org/10.1007/s10833-016-9290-8.

Engelbart, Douglas C. 1992. "Toward High-Performance Organizations: A Strategic Role for Groupware." San Jose, CA: Bootstrap Institute. http://www.dougengelbart.org/pubs/augment-132811.html#3.

Fishman, Barry J., William R. Penuel, Anna-Ruth Allen, Britte Haugan Cheng, and Nora Sabelli. 2013. "Design-Based Implementation Research: An Emerging Model for Transforming the Relationship of Research and Practice." In *Design-Based Implementation Research: Theories, Methods, and Exemplars*, edited by Barry J. Fishman, William R. Penuel, Anna-Ruth Allen, and Britte Haugen Cheng, 136–56. New York: Teachers College, Columbia University.

Fullan, Michael. 2016. "The Elusive Nature of Whole System Improvement in Education." *Journal of Educational Change* 17 (4): 539–44. https://doi.org/10.1007/s10833-016-9289-1.

Gallagher, Mary Jean, John Malloy, and Rachel Ryerson. 2016. "Achieving Excellence: Bringing Effective Literacy Pedagogy to Scale in Ontario's Publicly-Funded Education System." *Journal of Educational Change* 17 (4): 477–504. https://doi.org/10.1007/s10833-016-9284-6.

Glennan, Thomas K., Susan J Bodilly, J. R. Galegher, and K. A. Kerr. 2004. *Expanding the Reach of*

Education Reforms: Perspectives from Leaders in the Scale-Up of Educational Interventions. 1st ed. Rand Publishing.

Hatch, Thomas. 2002. "When Improvement Programs Collide." *The Phi Delta Kappan* 83 (8): 626–39.

Hatch, Thomas, and Noel White. 2002. "The Raw Materials of Reform: Rethinking the Knowledge of School Improvement." *Journal of Educational Change* 3 (2): 117–34.

Knapp, Michael S. 2008. "How Can Organizational and Sociocultural Learning Theories Shed Light on District Instructional Reform?" *American Journal of Education* 114 (4): 521–39.

LeMahieu, Paul G., Alicia Grunow, Laura Baker, Lee E. Nordstrum, and Louis M. Gomez. 2017. "Networked Improvement Communities: The Discipline of Improvement Science Meets the Power of Networks." *Quality Assurance in Education*, January.

Lewis, Catherine. 2015. "What Is Improvement Science? Do We Need It in Education?" *Educational Researcher* 44 (1): 54–61. https://doi.org/10.3102/0013189X15570388.

Louis, Karen Seashore. 2008. "Learning to Support Improvement—Next Steps for Research on District Practice." *American Journal of Education* 114 (4): 681–89. https://doi.org/10.1086/589320.

McLaughlin, Milbrey Wallin. 1987. "Learning from Experience: Lessons from Policy Implementation." *Educational Evaluation and Policy Analysis* 9 (2): 171. https://doi.org/10.2307/1163728.

Niesz, Tricia, and Kelli Ryan. 2018. "Teacher Ownership versus Scaling up System-Wide Educational Change: The Case of Activity Based Learning in South India." *Educational Research for Policy and Practice*, August. https://doi.org/10.1007/s10671-018-9232-8.

Nunnery, John A. 1998. "Reform Ideology and the Locus of Development Problem in Educational Restructuring: Enduring Lessons from Studies of Educational Innovation." *Education and Urban Society* 30 (3): 277–295.

Nunnery, John A., Linda Bol, Amy Dietrich, Lyndall Rich, Susan Kelly, Douglas Hacker, and Allan

Sterbin. 1997. “Teachers’ Initial Reactions to Their Pre-Implementation Preparation and Early Restructuring Experiences.” *School Effectiveness and School Improvement* 8 (1): 72–94.

Patton, Michael Quinn. 2002. *Qualitative Research and Evaluation Methods*. 3rd edition. Thousand Oaks, Calif: SAGE Publications, Inc.

Payne, Charles M. 2008. *So Much Reform, So Little Change: The Persistence of Failure in Urban Schools*. Third Printing, 2010 edition. Cambridge, Mass: Harvard Education Press.

Penuel, William R., Barry J. Fishman, Britte Haugan Cheng, and Nora Sabelli. 2011. “Organizing Research and Development at the Intersection of Learning, Implementation, and Design.” *Educational Researcher* 40 (7): 331–37. https://doi.org/10.3102/0013189X11421826.

Peurach, Donald J., and Joshua L. Glazer. 2012. “Reconsidering Replication: New Perspectives on Large-Scale School Improvement.” *Journal of Educational Change* 13 (2): 155–90.

Redding, Chris, Marisa Cannata, and Jason Miller. 2018. “System Learning in an Urban School District: A Case Study of Intra-District Learning.” *Journal of Educational Change* 19 (1): 77–101.

Rincón-Gallardo, Santiago, and Richard Elmore. 2012. “Transforming Teaching and Learning Through Social Movement in Mexican Public Middle Schools.” *Harvard Educational Review* 82 (4): 471–90. https://doi.org/10.17763/haer.82.4.46751717307t4j90.

Rosing, Kathrin, Michael Frese, and Andreas Bausch. 2011. “Explaining the Heterogeneity of the Leadership-Innovation Relationship: Ambidextrous Leadership.” *The Leadership Quarterly* 22 (October): 956–74. https://doi.org/10.1016/j.leaqua.2011.07.014.

Rowan, B., R.J. Correnti, R.J. Miller, and E.M. Camburn. 2009. “School Improvement by Design: Lessons from a Study of Comprehensive School Reform Programs.” In *AERA Handbook on Education Policy Research*, 637–51. New York, NY: Routledge.

Russell, Jennifer Lin, Anthony S. Bryk, Jonathan R. Dolle, Louis M. Gomez, Paul G. LeMahieu, and

Alicia Grunow. 2017. “A Framework for the Initiation of Networked Improvement Communities.” *Teachers College Record* 119 (7).

Sabelli, Nora, and Christopher J. Harris. 2015. “The Role of Innovation in Scaling Up Educational Innovations.” In *Scaling Educational Innovations*, edited by Chee-Kit Looi and Laik Woon Teh, 13–30. Singapore: Springer.

Sanders, Mavis G. 2014. “Principal Leadership for School, Family, and Community Partnerships: The Role of a Systems Approach to Reform Implemetnation.” *American Journal of Education* 120 (2): 233–55.

Sawchuk, Stephen. 2017. “With Latest Education Investments, Gates Pivots Again - Education Week.” *Education Week*, November 1, 2017. https://www.edweek.org/ew/articles/2017/10/23/with-latest-education-investments-gates-pivots-again.html.

Shiffman, Catherine Dunn, Matthew Riggan, Diane Massell, Matthew Goldwasser, and Joy Anderson. 2008. “Channeling Adaptation: The Role of Design in Enactment Patterns.” In *The Implementation Gap: Understanding Reform in High Schools*, edited by Jonathan A Supovitz and Elliot H. Weinbaum, 46–67. Columbia University, New York: Teachers College.

Spillane, James P., Brian J. Reiser, and Todd Reimer. 2002. “Policy Implementation and Cognition: Reframing and Refocusing Implementation Research.” *Review of Educational Research* 72 (3): 387–431. https://doi.org/10.3102/00346543072003387.

Stein, Mary Kay, and Cynthia E. Coburn. 2008. “Architectures for Learning: A Comparative Analysis of Two Urban School Districts.” *American Journal of Education* 114 (4): 583–626.

Stringfield, Sam, and Amanda Datnow. 1998. “Scaling Up School Restructuring Designs in Urban Schools.” *Education and Urban Society* 30 (3): 269–76.

Supovitz, Jonathan A. 2008. “Implementation as Iterative Refraction.” In *Implementation Gap:*

Understanding Reform in High Schools, edited by Jonathan A. Supovitz and Elliot H. Weinbaum, 151–72. New York, NY: Teachers College Press.

Tett, Lyn, Jim Crowther, and Paul O'Hara. 2003. "Collaborative Partnerships in Community Education." *Journal of Education Policy* 18 (1): 37–51.

Waterman, Robert H. 1992. *Adhocracy : The Power to Change*. 1st Norton ed. New York: W.W. Norton.

Watson, Cate, and Maureen K. Michael. 2016. "Translations of Policy and Shifting Demands of Teacher Professionalism: From CPD to Professional Learning." *Journal of Education Policy* 31 (3): 259–74. https://doi.org/10.1080/02680939.2015.1092053.

Zavadsky, Heather. 2016. "Bringing Effective Instructional Practice to Scale in American Schools: Lessons from the Long Beach Unified School District." *Journal of Educational Change* 17 (4): 505–27. https://doi.org/10.1007/s10833-016-9287-3.

Tables

Table 1 – Descriptive Information on Innovation Schools

	Wheatley	Cervantes	Walker
Enrollment	>1500	700-1200	>1500
Student race/ethnicity			
Hispanic	40-60%	>80%	>80%
African American	20-40%	<20%	<20%
White	20-40%	<20%	<20%
Percent economically disadvantaged	40-60%	>80%	>80%
Recent reform history	Teacher leadership team successfully designed and implemented a school-wide literacy initiative; New principal appointed at the start of Phase 2	School-wide literacy initiative was successfully implemented; New principal appointed at the start of Phase 1	Target of school turnaround efforts a few years prior to participating in this work; New principal appointed at the start of Phase 1
DIDT representatives	Two teachers selected by the principal who were members of the existing teacher leadership team	One teacher who was identified as a leader during the literacy initiative implementation; One teacher whose subject assignment was considered relevant for SOAR; both selected by principal	One non-classroom teacher selected by principal to minimize instructional disruption; two classroom teachers selected by principal at facilitator encouragement to appoint additional personnel
SIDT composition	Six teachers, most of whom were members of the existing teacher leadership team and one assistant principal; recruited by DIDT representatives	Six teachers recruited by DIDT representatives because of perceived interest as early adopters	Eight department chairs selected by principal due to their role on school leadership team

Source: District administrative data, 2012-2013 school year.

Table 2 – Coding Framework for Capacity Building and Innovation Design

Attitudes and engagement	• Attendance • Attitudes • Engagement
Delivery of Learning for Design and Implementation	• Learning about design challenge • Learning about design process • Learning about implementation and scale • Learning about continuous improvement*
Participant Understanding of Design and Implementation	• Understanding of design challenge • Understanding of design process • Understanding of implementation and scale • Understanding of continuous improvement*
Design Process	• Collaborative • Openness to new ideas • Needs-centered • Grounded in design challenge • Alignment with existing system components • Iterating on the design* • Piloting and PDSA*
Design Concept Itself	• Incorporation of design challenge core elements • Participants' perceptions of developed innovation
Understanding the Process	• Centrality of the capacity building framework • Emphasis of the DIDT as a district-wide structure • Integration of research team • Focus on relationship building with schools • Points of significant concern • Significant decision points

*These elements were added in Phase 2.